

T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
BİLGİ VE BELGE YÖNETİMİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

1 NUMARALI ERZURUM AHKÂM DEFTERİNİN
(M.1742 – 1749) METİN MADENCİLİĞİ
TEKNIĞIYLA DEĞERLENDİRİLMESİ

Murat ESRİNGÜ

2501180047

TEZ DANIŞMANI

PROF. DR. İshak KESKİN

İSTANBUL – 2022

ÖZ

1 NUMARALI ERZURUM AHKÂM DEFTERİNİN (M.1742–1749) METİN MADENCİLİĞİ TEKNİĞİYLE DEĞERLENDİRİLMESİ MURAT ESRİNGÜ

Teknolojik gelişmelere bağlı olarak büyük miktarda veri oluşmaya başlamıştır. Söz konusu veriler veri madenciliği ve metin madenciliği yöntemleri kullanılarak analiz edilebilmekte ve ayrıca bu verilerden faydalı bilgiler çıkarılabilmektedir. Veri madenciliği yapılandırılmış veriler üzerinde uygulanırken metin madenciliği yarı yapılandırılmış ya da yapılandırılmamış veriler üzerinde uygulanmaktadır.

Arşivlere ya da veri tabanlarına bakıldığında buralarda depolanan veri miktarının çok yüksek olduğu görülebilir. Bu bağlamda Osmanlı arşiv belgeleri de önemli bir veri kaynağı olarak kabul edilebilir. Bu verilerin çoğu yapılandırılmamış ya da yarı yapılandırılmış metinsel veri formundadır. Büyük miktardaki metinsel verilerin manuel olarak düzenlenmesi, analiz edilmesi ve dolayısıyla faydalı bilgilerin elde edilmesi oldukça zor ve zaman alıcı bir işlemdir.

Bu çalışmada, 1 Numaralı Erzurum Ahkâm Defterinin içerisinde bulunan metinsel hüküm verileri, metin madenciliği yöntemleri uygulanarak, otomatik bir şekilde konularına göre sınıflandırılmıştır. Sınıflandırma işleminde Decision Tree, Support Vector Machine, Gradient Boosted Trees, Random Forest ve K Nearest Neighbor algoritmaları kullanılmıştır. Kullanılan doküman frekansı, ikili ağırlıklandırma, terim frekansı, ters doküman frekansı ve ters sınıf frekansı ağırlıklandırma yöntemlerine göre algoritmaların sınıflandırma başarıları değerlendirilmiştir. En başarılı sonuç %0,812 doğruluk oranıyla ters doküman frekansı ağırlıklandırma yöntemi kullanılarak Gradient Boosted ağaçları algoritmasıyla elde edilmiştir. Oluşturulan sınıflandırma modelinin geliştirilebilmesi ve Osmanlı arşiv belgeleri üzerinde uygulanacak metin madenciliği uygulamaları için önerilerde bulunulmuştur.

Anahtar Kelimeler: Metin Madenciliği, Ahkâm Defterleri, Erzurum Ahkâm Defterleri, Metin Sınıflandırma, KNIME.

ABSTRACT

EVALUATION OF ERZURUM AHKÂM DEFTER (M.1742–1749) NUMBERED 1 WITH THE TEXT MINING TECHNIQUE

MURAT ESRİNGÜ

Depending on the technological developments, a large amount of data has begun to be formed. The data in question can be analyzed using data mining and text mining methods, and useful information can also be extracted from these data. While data mining is applied on structured data, text mining is applied on semi-structured or unstructured data.

When we look at the archives or databases, it can be seen that the amount of data stored in them is very high. In this context, Ottoman archival documents can also be considered as an important data source. Most of this data is in the form of unstructured or semi-structured textual data. The process of manually arranging and analyzing large amounts of textual data, and therefore obtaining useful information, is a rather difficult and time-consuming process.

In this study, the textual judgment data contained in the Erzurum Ahkâm Defter numbered 1 were automatically classified according to their subjects by applying text mining methods. Decision Tree, Support Vector Machine, Gradient Boosted Trees, Random Forest and K Nearest Neighbor algorithms were used in the classification process. The classification achievements of the algorithms were evaluated according to the document frequency, binary weighting, term frequency, inverse document frequency and inverse class frequency weighting methods used. The most successful result was obtained from the Gradient Boosted trees algorithm according to the inverse document frequency weighting method with an accuracy rate of %0,812. Suggestions were made for the development of the classification model created and for the text mining applications to be applied on the Ottoman archive documents.

Keywords: Text Mining, Ahkâm Defterleri, Erzurum Ahkâm Defterleri, Text Classification, KNIME.

ÖNSÖZ

Günümüzde teknolojik gelişmelere bağlı olarak büyük miktarda veri üretilmektedir. Meydana gelen verinin önemli bir kısmı ise metin formatındadır. Kurum, şirket ve kişiler gerçekleştirilecek iş ve işlemlerde, alınacak kararların belirlenmesinde ya da var olan bir durumun daha iyi anlaşılmasında belirleyici olabilecek analizlerin ve faydalı bilgilerin elde edilebilmesi için sahip oldukları veri üzerinde yoğunlaşmışlardır. Bu durum ise veri madenciliği ve metin madenciliği uygulamalarını oldukça önemli kılmıştır.

Metinsel veri üzerinde uygulanan metin madenciliği işlemleri ile veri yapılandırılmış hale getirilir. Böylece hedeflenen amaç doğrultusunda analizlerin ve faydalı bilgi elde etme işlemlerinin yapılabilmesi mümkün olmaktadır. Söz konusu analizlerin yapılabilmesi ve faydalı bilgilerin elde edilmesi işlemlerini mümkün kılan metin madenciliği işlemlerinden biri de sınıflandırma işlemidir. Metinsel verilerin manuel olarak sınıflandırılması oldukça zor ve zaman alıcı bir süreçtir. Ayrıca sınıflandırma işleminin sonuçları, söz konusu sınıflandırma işlemi gerçekleştiren kişi ya da kişilerin kararlarına/tercihlerine göre değişiklik gösterebilmektedir. Bu durum sınıflandırma işleminde belirli bir standardın oluşmasını olumsuz şekilde etkilemektedir.

Hacim olarak büyük miktardaki Osmanlı arşiv belgeleri veri kaynağı olarak kabul edilebilir. Söz konusu belgeler tarih, sosyoloji ve hukuk alanları başta olmak üzere farklı birçok alanda yapılacak çalışmalar için önemli bir kaynak niteliği taşımaktadır. Bu sebeple Osmanlı arşiv belgelerinin metin madenciliği yöntemiyle ele alınması ve gerçekleştirilen analizler doğrultusunda faydalı bilgilerin elde edilmesi işlemi önem kazanmaktadır. Bu çalışmada Osmanlı arşiv belgelerinden 1 numaralı Erzurum Ahkâm Defteri ele alınmıştır ve bu defter içerisinde bulunan hüküm verisi yapılandırılmış hale dönüştürülerek konularına göre otomatik olarak sınıflandırılmıştır. Böylece oldukça zaman alıcı ve zorlu bir işlem olan sınıflandırmanın hızlı bir şekilde gerçekleştirilmesi, amaçlara göre analizlerin yapılabilmesi ve faydalı bilgilerin elde edilebilmesi mümkün kılınmıştır. Ayrıca oluşturulan sınıflandırma modelinin ve çalışmada karşılaşılan kısıtlılıklar ile bu

kısıtlılıklara verilen önerilerin, sonrasında Osmanlı arşiv belgeleri üzerinde gerçekleştirilecek olan metin madenciliği çalışmaları için ilham verici olabileceği düşünülmektedir.

Tez çalışmam süresince bana her türlü imkânı sağlayan ve desteklerini hiçbir zaman esirgemeyen değerli hocalarım Prof. Dr. Tuba Çavdar Karatepe ve Prof. Dr. Burçak Şentürk hocalarıma teşekkür ederim.

Tez çalışmamda kullandığım hüküm verilerinin konu başlıklarını yeniden belirlerken, değerli zamanını ayıran ve takıldığım her an bana yardımcı olan kıymetli Hocam Prof. Dr. Kemalettin Kuzucu'ya ve yoğun çalışma temposunun arasında yardımlarını benden esirgemeyen ve metin madenciliğine dair sorularımı hiçbir zaman cevapsız bırakmayan saygıdeğer Arş. Gör. M. Emin Gedikli'ye teşekkürü borç bilirim.

Son olarak vaktimin çoğunu tez çalışmama ayırmam sebebiyle önemli birçok anda yanında olamadığım ve kendisine yeterince zaman ayıramadığım kıymetli eşime desteklerinden ve anlayışından dolayı teşekkür ederim. Benden dualarını hiçbir zaman esirgemeyen anne ve babama da minnettarım.

İstanbul, 2022

Murat Esringü

İÇİNDEKİLER

ÖZ.....	ii
ABSTRACT.....	iii
ÖNSÖZ.....	iv
İÇİNDEKİLER	vi
TABLolar LİSTESİ.....	x
ŞEKİLLER LİSTESİ.....	xi
RESİMLER LİSTESİ.....	xii
DENKLEMLER LİSTESİ	xiv
KISALTMALAR	xv
GİRİŞ	1

BİRİNCİ BÖLÜM METİN MADENCİLİĞİ

1.1. Tanımı ve Kapsamı	6
1.2. Tarihçesi	8
1.3. Metin Madenciliğinin Uygulama Sahası.....	8
1.3.1. Bilgi Erişim	10
1.3.2. Bilgi Çıkarımı	12
1.3.3. Doğal Dil İşleme	13
1.3.4. Web Madenciliği.....	14
1.3.5. Kavram Çıkarımı.....	15
1.3.6. Doküman Kümeleme	15
1.3.7. Doküman Sınıflama	16
1.4. Metin Madenciliği Süreci.....	17
1.4.1. CRISP-DM Modeli	19
1.4.1.1. İş Anlama	20
1.4.1.2. Veriyi Anlama.....	20

1.4.1.3. Veriyi Hazırlama.....	21
1.4.1.4. Modelleme.....	21
1.4.1.5. Değerlendirme.....	22
1.4.1.6. Yayılım.....	23
1.4.2. Metin Koleksiyonu Oluşturma.....	23
1.4.3. Metin Ön İşleme.....	24
1.4.3.1. Dönüştürme ve Temizleme	25
1.4.3.2. İşaretleme	26
1.4.3.3. Durak Kelimelerinin Çıkarılması.....	27
1.4.3.4. Gövdeleme (Kök Bulma)	27
1.4.3.5. Terimlerin Ağırlıklandırılması	28
1.4.3.5.1. Bitsel Ağırlıklandırma.....	29
1.4.3.5.2. Frekansa Göre Ağırlıklandırma (Weighting by Frequency)	30
1.4.3.5.3. Terim Frekansı (Term Frequency-TF) ve Ters Doküman Frekansı (Inverse Document Frequency-IDF)	30
1.4.3.5.4. Ters Sınıf Frekansı (Inverse Class Frequency-ICF)	31
1.4.3.5.5. Normalizasyon (Term Frequency Component-TFC).....	31
1.4.3.5.6. Logaritmik Terim Bileşeni (Logarithmic Term Component-LTC)	31
1.4.4. Özellik Seçimi.....	32
1.4.5. Veri Madenciliği	33
1.4.6. Görselleştirme	33
1.4.7. Değerlendirme.....	33

İKİNCİ BÖLÜM AHKÂM DEFTERLERİ

2.1. Tanımı ve Tarihsel Gelişimi.....	35
2.2. Ahkâm Defterlerinin Önemi	40

2.3. Ahkâm Defterlerinin Diplomatik ve Fiziki Özellikleri	42
2.4. Ahkâm Defterlerinin İçeriği	45
2.5. Erzurum 1 Numaralı Ahkâm Defteri	48
2.5.1. Defterin Fiziki Yapısı.....	48
2.5.2. Defterin Diplomatik ve Teknik Özellikleri	52
2.5.3. Defterin Dil Özellikleri	53
2.5.4. Defterde Ele Alınan Konular	53
2.5.4.1. Asayiş	55
2.5.4.2. Sosyal, Ekonomik ve İdarî Konular	55
2.5.4.3. Arazi Meselesi ve Vergi.....	56

ÜÇÜNCÜ BÖLÜM

METİN MADENCİLİĞİ UYGULAMASI

3.1. Uygulamada Kullanılan Programlar	60
3.1.1 Microsoft Excel	60
3.1.2. KNIME.....	61
3.2. Veri Toplama Süreci	63
3.3. Verilerin Hazırlanması ve Veriyi Anlamlandırma	64
3.4. Veri Ön İşleme	72
3.4.1. Excel Reader (Excel Okuyucu).....	73
3.4.2. Strings To Document (Doküman Dizeleri)	74
3.4.3. Case Converter (Büyük-Küçük Harf Dönüştürücü).....	75
3.4.4. Punctuation Erasure (Noktalama İşaretlerini Silme)	76
3.4.5. Zemberek Stemmer (Zemberek Kök Bulucu).....	77
3.4.6. Number Filter (Sayı Filtresi)	78
3.4.7. Stop Word Filter (Durak Kelime Filtresi)	79
3.4.8. Dictionary Replacer (Sözlük Değiştirici).....	80

3.4.9. Table Creator (Tablo Oluşturucu).....	81
3.4.10. N Chars Filter (N Karakter Filtresi).....	82
3.4.11. Dictionary Tagger (Sözlük Etiketleyici)	83
3.4.12. Bag of Words Creator (Kelime Çantası Oluşturma).....	83
3.4.13. DF (Document Frequency, Doküman Frekansı).....	84
3.4.14. Frequency Filter (Frekans Filtresi).....	85
3.5. Sınıflandırma İşlemi.....	91
3.5.1. Hükümlerin Sayısal Olarak İfade Edilmesi.....	91
3.5.2. Decision Tree Learner (Karar Ağacı Öğrenicisi).....	93
3.5.3. Decision Tree Predictor (Karar Ağacı Tahmincisi)	95
3.5.4. Support Vector Machine (SVM) Learner (Destek Vektör Makinesi Öğrenicisi).....	96
3.5.5. Support Vector Machine (SVM) Predictor (Destek Vektör Makinesi Tahmincisi)	97
3.5.6. Gradient Boosted Trees Learner (Gradyan Artırma Ağaçları Öğrenicisi).....	97
3.5.7. Gradient Boosted Trees Predictor (Gradyan Artırma Ağaçları Tahmincisi)	98
3.5.8. Random Forest Learner (Rastgele Orman Öğrenicisi).....	98
3.5.9. Random Forest Predictor (Rastgele Orman Tahmincisi).....	99
3.5.10. K Nearest Neighbor (K En Yakın Komşu)	100
3.5.11. Scorer (Performans Hesaplayıcı)	101
3.6. Sınıflandırma Sonuçları	103
SONUÇ VE ÖNERİLER.....	110
KAYNAKÇA	119
EKLER.....	138

TABLÖLAR LİSTESİ

Tablo 2. 1: Ahkâm Defterlerinin Eyaletlere Göre Listesi.	45
Tablo 2. 2: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Aylara Göre Yüzdesel Dağılımı.....	49
Tablo 2. 3: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Aylara ve Yıllara Göre Dağılımı.....	50
Tablo 2. 4: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Genel Konulara Göre Dağılımı.....	54
Tablo 2. 5: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Konulara Göre Detaylı Dağılımı.....	58
Tablo 3. 1: Hükümlerin Düzenlenmiş Konularına Göre Dağılımı.....	66
Tablo 3. 2: En Çok Kullanılan 50 Kelime ve Sıklıkları.	87
Tablo 3. 3: Hükümlerde En Çok Geçen Kişi İsimleri.	89
Tablo 3. 4: Hükümlerde Geçen Şehir İsimleri ve Sıklıkları.....	90
Tablo 3. 5: Sözlük	92
Tablo 3. 6: Hüküm-Kelime Eşleşmesi.	92
Tablo 3. 7: Karışıklık Matrisi.....	103
Tablo 3. 8: DF Ağırlıklandırma Yöntemiyle Sınıflandırma Algoritmalarının Başarısının İncelenmesi.	105
Tablo 3. 9: Binary Ağırlıklandırma Yöntemiyle Sınıflandırma Algoritmalarının Başarısının İncelenmesi.	106
Tablo 3. 10: TF Ağırlıklandırma Yönteminde Sınıflandırma Algoritmalarının Başarısının İncelenmesi.	106
Tablo 3. 11: IDF Ağırlıklandırma Yönteminde Sınıflandırma Algoritmalarının Başarısının İncelenmesi.	107
Tablo 3. 12: ICF Ağırlıklandırma Yönteminde Sınıflandırma Algoritmalarının Başarısının İncelenmesi.	107
Tablo 3. 13: Gradient Boosted Ağaçları Algoritması Sonucunda Elde Edilen Karışıklık Matrisi (n=362).	108

ŞEKİLLER LİSTESİ

Şekil 1. 1. Metin Madenciliğinin Uygulama Sahaları ve İlişkili Olduğu Alanlar.	10
Şekil 1. 2. Metin Madenciliği Süreci.	18
Şekil 1. 3. Veri Madenciliği için Çapraz Endüstri Standart Süreci.	20
Şekil 1. 4. Metin Madenciliği Süreçleri.	22
Şekil 3. 1: Yıllara Göre Hüküm Konularının Frekans Dağılımı.....	69
Şekil 3. 2: Yıllara Göre Hüküm Sayılarının Değişimine Ait Çizgi Grafiği.....	70
Şekil 3. 3: Aylara Göre Hüküm Konularının Sayıları.	71
Şekil 3. 4: Aylara Göre Hüküm Sayılarının Değişimi.	72
Şekil 3. 5: Karar Ağacı Örneği.	95

RESİMLER LİSTESİ

Resim 2. 1: Erzurum 1 Numaralı Ahkâm Defteri'nin Kapağı.....	48
Resim 2. 2: 1 Numaralı Erzurum Ahkâm Defterinin Girişi.	51
Resim 3. 1: Excel Reader Düğümü.....	73
Resim 3. 2: Excel Reader Düğümüne Aktarılan Hüküm Verileri.	73
Resim 3. 3: String To Document Düğümü.	74
Resim 3. 4: Dokümana Çevrilen Metin Verileri.	74
Resim 3. 5: Case Converter Düğümü.	75
Resim 3. 6: Küçük Harfe Dönüştürülen Hüküm Verileri.....	75
Resim 3. 7: Punctuation Erasure Düğümü.	76
Resim 3. 8: Noktalama İşaretlerinin Kaldırıldığı Hüküm Verileri.....	76
Resim 3. 9: Zemberek Stemmer Düğümü.	77
Resim 3. 10: Zemberek Stemmer ile Köklerine Ayrılan Kelimeler.....	77
Resim 3. 11: Number Filter Düğümü.	78
Resim 3. 12: Stop Word Filter Düğümü.....	79
Resim 3. 13: Durak Kelimelerin Hüküm Verilerinden Çıkarılması.....	79
Resim 3. 14: Dictionary Replacer Düğümü.	81
Resim 3. 15: Dictionary Replacer Düğümü İle Değiştirilmiş Kelime Verileri.	81
Resim 3. 16: Table Creator Düğümü.....	81
Resim 3. 17: Table Creator Düğümüne Aktarılan Veriler.	82
Resim 3. 18: N Chars Filter Düğümü.	82
Resim 3. 19: Dictionary Tagger Düğümü.	83
Resim 3. 20: Bag of Words Creator Düğümü.	83
Resim 3. 21: Kelime Çantası Oluşturma İşlemi.	84
Resim 3. 22: DF Düğümü.....	84
Resim 3. 23: Hüküm Verileri İçerisindeki Kelimelerin DF Ağırlıklandırma Yöntemine Göre Hesaplanan Frekans Değerleri.....	85
Resim 3. 24: Frequency Filter Düğümü.	85
Resim 3. 25: Frekanslarına Göre Filtreleme İşleminin Çıktıları.	86
Resim 3. 26: KNIME Sınıflandırma Modeline Ait Veri Ön İşlem Aşamasının Genel Görünümü.	87

Resim 3. 27: En Çok Kullanılan 1.000 Kelimenin Kelime Bulutu İle Gösterimi.	88
Resim 3. 28: Document Vector Döğümü Çıktısının Örneğ Gösterimi.	93
Resim 3. 29: Decision Tree Learner Döğümü.	95
Resim 3. 30: Decision Tree Predictor Döğümü.	96
Resim 3. 31: SVM Learner Döğümü.	96
Resim 3. 32: SVM Predictor Döğümü.	97
Resim 3. 33: Gradient Boosted Trees Learner Döğümü.	98
Resim 3. 34: Gradient Boosted Trees Predictor Döğümü.	98
Resim 3. 35: Random Forest Learner Döğümü.	99
Resim 3. 36: Random Forest Predictor Döğümü.	99
Resim 3. 37: K Nearest Neighbor Döğümü.	100
Resim 3. 38: Scorer Döğümü.	101
Resim 3. 39: KNIME İle Oluşturulan Sınıflandırma Modeli.	102

DENKLEMLER LİSTESİ

Denklem 3. 1: Doğruluk Oranı.....	104
Denklem 3. 2: F-ölçütü.	104
Denklem 3. 3: Kesinlik.	104
Denklem 3. 4: Duyarlılık.....	105



KISALTMALAR

B: Receb

BOA: Bařbakanlık Osmanlı Arřivi

C: Cemaziyelahir

Ca: Cemaziyelevvel

CRISP-DM: Cross-Industry Standard Process For Data Mining

DF: Doküman Frekansı

HTML: Hyper Text Markup Language

ICF: Ters Sınıf Frekansı

IDF: Ters Doküman Sıklığı

KNN: K-Nearest Neighbors

L: řevval

LTC: Logaritmik Terim Bileřeni

M: Muharrem

N: Ramazan

NLP: Natural Language Processing

R: Rebiyülahir

Ra: Rebiyülevvel

S: Safer

SOM: Self Organizing Maps

SVM: Support Vector Machine

Ş: Şaban

TF: Terim Sıklığı

Z: Zilhicce

Za: Zilkade

XML: Extensible Markup Language



GİRİŞ

Hayatın normal akışı içerisinde çok farklı aktörler tarafından gerçekleştirilen iş ve işlemler sonucunda oluşan veriler, teknolojik gelişmelerin de etkisiyle, zamanla nitelikleri farklı olan yığınlar haline gelmiş durumdadır. Bu durum, hacmi büyük miktarlara varan verilerin depolanmasını ve düzenlenmesini gerektirmektedir. Ayrıca, ihtiyaç duyulması halinde, biriken büyük hacimdeki veriye tekrar başvurulduğunda en doğru ve kullanılabilir veriye en kısa zamanda ulaşılabilmesi gerekmektedir. Söz konusu gereklilikler ile verilerin sahip olduğu değerler, mevcut (ve gelecekte popüler olması muhtemel) bilim dallarının birçok dalını ilgilendirebilecek konulardaki çalışmalar için bilgisel kaynak ihtiyacını karşılayabilir durumdadır. Bu çalışmaların temel amacı, uygun veriler üzerinde ulaşmak istenilen hedeflere yönelik işlemlerin yapılması ve böylece yeni bilgilerin elde edilmesidir. Depolanan ve düzenlenen veriler ilk bakışta bir anlam ifade etmeyebilmektedir. Veri içerisinde gömülü halde bulunan bilgileri ortaya çıkarabilmek ve anlamlandırmak için, oluşan veriler üzerinde bir dizi teknik ve analiz işleminin yapılması gerekebilmektedir. Verilerin analiz edilebilmesine olanak sağlayan program ve algoritmalar ile gerçekleştirilen detaylı incelemeler sonucunda, veri içerisinde bulunan gömülü bilgiler (daha önce fark edilmemiş bilgiler) ortaya çıkarılabilmektedir.

Verileri anlamlandırmak ve analiz ederek yararlı bilgilerin elde edilmesini sağlamak amacıyla yapılan işlemler bulunmaktadır. Bu işlemler veri madenciliği ve metin madenciliğidir. Akgöbek ve Kaya (2011: 238) çalışmalarında veri madenciliğini “büyük ölçekli veriler arasından bilgiye ulaşma ve bu verilerden genellenmiş bilgiler elde etme” şeklinde tanımlamıştır. Gümüşsu ve Murat (2019: 288) metin madenciliğini benzer bir ifade ile “çok büyük metin verileri içerisinde bilgiye ulaşma işlemi” olarak tanımlamıştır. Dolgun, Özdemir ve Oğuz (2009: 49) ise veri madenciliğinin gerçekleştirilebilmesi için verinin yapısal olması gerektiğine, metin madenciliğinin de web madenciliği ile birlikte verileri yapısal duruma getirdiğine değinmiştir. Böylece veri madenciliğinin yapısal veriler üzerinde uygulandığı fakat metin madenciliği ve web madenciliğinin yapısal olmayan ya da yarı yapısal halde bulunan verilere uygulandığı sonucuna ulaşmak mümkündür. Söz konusu tanımlar ve bilgiler

doğrultusunda veri madenciliği ve metin madenciliğinin uygulandıkları veri türüne göre tercih edildiği ve yukarıda bahsedildiği üzere veri içerisindeki gömülü bilgilerin ortaya çıkarılmasını hedefledikleri ifade edilebilmektedir.

Üzerinde işlem yapılacak olan veri kaynağının niteliği ve niceliği, ortaya çıkarılacak gömülü bilginin sınırlarını belirler. Arşivler araştırmacılara nitelikli ve büyük miktarda veri kaynağı sunmaları bakımından veri merkezi olarak kabul edilebilmektedir. Osmanlı İmparatorluğu’ndan günümüze intikal etmiş olan büyük miktardaki arşiv materyalini de içerisinde bulunduran Devlet Arşivleri Başkanlığı, araştırmacıların başvuracağı önemli bir veri kaynağı merkezi olarak görülebilir. Osmanlı İmparatorluğu sınırları içerisinde bulunmuş olan farklı milletler, coğrafik bölgeler ve ilişki içerisinde bulunulmuş olan devletler hakkında bilgiler barındırabilmesinin yanı sıra toplumun sosyal, kültürel ve ekonomik yaşamı ile ilgili dönem analizlerinin yapılabilmesi dolayısıyla birçok gizli kalmış bilgi ve örüntünün ortaya çıkarılmasına imkân sağlayabilecek olması, Osmanlı dönemi kurumlarında üretilen arşiv materyalinin veri kaynağı olarak ele alınmasını değerli kılmaktadır. Söz konusu veri kaynaklarının düzenlenebilmesi, analiz edilebilmesi ve gizli kalmış örüntülerin keşfedilebilmesi için veri ve metin madenciliği işlemlerinden faydalanılabilmektedir.

Yukarıda bahsedilen arşiv kaynağı içerisinde, oluşmaları ve içerikleri bakımından farklılık gösteren birçok kaynak türü mevcuttur. Ahkâm Defterleri bu kaynaklardan birisidir. Ahkâm Defterleri, halkın yerel mahkemelerde çözüme kavuşturulamayan sorunlarının çözülmesi için Dîvân-ı Hümâyûn’a başvurması (Tunç, 2022: 4) ve bildirilen sorunların çözümüne yönelik verilen hükümlerin kaydedilmesi sonucunda oluşmuştur (Günay, 2013a: 9). Söz konusu defterler resmî belge niteliği taşımaktadır ve “resmî belgeler kurumsal kültürden, siyasal, toplumsal ve tarihi bağlamlardan bağımsız olarak yazılıp oluşturulamaz” (Özdemirci ve Torunlar, 2015: 78). Defter içerisinde bulunan hükümler üzerinde metin madenciliği işlemleri uygulanarak, materyalde bulunan bilginin/verinin bağlam(lar)ı hakkında bilgi çıkarımlarında bulunulması mümkündür. Bilgi çıkarımına imkân sağlayan metin madenciliği işlemlerinden birisi sınıflandırmadır. Bir araya getirilen ve düzenlenen

hüküm verilerinin sınıflandırılması ile istatistiksel sonuçlar elde edilebilmektedir. Bu sonuçlar anlamlandırılarak bilgi çıkarımı işlemi gerçekleştirilebilir.

Tez çalışmasının temel amacı veri kaynağı olarak seçilen hüküm verilerinin başarılı bir şekilde sınıflandırılabilmesi için model oluşturulmasıdır. Çalışmanın sonunda hüküm verilerinin sınıflandırılması için bir model oluşturulmuştur. Söz konusu model ile manuel olarak yapılacak sınıflandırma işleminin neden olacağı iş yükünün ve zaman kaybının önüne geçilmesine, çeşitli istatistiksel sonuçların elde edilebilmesi mümkün kılınarak bu sonuçların değerlendirilmesine ve dolayısıyla veri içerisinde gizli kalmış örüntülerin ortaya çıkarılmasına imkân sağlanması hedeflenmektedir. Oluşturulmuş model ile hüküm verileri manuel olarak yapılan sınıflama işleminden daha hızlı bir şekilde sınıflandırılabilmiştir ve yapılandırılmış veriye dönüştürülen hüküm verileri üzerinde çeşitli analiz ve işlemlerin yapılması mümkün kılınmıştır. Dolayısıyla yapılabilecek analizler ile gizli kalmış bilgilerin çıkarımının yapılması mümkün olmuştur. Ayrıca gerçekleştirilecek yeni çalışmalara fikir sağlanması da amaçlar arasındadır.

Yukarıda bahsedilen amaçlara dayanarak çalışmanın hipotezi şu şekilde belirlenmiştir; “Osmanlı Türkçesi ile düzenlenmiş belgeler üzerinde metin madenciliği işlemlerinin uygulanması mümkündür ve bu işlemler ile Osmanlı Türkçesi belgelerin otomatik olarak konularına göre sınıflandırılabileceği bir model oluşturulabilir.”

Çalışmanın evrenini Osmanlı bürokrasisinde üretilen Ahkâm Defterleri oluşturmaktadır. Çalışmanın örneklemini ise 1 numaralı Erzurum Ahkâm Defteri’dir. Söz konusu defter M.1742 ile M.1749 (H.1155-H.1162) tarihleri arasında deftere yazılmış olan hükümleri kapsamaktadır. 1 numaralı Erzurum Ahkâm Defteri incelendiğinde kullanılan alfabenin Osmanlı İmparatorluğu döneminde kullanılmış olan ve ağırlıklı olarak Arap alfabesi temelinde geliştirilen harflerden oluştuğu ve Osmanlı Türkçesi ile yazıldığı görülmektedir. Defterde bulunan hükümlerin Latin harflerine çevrilmesi ve konularının belirlenmesi işleminin, tez çalışmasının süresinin uzamasına neden olacağı öngörülmüştür. Bu sebeple defter içerisindeki hüküm verileri, daha önce söz konusu defterin transkript edilmesi amacıyla oluşturulmuş dört farklı tezden elde edilmiştir. Bu tezler; Muhittin Kul ve Hikmet Çiçek’in 2011 tarihli

yüksek lisans tezleri, 2019 tarihli Feyzanur Susam ve Nazmiye Gül Kantaroğlu'nun yüksek lisans tezleridir. Tezlerde 1 numaralı Erzurum Ahkâm Defteri içerisindeki hükümler Latin harflerine çevrilmiş ve ayrıca konu başlıkları da belirlenmiştir. Toplamda 1216 adet hüküm verisi Excel programında bir araya getirilmiştir. Bu verilerin içerisinde 4 adet derkenâr¹, 3 adet ise iptal edilmiş hüküm bulunmaktadır. Hükümlerin sınıflandırılması işleminde 3 adet hükmün zaten iptal edilmiş olması, derkenârların ise kısa notlar olması ve hükümlerin konularına göre sınıflandırılmasında sonuca herhangi bir katkısı olmayacağı öngörüldüğünden söz konusu 7 veri çalışmanın kapsamı dışında bırakılmıştır. Sınıflandırma işlemi 1209 adet hüküm verisi üzerinde gerçekleştirilmiştir.

Çalışmada hüküm verilerinin algoritmalarca anlaşılabilir bir yapıya dönüştürülmesi için, verilere bir dizi ön işlem uygulanmıştır. Söz konusu işlemler için metin madenciliği yöntemleri kullanılmıştır. Ardından literatürde sıklıkla kullanılan ağırlıklandırma yöntemleri ve yine sık kullanılan sınıflandırma algoritmaları ile yapılandırılmış hüküm verileri sınıflandırılmıştır. Gerçekleştirilen işlemlerde araç olarak açık kaynak kodlu bir metin madenciliği programı olan KNIME kullanılmıştır.

Tez dört bölümden oluşmaktadır. Birinci Bölüm'de yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış veri türleri hakkında kısaca açıklayıcı bilgiler verilmiştir. Söz konusu verilerin anlamlandırılarak yeni faydalı bilgilerin çıkarılmasına imkân sağlayan veri madenciliği ve metin madenciliği yöntemlerinin neler olduğundan kısaca bahsedilmiştir. Gerekli tanımlamaların ardından, çalışmada, metin verilerinden oluşan veri kaynağı üzerinde yapılacak işlemlerde başvurulacak metin madenciliğinin tarihsel gelişimi hakkında bilgi verilmiştir. Ardından metin madenciliğinin uygulama sahalarının neler olduğu açıklanmıştır. Bölümün sonunda metin madenciliğinde uygulanan süreçler hakkında bilgi verilmiştir.

Tezin İkinci Bölüm'ünde veri kaynağı olarak ele alınan Ahkâm Defterlerinin tanımı ve tarihsel gelişimi hakkında bilgi verilmiştir. Ahkâm Defterlerinin önemi ve teknik özellikleri hakkında bilgiler verildikten sonra içeriği hakkında açıklamalarda

¹“Bir yazının kenarına yazılmış mütâlaa ve ifâde” (Sami, 2011: 476).

bulunulmuştur. Ardından 1 numaralı Erzurum Ahkâm Defteri hakkında genel bilgiler verilerek defterin içerdiği konular ele alınmıştır.

Üçüncü Bölüm’de sınıflandırma işlemi ile elde edilebilecek istatistiksel analiz tabloları verilmiştir. Bu tablolar vasıtasıyla elde edilebilecek bilgi çıkarımlarına da kısaca değinilmiştir. Ardından metin madenciliği işlemlerinin seçilen hüküm veri kümesi üzerinde uygulanması ele alınmıştır. Çalışmada kullanılan araçlar tanıtılmıştır. Veri kümesi üzerinde KNIME programı vasıtasıyla gerçekleştirilen veri ön işleme adımları program ara yüzünde kullanılan araçların kısaca tanıtılması ile açıklanmıştır. Daha sonra verilerin sınıflandırılması işlemi için kullanılan algoritmalar kısaca değerlendirilmiştir. Son olarak KNIME programında oluşturulan sınıflandırma modelinin genel görüntüsü ile birlikte elde edilen sınıflandırma işlemlerinin başarı oranları ağırlıklandırma yöntemlerine göre sunulmuştur.

Tezin son bölümü olan Dördüncü Bölüm’de ise çalışmanın sonuçları değerlendirilmiştir. Osmanlı Türkçesi ve dönem alfabesi ile düzenlenmiş metin verileri üzerinde başta sınıflandırma işlemleri olmak üzere metin madenciliği yöntemleri uygulanarak yapılacak olan yeni çalışmalara fikir sağlaması amacıyla önerilerde bulunulmuştur.

BİRİNCİ BÖLÜM

METİN MADENCİLİĞİ

1.1. Tanımı ve Kapsamı

Bilişim teknolojilerinin hızlı bir şekilde gelişmesi ile verinin üretildiği, kullanıldığı ve saklandığı ortamlarda zamanla değişiklikler meydana gelmiştir. Büyük bir hacim ve karmaşık bir yapıya sahip olan veriler, doğru bilgiyi hızlı şekilde elde etmeyi zorlaştırmaktadır. Kurumlar, şirketler ve kişiler yığın haline gelen verilerden anlamlı sonuçlar çıkarma, aralarında dikkat çekici ilişkiler kurma ve sonuç olarak verileri kendileri için faydalı birer çıktı haline getirme isteği içerisinde olmuşlardır.

Veri tabanlarında depolanan verilerden bilimsel, yönetsel, hukuksal vs. bakımlarından faydalı yeni çıktılar elde edilebilmesi için çeşitli araç ve teknikler geliştirilmiştir. Amaca uygun araç ve tekniklerin belirlenmesi, verinin yapılandırılmış, yarı-yapılandırılmış veya yapılandırılmamış bir şekilde olmasıyla ilişkilidir. Yapılandırılmış veriler “standart bir veri tabanı içindeki satır ve sütunlara uyan geleneksel verilerdir” (Babadağ, 2021: 47) şeklinde açıklanabilir. Yapılandırılmış verinin bir başka tanımı Gahi, Guennoun ve Mouftah (2016: 953) tarafından modellenabilir, girdi olarak sokulabilir, depolanabilir, sorgulanabilir, işlenebilir ve görselleştirilebilir olan tüm veri türleri şeklinde yapılmış; ayrıca bu tür veriler, üzerlerinde yapılacak işlemler için yüksek performanslı yetenekler ile paralel teknikler gerektirmediklerinden yararlı bilgilerin elde edilmesinin daha kolay olduğu şeklinde açıklanmıştır.

Yapılandırılmış verilerden yararlı bilgi elde edilebilmesi için veri madenciliği araçlarından ve tekniklerinden yararlanılmaktadır. Literatürde veri madenciliğinin birden çok tanımı bulunmaktadır. Temel olarak benzer tanımların yapıldığı fakat ortak bir tanım üzerinde fikir birliğinin sağlanamadığı gözlemlenmiştir. Veri madenciliği uygulamalarının kullanıldığı alanlar, yapılan tanımların şekillenmesinde önemli bir rol oynamaktadır. Literatürde veri madenciliği şu şekilde tanımlanmaktadır;

Veri madenciliği hakkında ilk yapılan tanımlardan biri “Verilerden örtük, daha önce bilinmeyen ve potansiyel olarak yararlı bilgilerin farklı yöntemlerle çıkarılması

işlemdir” (Frawley, Piatetsky-Shapiro ve Matheus, 1992: 58) şeklindedir. Bir başka tanımda “verilerden, belirli ancak bilinmeyen bir sınıfta yer alan nesne veya olayları ifade eden örüntülerin çıkarılması amacıyla belirli algoritmaların uygulanmasıdır” (Şentürk, 2006: 3) şeklinde açıklanmıştır. Diğer bir tanıma göre ise veri madenciliği, “büyük, gürültülü ve dağınık veri kümelerinde bulunan veri öğeleri arasındaki zayıf ilişki kalıplarını bulmak için makine öğrenimi algoritmalarının kullanılmasıdır” (Nisbet, Elder ve Miner, 2009: 17) şeklinde ifade edilmiştir.

Yapılandırılmamış veya yarı yapılandırılmış verilerin analizini yapabilmek için veri madenciliğinin bir alt türü olan metin madenciliği tekniklerinden yararlanılmaktadır. Literatürde yapılandırılmamış veri “düzgün bir şekilde satır ve sütunlara sığmayan, ilişkisel tablolara kolayca endekslenemeyen verilerdir” (Babadağ, 2021: 47) şeklinde açıklanmaktadır. Yarı yapılandırılmış veri ise Gahi ve diğerleri (2016: 953) tarafından yapılandırılmış verinin bir türü olarak ifade edilmiş fakat yapılandırılmış veriler gibi modellenemez olduğu sonucuna varılmaması gerektiği de belirtilerek “yapıyı tanımlayan modele ek olarak verilerin belirli alanlarını tanımlamak için kullanılan etiketler ve işaretçiler gibi bir tür tanımlayıcı model içeren veriler” şeklinde değerlendirilmiştir. Bu tür verilere örnek vermek gerekirse çalışmada ele alınan hükümler yarı yapılandırılmış türdedir. Hükümlerin tarihleri ve konuları belirli satır ve sütunlar içerisinde düzenlenebilirken hükmü oluşturan metin verisi tamamen yapılandırılmamış yapıdadır. Literatürde metin madenciliği tanımları ise şu şekildedir;

Oğuzlar (2011: 7)’a göre metin madenciliği, belirgin bir formatı olmayan, yazı biçimindeki veriler içerisinde gizli nitelikteki bilginin çıkarılması ve düzensiz haldeki verinin formatlanması sürecidir. Hearts (2003: 1), metin madenciliğini, “farklı yazılı kaynaklardan otomatik olarak bilgi ayıklayarak, daha önce bilinmeyen yeni bilgilerin bilgisayar tarafından keşfedilmesidir” şeklinde tanımlamıştır. Sukanya ve Biruntha (2012: 269) kavramı, “daha önce bilinmeyen bilgileri farklı yazılı kaynaklardan otomatik olarak çıkararak bulma keşfidir” biçiminde açıklamışlardır. Kobayashi, Mol, Vrolijk ve Kismihok (2021: 140) ise “metinler içerisindeki gizli kalıpların ve özelliklerin ayıklanması amacıyla çeşitli algoritmaların uygulanması”nı ifade ettiğini belirtmektedirler.

Metin madenciliği, yapılan tanımlardan anlaşılacağı üzere, veri madenciliği araç ve yöntemlerinin uygulanamadığı, çoğunlukla doğal dil ile oluşturulmuş metin formatında bulunan kaynakların ihtiyaca göre çözümlenmesi ve kullanılabilecek hale getirilmesi sürecidir şeklinde tanımlanabilir.

1.2. Tarihçesi

Metin madenciliği, duyulan ihtiyaçlar doğrultusunda meydana gelmiştir. Öncelikle sahip olunan büyük miktardaki verilerin kategorize edilmesi ihtiyacı doğmuştur. “Metin madenciliği gelişmeleri, metin belgelerini (örneğin, bir kütüphanedeki kitaplar) kataloglama ihtiyacı ile başlamıştır” (Miner, Delen, Elder, Fast, Hill, Nisbet, 2012: 3). Bilginin/verinin gruplandırılması şeklinde ortaya çıkan çözüm, bilgi ihtiyacında olanlar için her ne kadar kolaycı ve nispeten hızlı veri sağlasa da bilginin/verinin ayrıntılı ve ilişkili olanın bulunması yönünde fazla kolaylık sağlamamıştır. Ancak kategorize edilen bilgiye hızlı ve ayrıntılı bir şekilde erişme isteği beraberinde yeni çalışmalara zemin hazırlamıştır. Bu duruma bağlı olarak, “Bilgi erişim ile ilgili çalışmalar 1960’lı yıllarda başlamış, doğal dili anlamaya ve sayısallaştırmaya yönelik programlar geliştirilmeye çalışılmıştır” (Oğuz, Bilge, Saka, 2007: 108). 1990’lı yıllardan itibaren ise internetin yaygınlaşması ve popüler hale gelmesiyle kişisel bilgisayarlarda veri üretimi hızla artmaya başlamıştır. Böylece söz konusu ihtiyaçlar sunulan çözümlerle birlikte daha fazla çeşitlenmiş ve artık elde edilen verilerden bilgi çıkarımı yapılabilmesi için çalışmalara başlanmıştır. Oğuz ve diğerleri (2007: 108), veri hızının arttığı 1990’lı yıllarda metinlere erişim, metinlerden bilgi çıkarımının yapılması, metinlerin kategorilere ayrılması ve metinlerin yapısal hale getirilmesi gibi durumlara yönelik çalışmaların arttığını belirtmektedir. Organizasyonların ve kişilerin istek ve ihtiyaçları doğrultusunda çalışmalar kümülatif olarak detaylandırılarak metin madenciliği uygulamaları mevcut halini almıştır.

1.3. Metin Madenciliğinin Uygulama Sahası

Web ortamında oluşturulmuş bir haber yazısı, sosyal medya hesaplarından yazılan metin içerikli paylaşımlar ve mesajlar, elektronik belge yönetim sistemleri içerisinde oluşturulan yazışmalar, HTML vb. günlük olarak yapılandırılmamış veya

yarı yapılandırılmış bir biçimde sürekli üretilen verilerden bilinmeyen, gizli kalmış örüntülerin çıkarımını yaparak fayda elde etme isteği, teknolojinin gelişimiyle birlikte hız kazanmıştır.

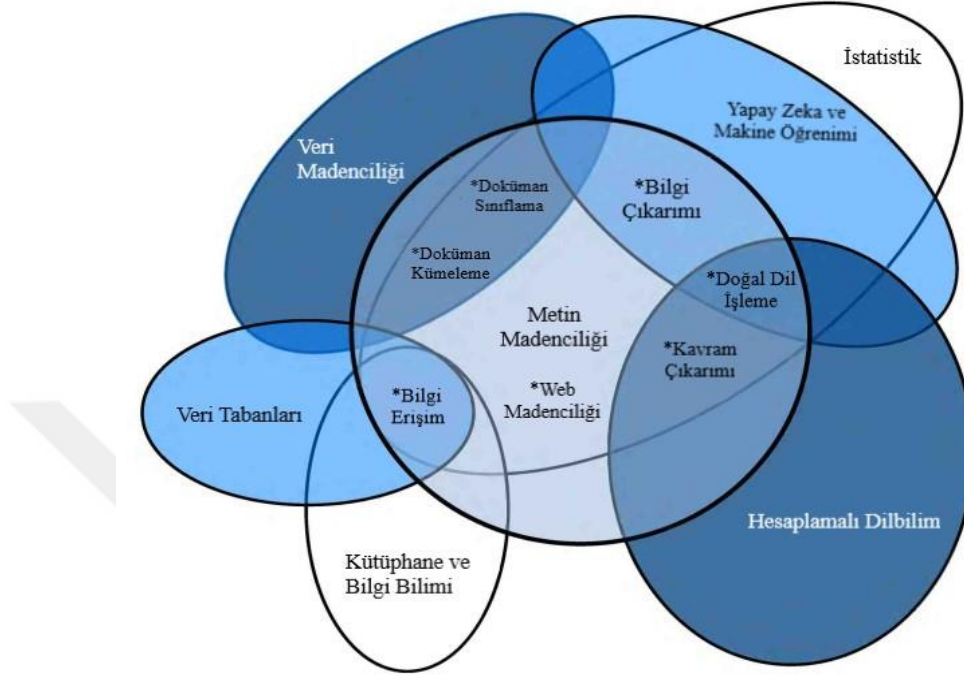
Günümüzde basılı ya da elektronik ortamda oluşturulan metinlerin yaklaşık %80'i yapılandırılmamış ya da yarı yapılandırılmış metin verilerinden oluşmaktadır. Meydana gelen büyük miktardaki veri yığını farklı tekniklerle ve araçlarla analiz edilerek organizasyonlar için faydalı birer çıktı haline getirilmektedir. Veri yığınlarını anlamlı şekilde analiz etme imkânı sunan farklı teknik ve araçlar vardır. “Özellikle web ortamında elde edilen metin halindeki büyük verinin analiz edilmesinde metin madenciliği önemli bir istatistiksel tekniktir” (Doğan ve Arslantekin, 2017: 71).

Birçok alanda metin madenciliği araç ve uygulamaları kullanılmaktadır. Metin madenciliğinin disiplinler arası olması ve doğal dil ile oluşturulmuş veriler ile ilgilenmesi ve bu verilerin miktarındaki hızlı artış, kullanıldığı uygulama sahalarında ya da çalışma alanlarında fazlaca çeşitlilik göstermesini sağlamaktadır. Metin madenciliği uygulamaları aşağıdaki alanlarda sıklıkla kullanılmaktadır:

- “Yayıncılık ve medya
- Bankalar, sigorta ve finans piyasaları
- Telekomünikasyon, enerji ve diğer hizmet sektörleri
- Bilişim sektörü ve internet
- Eczacılık ve sağlık hizmetleri
- Siyasî ve adlî kurumlar gibi alanlarda sıklıkla kullanılmaktadır” (Bolasco, Canzonetti, Capo, Rinaldi, Singh, 2005: 36).

Bilişim hizmetlerinin ve araçlarının çeşitlenmesiyle birlikte ilişkili olduğu alanlar da sürekli artmaktadır. Bu sebeple literatürde, metin madenciliği uygulama sahaları ve ilişkili olduğu alanlar ile ilgili farklı başlık ve açıklamalar ile karşılaşılmuştur. Genel hatlarıyla Miner ve arkadaşları (2012: 31) metin madenciliğinin yedi uygulama sahası olduğunu ve altı alanla ilişkili olduğunu belirtmiştir.

Şekil 1. 1. Metin Madenciliğinin Uygulama Sahaları ve İlişkili Olduğu Alanlar.



Kaynak:(Miner ve diğerleri, 2012: 31).

Metin madenciliği uygulama sahaları bilgi erişim, bilgi çıkarımı, doğal dil işleme, web madenciliği, kavram çıkarımı, doküman kümeleme ve doküman sınıflamadır (Miner ve diğerleri, 2012: 31).

1.3.1. Bilgi Erişim

İnternet ve bilgisayarların, günlük hayatta önemli birer araca dönüşmesi ile birlikte kişisel ve kurumsal verinin/bilginin çoğu sayısal ortamda üretilmeye başlamıştır. Verinin/bilginin aktarılması işlemi de benzer şekilde bu gelişmelerden derinden etkilenmiştir. Teknolojik araç ve gereçlerin gelişmesiyle bilgi aktarımı daha kolay bir hale gelmiştir. Verinin/bilginin üretilmesinde ve paylaşılmasında yaşanan etkin gelişmeler bilgi üretim hızını da önemli ölçüde etkilemiş ve verinin/bilginin üretiminde hızlı bir patlama yaşanmıştır. Bu durum beraberinde istenilen doğru bilgiye hızlı bir şekilde erişilmesi konusunda problemler meydana getirmiştir.

Bilgi erişim kavramı genel olarak “depolanmış halde bulunan bilgiler içerisinde alakalı bilgilerin bulunması” (Dominich, 2008: 8) şeklinde

tanımlanabilmektedir. Literatürdeki bir diğer tanım incelendiğinde ise kavramın “kullanıcının belirli bir alanda mevcut bilgi birikimi içinden, kendisi için en uygun bilgiye en kısa zamanda ulaşmasıdır” (Taner, 2008: 3) şeklinde açıklandığı görülmektedir. Kavram hakkında yapılan tanımlar incelendiğinde bir bilgi kaynağı/birikimi içerisinden bilgiye erişildiği anlaşılmaktadır. Yukarıda bahsedildiği üzere bilginin/verinin kolaylıkla ve hızlı bir şekilde üretilebilmesi, bilgiye erişilecek olan kaynağın/birikimin hacmini artırmaktadır. Düzensiz bir yapı içerisinde erişilmek istenilenin bulunması doğal olarak kolay olmayacaktır. Bu sebeple kaynağın/birikimin belirli kurallara göre düzenlenmesi ve sınıflandırılması, bilgi erişim işleminin hızlı ve doğru bir şekilde gerçekleştirilebilmesi için önemlidir çıkarımı yapılabilmektedir. Söz konusu kavram, bilgiye erişilmek istenildiğinde başvurulabilecek, oluşumlarına neden olan durumlara göre büyük miktarda bilgi kaynağını/birikimini içerisinde bulunduran, önemli bilgi merkezlerinden arşivler ve kütüphaneler açısından da ele alınmıştır. Bu durumun sebebi “Bilginin sağlanması, düzenlenmesi ve eriştirilmesi bilgi kurumlarının temel sorumluluklarındandır” (Çapkın, 2016: 679) şeklinde ifade edilebilir. Tonta (1996: 215-230) çalışmasında bilgi erişimi kütüphaneler açısından ele almış, elektronik ortamda bulunan bilgi kaynaklarının sayısının arttığına ve bu kaynakların çeşitlendiğine değinmiştir. Söz konusu çalışmada internetin ve bilişim teknolojilerinin gelişmesiyle birlikte kütüphanelerin ve ayrıca bilgi merkezlerinin sunmuş oldukları hizmetleri tekrar gözden geçirmeleri gerektiğine ve bu kaynaklara kullanıcıların erişebilmeleri için çalışmaların yapıldığına da değinmiştir. Çoğunlukla metinsel formatta bulunan kaynakların kullanıcıların hizmetine sunulmasından farklı olarak ses, görüntü ve grafik içeren elektronik kaynakların da erişime sunulması için çalışmaların yapılmasını belirtmektedir. Menne-Haritz (2008: 71-79) ise bilgi erişim kavramını arşivler açısından ele almaktadır. Söz konusu çalışmada arşivlerin, geçmişini inceleyerek günümüz problemlerinin anlaşılmasını ve geleceğe yönelik planların yapılmasını mümkün kıldıklarına değinmiştir. Ayrıca arşivlerin kullanıcılara sunduğu başlıca hizmetin ise üzerinde bilgi çıkarımı yapılacak olan ham maddeye erişim olduğuna değinilmiştir. Burada kullanılan ham madde kavramı arşiv kayıtlarını ifade etmektedir. Arşivlerin potansiyel enformasyonu sunarak araştırmaya zemin hazırladıkları böylece kullanıcıların yaptıkları çalışmalar sonucunda bilgi elde ettikleri de belirtilmiştir. Yukarıdaki bilgiler doğrultusunda ve metin madenciliğinin tanımları

da göz önünde bulundurularak bilgi erişim kavramı metin madenciliği açısından ele alınacak olunursa; var olan bilgi kaynağının/birikiminin içerisinde daha öncesinde bilinmeyen, gizli kalmış bilgileri ortaya çıkarmak için alakalı olabilecek bilgilere/verilere erişilmesidir denilebilir.

1.3.2. Bilgi Çıkarımı

Bilgi çıkarımı kavramı “doğal dil ile oluşturulmuş bir metinden, varlıkların tanımlanması, tanımlanan varlıklar arası ilişkilerin belirlenmesi ve olayların belirlenmesi sürecidir” (Arslan, 2011: 11) şeklinde tanımlanmaktadır. Bilgi çıkarım, büyük miktardaki metin kaynağı üzerinde doğal dil işlemenin kullanılması sonucu insan müdahalesini minimize ederek belirli kriterler hakkında bilgi edinilmesini amaçlamaktadır (Pasin, 2018: 18). Bazen belgenin tamamına erişmek istenilmesinin dışında belgenin bir kısmına ve belirli verilere ulaşmak istenebilir. Bilgi çıkarımına ise bu durumlarda başvurulmaktadır. Bilgi çıkarımı doğal dil ile oluşturulmuş metinsel verilerden öncesinde belirlenmiş olan bilgilerin ya da verilerin otomatik olarak tanımlanmasını, tanımlanan bu bilgi/verilerin bir araya getirilerek yapılandırılmış veritabanında saklanmasını ifade eden işlemlerdir (Vlahovic, 2010: 575). Bilgi çıkarımının metinsel veri içerisinde bulunan kişi adları, kurum ve yer bilgileri, tarih ve saat detayları, para birimleri ya da miktarları ile kısaltmaların adları ve bunların yanı sıra ilişkisel bağlantılar ile kopuklukları (Ergün, 2017: 182) kullandığı ifade edilmektedir. Bilgi çıkarımı işleminin işlevselliği bilgi kaynağının yapılandırılma süreciyle doğrudan ilişkilidir. “Kaynak iyi yapılandırılmış ise işlem oldukça basit ve anlaşılır olabilir. Bilgi kaynağı az yapılandırılmış ise ya da doğal dil halinde ise doğal dil tanıma ve benzeri süreçleri içerdiğinden çıkarma işleminin karmaşıklığı artar” (Vlahovic, 2010: 575).

Bilgi çıkarımı ile bilgi erişim kavramlarının karıştırılması muhtemeldir. Kavramlar hakkında yapılan tanımlamalar incelendiğinde aralarındaki fark şu şekilde ifade edilebilir; bilgi erişim kavramı uygun sorgulamalar sonucunda ihtiyaç duyulan bilgi ile ilişkili kaynaklara erişmeyi ifade etmektedir. Bilgi çıkarımı ise erişim sağlanan söz konusu kaynaklar içerisindeki bilgileri yapısal hale dönüştürerek sunma işlemidir. Ayrıca Vlahovic (2010: 575) bilgi çıkarım ve bilgi erişim kavramları arasında bulunan

temel farkın; bilgi çıkarım işleminde ilgilenilen gerçeklerin önceden belirtilmiş iken bilgi erişim işleminde kullanıcının bilmediği ve ilgisini çekebilecek gerçeklere sahip olabilen kaynakların keşfedilmeye çalışılması olduğunu belirtmiştir.

1.3.3. Doğal Dil İşleme

İlk olarak yapay zekânın bir alt dalı olarak görülmesine karşın doğal dil işleme, insana ait olan dilin bilgisayarlar tarafından tanımlanabilmesi, işlenmesi ve yeniden üretilmesi gibi konularda yapılan çalışmaların artması ve bu çalışmaların başarılı sonuçlar elde etmesiyle birlikte bilgisayar bilimlerinde kendine önemli bir yer edinmiştir. Doğal dil işleme, “ana işlevi bir doğal dili çözümleme, anlama, yorumlama ve üretme olan bilgisayar sistemlerinin tasarımını ve gerçekleştirilmesini konu alan bir mühendislik alanıdır” (Oğuzlar, 2011: 11) şeklinde ifade edilmektedir. Literatürde yer alan bir diğer tanımda ise “Doğal dil işleme (NLP), bilgisayarların metni bir insan gibi parçalamasına ve anlamasına olanak tanıyan bir makine öğrenim tekniğidir” (Akyüz ve Gülten, 2022: 13) şeklinde açıklanmaktadır.

Başarılı bir doğal dil işleme sürecinin gerçekleştirilebilmesi için, işleme alınacak metin verilerinin yazıldığı dilin özellikleri hakkında sahip olunan bilgi belirleyici olmaktadır. Covington (1994: 5) dilin yapısını beş seviyeye ayırmaktadır. Söz konusu seviyeler;

- Sesbilim
- Biçimbilim
- Söz dizimi
- Anlambilim
- Kullanımbilimdir.

Dil bilimcilerin yukarıda ifade edilen konular üzerinde, dilin özellikleri ve kuralları hakkında yapmış oldukları belirleyici çalışmalar, doğal dil işleme çalışmalarında yol gösterici olma niteliği taşımaktadır. Doğal dil işleme kapsamında genel olarak çalışılan konular şu şekilde sıralanabilir (Adalı, 2012: 4);

- “Yazım yardımcı araçlarının geliştirilmesi
- Yazım yanlışlarının düzeltilmesi
- Bul ve değiştir
- Basılı bir metni okuma (optik olarak metin okuma) ve okuma yanlışlarını düzeltme
- Bir metnin özetini çıkarma
- Metnin içerdiği bilgiyi çıkarma
- Bilgiye erişim
- Metni anlama
- Bilgisayarla sesli iletişim
- Bilgisayarın konuşması (metni seslendirme)
- Konuşmayı anlama (konuşmayı metne dönüştürme)
- Soru yanıt dizgeleri
- Yabancı dil okuma yardımcı araçları
- Yabancı dilde yazma yardımcı araçları
- Doğal diller arası çeviri”

1.3.4. Web Madenciliği

“Web madenciliği, web sitelerinin tasarımlarının geliştirilmesi, e-ticaretin geliştirilmesi ve bilgi çıkarımı yapılabilmesi için web sayfalarının içeriğinden, e-ticaret kaynaklarından ve web sayfalarının linklerinden elde edilen verilere veri madenciliği tekniklerinin kullanılarak ilgi çekici ve potansiyel bilgileri elde etme işlemidir” (Gupta ve Lehal, 2009: 73). Web ortamında bulunan veri yığınının düzenlenmesi ve kullanılması, e-ticaretin de gelişmesiyle birlikte önem kazanmıştır. Web madenciliği uygulamaları ile web ortamında üretilen içerikler, log kayıtları, kullanıcıların kayıtlı bilgileri, oturum süreleri ve hareketleri gibi veri kaynakları incelenerek çeşitli analizler elde edilebilmektedir. Organizasyonlar elde edilen analiz sonuçlarına göre iş süreçlerinde iyileştirmeler yapmaktadırlar.

Web içerisinde bulunan veri çeşitliliği oldukça fazladır. Heterojen bir yapıda bulunan verilerin sorunsuz olarak incelenebilmesi için farklı uygulamalar gerekmektedir. Dolgun ve diğerleri (2009: 51-52) web madenciliği uygulamalarını;

- Web yapı madenciliği
- Web içerik madenciliği ve
- Web kullanım madenciliği olmak üzere üç başlık altında ele almıştır.

Web yapı madenciliği, web siteleri, web sayfaları ya da söz konusu sayfalarda bulunan bağlantılar arasındaki ilişkileri incelemektedir. Web içerik madenciliği ile web sayfalarını oluşturan içerikler incelenmektedir. Böylece içerik bakımından incelenen sitelerin belirli sınıf ya da kümelere ayrılması mümkün olmaktadır. Web kullanım madenciliği ise web sunucularında bulunan kullanıcıların erişim kayıtları incelenerek bir web sitesindeki ziyaretçilerin site içerisinde yapmış oldukları faaliyetler hakkında bilgi sahibi olunmasına imkân tanımaktadır (Dolgun ve diğerleri, 2009: 52-53).

1.3.5. Kavram Çıkarımı

Kavram çıkarımı, “Kelime ve deyimlerin anlamsal olarak benzer gruplara ayrılması” (Miner ve diğerleri, 2012: 32) şeklinde tanımlanmıştır. Bir diğer tanımda ise “yapılandırılmamış metindeki ilgi alanlarına ilişkin kavramlara atıfta bulunan ifadelerin belirlenme süreci” (Torii, Waghlikar ve Liu, 2011: 580) olarak ifade edilmiştir. Söz konusu tanımlar doğrultusunda şu çıkarımları yapmak mümkündür; Kavram çıkarımı kaynak veriler içerisinde anlamsal olarak benzer kelimelerin gruplandırılarak erişilmesine imkân sağlamaktır. Erişilmek istenen bilgilerin yerleri hakkında bilgi verir. Kelime birlikteliği analizlerinin ve duygu analizlerinin yapılmasında kolaylıklar sağlamaktadır. Metinler kelime torbası ya da kelime kümesi olarak düşünülebilir ve metin içerisinde geçen kelimelerin çoğu metin hakkında bilgi vermemektedir. Bir doküman, bir ya da daha fazla kavramsal konu içerebilir ve bu kavram/kavramlar, bazı karakteristik terimler ile temsil edilebilir. Dolayısıyla doküman da bir dizi karakteristik terimle tanımlanabilir. Dokümanların sınıflama işlemleri de genellikle konu kavramlarına odaklanmaktadır. Dolayısıyla kavram çıkarma işlemi kavramlara dayalı sınıflama için çok önemlidir (Zhang, Gongi Wang ve Yin, 2003: 66-67). Bu sebeplerle söz konusu kavram, üzerinde çalışma yapılacak olan dokümanlardan bir dizi kavramın çıkarılmasını ifade etmektedir denilebilir.

1.3.6. Doküman Kümeleme

“Metin Kümeleme, doküman koleksiyonlarının doküman benzerliklerine bağlı olarak gözetimsiz ve otomatik biçimde gruplara ayrıştırılmasıdır” (Tunalı ve Bilgin,

2012: 598). Söz konusu işlem geniş doküman verilerinde içerikleri bakımından benzer dokümanların tespit edilerek gruplandırılmasını amaçlamaktadır (Canhası, 2007: 30). Üzerinde analiz işlemi gerçekleştirilecek olan dokümanların çeşitli kümeleme algoritmalarının kullanılması ile gruplandırılması için kullanılan denetimsiz bir işlemdir ve verilerdeki içsel yapıların ortaya çıkarılması, söz konusu yapılar hakkında çalışmaların yapılması ve bu yapıların analiz edilmesi için verilerin gruplara ayrılmasını amaçlamaktadır (Öztürk, 2021: 28). Bu bilgiler doğrultusunda doküman kümeleme, herhangi bir kategoriyle ilişkilendirilmemiş ya da belirli bir sınıfa ait olmayan dokümanların bir araya getirilerek gruplanması işlemidir şeklinde ifade edilebilir. Babadağ (2021: 44), kümeleme işleminde grupların neler olacağı ve söz konusu grupların özelliklerinin ne olduğu hakkında bilgi sahibi olunmadığını belirtmiştir. Ayrıca kavramı, dokümanların benzerliklerine göre grupların oluşturulmasının ardından dokümanın hangi gruba ait olduğunun belirlenmesi işlemi olarak ifade etmiştir. “Metin kümeleme işlemlerinde kullanılan birçok denetimsiz öğrenme algoritmaları vardır. K-means, hiyerarşik kümeleme ve SOM (Self Organizing Maps) yöntemleri sıklıkla kullanılmaktadır” (Gümüüşsu, 2019: 56-57). Bir kümeleme algoritması, her bir belgenin konusu için vektör oluşturmakta ve belgelerin bu kümelerle olan uyumunu ölçmektedir (Gupta ve Lehal, 2009: 64). Kümeleme işlemi ile verilerin dağılımı anlaşılabilir ve oluşan kümelerin özellikleri izlenebilir (Canhası, 2007: 31).

1.3.7. Doküman Sınıflama

Teknolojik gelişmelere bağlı olarak belge üretimi kolaylaşmıştır ve bu durum belge sayısının artmasını meydana getirmiştir. Yığın haline gelen belgelerin içerisinde bulunan değerli bilgilerin ortaya çıkarılabilmesi için söz konusu belgelerin içeriklerinin belirlenmesi ve işlenmesi (sınıflandırılması) gerekmiştir (Haltaş, Alkan ve Karabulut, 2015: 419). Metin madenciliğinde doküman sınıflaması belirli özellikler bakımından ilişkili kavramların ve temaların aynı özellikler bakımından bir araya getirilerek gruplandırılmasıdır. Metin sınıflandırmasının amacı bir dizi belgeyi önceden tanımlanmış sabit sayıda kategoride sınırlamaktır ve her belge birden fazla sınıfa ait olabilmektedir (Gupta ve Lehal, 2009: 64). Kategorisi belli olmayan yeni bir

veri eklenirken, önceden tanımlanmış olan kategorilerden elde edilen model belirleyici olmaktadır. Dokümanların otomatik olarak sınıflanması işlemi, söz konusu dokümanların kullanıcıların ihtiyacına uygunluğuna göre önceliklendirmesinde yardımcı olmaktadır (Zhang ve diğerleri, 2003: 66).

Doküman sınıflama ile doküman kümeleme birbiriyle karıştırılması muhtemel olan iki farklı kavramdır. Söz konusu kavramlar ile ilgili yapılan tanımlar incelendiğinde iki kavram arasındaki farkları şu şekilde ifade edebilmek mümkün olabilmektedir; Doküman sınıflamada yeni bir veri ya da doküman, önceden belirlenmiş olan gruplara dâhil edilmektedir, fakat doküman kümelemede hangi gruba dâhil edileceği önceden belirlenmemiştir. Doküman sınıflamada veri kaynakları benzerliklerine göre bir araya getirilirken doküman sınıflandırmada ortak özelliklerine göre bir araya getirilme işlemi gerçekleştirilmektedir. Pilavcılar (2011: 2), metin sınıflama işleminin temelde iki adımdan oluştuğuna değinmiştir. Söz konusu adımları ise, yapılmış olan sınıflama işleminin performansının değerlendirileceği algoritmanın seçilmesi ve seçilmiş olan algoritmanın uygulanacağı veri kümesinin seçimi şeklinde ifade etmiştir. Karaca, Günel ve Taştan (2015: 808-809) ise sınıflamanın zaman alıcı bir işlem olduğunu ifade etmektedirler. Manuel olarak yapılan sınıflama işlemlerinde sonuçların, sınıflamayı yapan uzmanların kararları doğrultusunda değişiklik gösterebildiğine de değinmişlerdir. Ayrıca sınıflama işleminin, veri setinin oluşturulması, ardından bu verilerin ön işlem aşamalarından geçirilmesi, özellik seçiminin uygulanması, sınıfların özellik ve doküman vektörlerinin oluşturulması ve son olarak benzerliklerin hesaplanması gibi adımlar sonucunda gerçekleştirildiğini belirtmektedirler. Doküman sınıflamada kullanılan algoritmalara örnek olarak K-en yakın komşu (KNN), Naive Bayes, karar ağaçları ve yapay sinir ağları verilebilir.

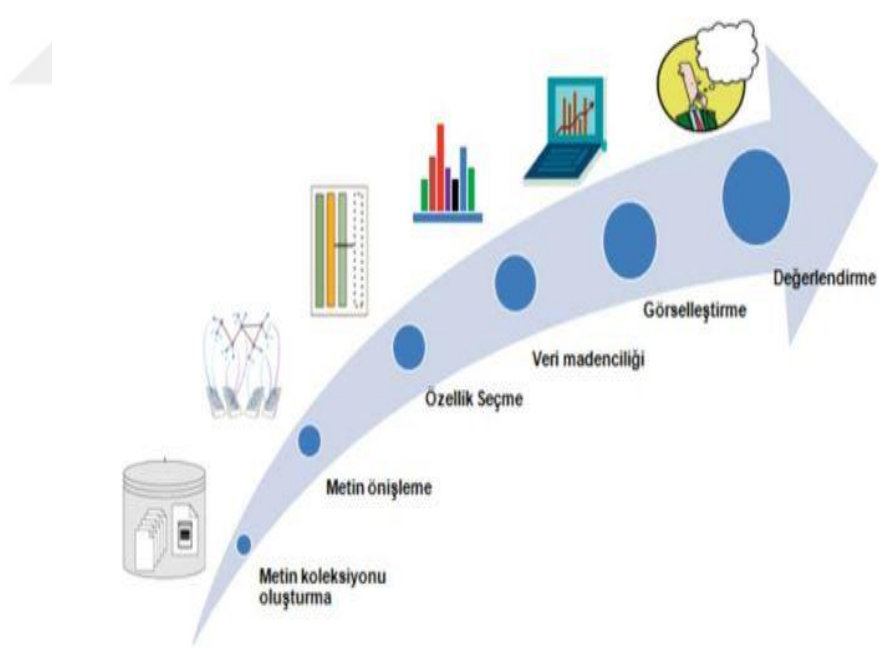
1.4. Metin Madenciliği Süreci

Metin madenciliği yapılırken izlenecek olan ortak bir süreç modeli bulunmamaktadır. Yapılan literatür incelemesinde veri madenciliği süreçleri üzerinde genel bir fikir birliğinin mevcut olduğu görülmektedir, fakat metin madenciliği için aynı durum söz konusu değildir. Literatürdeki çalışmalar incelendiğinde metin madenciliği süreçlerinin sıklıkla farklılık gösterdiği saptanmıştır. Bu sebeple veri

madenciliği yapılırken izlenilen süreçler metin madenciliği süreçlerine göre daha fazla olgunlaşmıştır sonucu çıkarılabilmektedir. Bu duruma neden olarak metin madenciliğinin veri madenciliğinden sonra ortaya çıkmış olması, işlem yapılacak veri kaynaklarının ve elde edilmek istenilen sonuçların türleri gösterilebilir.

Oğuzlar (2011: 9-10) metin madenciliği süreçlerinin beş adımda gerçekleştirildiğini belirtmektedir. Bu adımlar; metin koleksiyonu oluşturma, metin ön işleme, veri madenciliği, değerlendirme ve yorumlamadır. Akyüz ve Gülten (2022: 17) ise metin madenciliği süreçlerini üç adımda ele almaktadır. Bunlar; metin ön işleme, metin sayısallaştırılması ve bilgi keşfi aşamalarıdır. Oğuz ve diğerleri (2007: 109-110) ise metin madenciliğinin altı adımdan oluştuğunu belirtmiştir. Söz konusu adımlar; metin toplama, metin ön işleme, özellik seçme, veri madenciliği, görselleştirme ve değerlendirmedir.

Şekil 1. 2. Metin Madenciliği Süreci.



Kaynak: Oğuz ve diğerleri, 2007: 110.

Bu çalışmada veri madenciliği ve metin madenciliği hakkında yapılmış olan tanımlar ve açıklamalar doğrultusunda şu çıkarımlarda bulunulabilir: Veri madenciliği ve metin madenciliği alanları, temel prensipleri ve problemleri ele alış açıları bakımından birbirlerine oldukça yakındır. Literatürde veri madenciliği ile ilgili

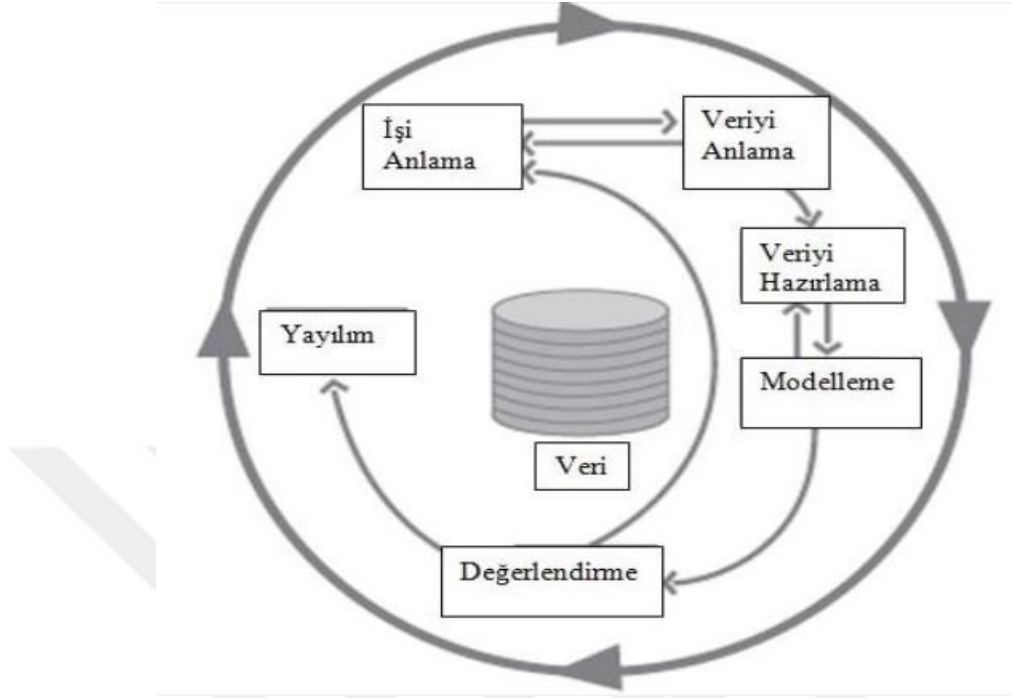
yapılmış olan çalışmalar incelendiğinde Veri Madenciliği için Çapraz Endüstri Standart Süreci'nin (CRISP-DM) sıklıkla kullanılan bir yöntem olduğu sonucuna varılmaktadır. Söz konusu çalışmalara Şekeroğlu'nun (2010: 95-105), Şık'ın (2014: 71-102) ve Küçükşille'nin (2009: 59-84) yapmış olduğu çalışmalar örnek gösterilebilir. Dolayısıyla CRISP-DM modeli esas alınarak metin madenciliği yapılması bir sorun oluşturmamaktadır. Söz konusu iki alan arasındaki farklılıklar sebebiyle süreç basamaklarında esneklikler ya da değişiklikler yapılabilmektedir.

Çalışmanın bu kısmında CRISP-DM modeli ile birlikte metin madenciliği süreçleri de açıklanacaktır.

1.4.1. CRISP-DM Modeli

1999 yılında, otomobil üreticisi Daimler-Benz, sigorta sağlayıcısı OHRA, donanım ve yazılım üreticisi olan NCR ve istatistiksel yazılım üreticisi olan SPSS gibi farklı sektörlerden bir araya gelmiş büyük şirketler, veri madenciliğine yönelik bir yaklaşımı resmîleştirmek ve standart hale getirmek için birlikte çalışmalara başlamışlardır. Söz konusu çalışmaların sonucu olarak CRISP-DM süreçleri oluşturulmuştur (Nort, 2012: 5). Birbiriyle ilişkili ve belirli bir döngüye sahip olan CRISP-DM süreçleri altı aşamadan oluşmaktadır. Bu aşamalar; işin anlaşılması, verinin anlaşılması, verinin hazırlanması, modelleme, değerlendirme ve yayılmadır (Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer ve Wirth, 2000: 10; aktaran Aytaç ve Bilge, 2013: 29-30). Süreç basamakları takip edilirken, yaşanılacak olan problemlerin çözümü için, bir önceki basamağa geri dönülerek gerekli incelemeler yapılır. Elde edilmek istenen sonuçlara ulaşma konusunda karşılaşılan problem devam etmekte ise sürecin en başına yani işi anlama basamağına ulaşıncaya kadar geri dönme işlemi gerçekleştirilerek incelemelere devam edilir. Süreç basamakları ve birbirleri arasında var olan ilişki Şekil 1.3'te görselleştirilmiştir.

Şekil 1. 3. Veri Madenciliği için Çapraz Endüstri Standart Süreci.



Kaynak: Chapman ve diğerleri, 2010: 10; aktaran Aytaç ve Bilge, 2013: 29.

1.4.1.1. İş Anlama

Süreç ilk olarak problemin ne olduğunun saptanması (Şeker, 2018: 12) ve yapılacak olan işin iyi bir şekilde algılanması ile başlar. Ayrıca bu bölümde, gerçekleştirilecek madencilik işlemi ile elde edilmek istenilen sonuçların ve bu işlemlerden beklentilerin neler olduğunun açık bir şekilde ifade edilmesi gerekmektedir (Aytaç ve Bilge, 2013: 29).

1.4.1.2. Veriyi Anlama

Veriyi anlama sürecinde probleme uygun veriler bir araya getirilir ya da sahip olunan veri kaynağı incelenir. Ayrıca kaynak içerisinde eksik ya da yanlış verilerin var olma durumu incelenir ve verinin türü (yapılandırılmamış, yarı yapılandırılmış, yapılandırılmış) de belirlenir. Eksik ya da yanlış veri tespit edildiği takdirde söz konusu verilerin sisteme eklenmemesi, eksikliklerin tamamlanması ve bu tamamlama işleminin nasıl yapılması gerektiğine dair kararlar bir sonraki basamakta verilmektedir (Şeker, 2018: 12). Sahip olunan verilerin incelenmesi sonucu veriler hakkında elde

edilmiş olan bilgi düzeyi gerçekleştirilecek olan madencilik süreçlerinin başarısı ile doğru orantılıdır (Aytaç ve Bilge, 2013: 29).

1.4.1.3. Veriyi Hazırlama

Veriyi hazırlama, veri kaynağının, üzerinde işlem yapılabilir duruma getirilmesi basamağı şeklinde açıklanabilir (Aytaç ve Bilge, 2013: 29). Söz konusu aşama ile nihai veri kümesini ya da toplanmış olan ilk ham verilerden modelleme araçlarını besleyecek olan verileri oluşturmak için yapılan işlemleri kapsamaktadır (Shearer, 2000: 16).

1.4.1.4. Modelleme

Bu basamak teknik işlemlerin gerçekleştirildiği süreci ifade etmektedir. Uygun modellerin seçilmesiyle birlikte veri seti üzerinde uygulanması basamağıdır. CRISP-DM süreç modellerinde veri hazırlama ve modelleme basamakları veri madenciliği ve metin madenciliği arasındaki en önemli farklılığı oluşturmaktadır (Binici, 2016: 17). Bu farklılık sebebiyle CRISP-DM süreçlerine dayandırılarak oluşturulan metin madenciliği süreçlerinde Miner ve arkadaşları (2012: 76), metin madenciliği süreçleri veri hazırlama ve modelleme basamaklarının metin madenciliğine özel olarak ve daha ayrıntılı bir şekilde tanımlanması gerektiğini belirtmektedir.

Şekil 1. 4. Metin Madenciliği Süreçleri.



Kaynak: Miner ve diğerleri, 2012: 76.

1.4.1.5. Değerlendirme

Oluşturulan modelin yayılım basamağından önce detaylı bir şekilde ele alınması, model oluşturulurken uygulanan adımların gözden geçirilmesi ve işi anlama kısmında belirlenmiş olan hedeflere ulaşılma durumunun kontrol edilmesi oldukça önemlidir (Wirth ve Hipp, 2000: 34). Değerlendirme adımı, önceki adımlar genel

olarak değerlendirilir ve ilk adım olan işi anlama basamağında belirlenen hedeflere ulaşılma durumu incelenir (Şeker, 2018: 12).

1.4.1.6. Yayılım

CRISP-DM süreçlerinin son basamağı olan yayılım adımında, başarılı sonuçlar veren modelin organizasyon ya da kişiler tarafından kullanılması olarak açıklanabilir (Shearer, 2000: 18) Bir organizasyonun internet aboneliğini iptal ettirme ihtimali yüksek olan müşterilerin tespit edilmesi için yaptırdığı bir çalışma sonucunda belirlenen müşterilere yeni paketlerin sunulması örnek verilebilir.

1.4.2. Metin Koleksiyonu Oluşturma

Metin madenciliğinde başarılı sonuçlar elde edilebilmesi için süreç adımlarının doğru bir şekilde uygulanması gerekmektedir. Metin madenciliği yapılırken karşılaşılan problemin iyi anlaşılması, konu hakkında yeterli bilgiye sahip olunması ve elde edilmek istenilen sonuçların belirlenmesi önem taşımaktadır. Burada yapılacak olan tanımlamalar ve sınırlamalar, gerçekleştirilecek olan metin madenciliği çıktılarının başarı oranını doğrudan etkilemektedir. Metin madenciliği süreçleri metin koleksiyonunun oluşturulmasıyla başlamaktadır. Oğuz ve diğerleri (2007: 109) koleksiyon oluşturma sürecini, araştırılmak istenilen konuda bilgi erişim araçları kullanılarak metin koleksiyonu oluşturma süreci olarak tanımlamaktadırlar. Söz konusu tanım doğrultusunda araştırma yapılacak olan konu ile ilişkili veri kaynaklarının bir araya getirilmesi gerekmektedir çıkarımına ulaşılabilmektedir.

Koleksiyon oluşturulması diğer bir ifade ile verilerin toplanması süreci, bir araya getirilen verilerin incelenmesine ve verilerin anlaşılmasına olanak sağlamaktadır. Kısaca “Verinin toplama süreci, veriyi anlama sürecidir” (Göker ve Tekedere, 2007: 293) denilebilir.

Argüden ve Erşahin (2008: 21) veri toplama sürecini;

- a) Başlangıç verilerinin toplanması,

- b) Toplanan verilerin tanımlanarak ihtiyaçların karşılanması durumundaki yeterliliklerinin değerlendirilmesi,
- c) Veri keşfinin yapılması diğer bir ifade ile çalışmanın gerçekleştirilmesi için veri anlamında eksikliklerin tespit edilmesi,
- d) Kullanılacak olan verilerin tam, doğru, hatasız ve eksiksiz olma durumlarının sorgulanarak veri kalitesinin ölçülmesi, şeklinde tanımlamaktadır.

Metin madenciliğinde temel veriler metinlerden oluşmaktadır. Bu kaynaklara erişebilme durumu gelişen teknolojik yenilikler sayesinde giderek kolaylaşmaktadır. Koleksiyonlar oluşturulurken, özellikle metin verilerinin sıklıkla oluşturulduğu ortam olan internet önemli bir veri kaynağıdır. İnternette bulunan veri kaynaklarının türleri farklılık gösterebilir. Bu farklılıklar dikkate alınmadan oluşturulan koleksiyonun içerisine dâhil edilen veriler, zaman kaybı gibi sorunları meydana getirebilmektedir. Facebook, Twitter gibi sosyal medya platformları, Google ve Yandex gibi arama motorları ve ayrıca veri tabanları sahip oldukları büyük veri yığını ile araştırmacıların sıklıkla başvurdukları kaynaklar olarak gözlemlenmektedir.

1.4.3. Metin Ön İşleme

Metin ön işleme, metin madenciliğinin uygulanması ve başarılı sonuçların elde edilebilmesi açısından en önemli adımlarından biri olarak kabul edilebilir. “Bu aşamada veri üzerinde yapılacak işlemlere ve bu işlemlerin hangi yöntemlerle yapılacağına karar verilir” (Şeker, 2018: 12). Gerçekleştirilecek olan işlemlerin detaylı ve eksiksiz yapılması metin madenciliğinin başarıyla uygulanmasına etki edebilecektir. Veri kümesine metin ön işleme adımlarının yapılmaması ya da eksik yapılması durumunun analiz kısmında sorunlarla karşılaşılmasına sebep olabileceği ve bu sorunların çözümü için sürecin tekrar gözden geçirilmesini gerektirebileceği kabul edilebilir. Dolayısıyla söz konusu durumun ciddi zaman kayıplarına sebep olabileceği sonucuna ulaşılabilir.

İlhan (2001: 35), veri ön işleme aşamasının fazla zaman aldığını belirtmektedir. Miner ve diğerleri (2012: 41) veri ön işleme aşamalarının oldukça zaman alıcı olduğunu ve proje süresinin %80 ya da %90'ını aldığını belirtmektedir. Ayrıca Binici (2016: 20) aşamanın zor ve zaman alıcı olduğunu deneyimlerine dayanarak belirtmektedir. İlhan (2001: 35), veri ön işleme aşamasının uygulanmasını;

- a) Veri üzerinde yapılacak olan analizleri engelleyebilecek veri problemlerini çözmek,
- b) Verilerin doğasını anlayarak daha anlamlı ve başarılı bir veri analizi yapmak,
- c) Veri kaynağından bilgi elde edebilmek, gibi nedenlere dayandırmaktadır.

Metin madenciliğinin ilgilendiği veri türü yapılandırılmamış ya da yarı yapılandırılmış verilerdir. Söz konusu verilerden bilgi elde edilebilmesi için bu verilerin makinalarca anlaşılabilir bir yapıya dönüştürülmesi diğer bir ifade ile yapılandırılmış veri haline getirilmesi gerekmektedir (Babadağ, 2021: 49). Kaşıkçı ve Gökçen (2014: 26) ön işleme adımlarını dönüştürme, tarama ve işaretleme, durak kelimelerinin çıkarılması ve kök bulma olmak üzere dört başlık altında ele almaktadır. Literatürde metin madenciliği ile ilgili yapılmış olan çalışmaların incelemesine dayanarak genel itibariyle metin ön işleme aşamaları aşağıdaki gibi açıklanabilmektedir;

1.4.3.1. Dönüştürme ve Temizleme

Metin madenciliğinin uygulanması için toplanmış olan veriler farklı formatlarda bulunabilmektedir. Resim, video ya da ses gibi formatlarda bulunan verileri metin formatına dönüştürmek gerekmektedir (Bayer, 2011: 47). Ayrıca internet ortamında bulunan veri kaynakları çeşitli kodlamalar ya da etiketler barındırdığından (XML, HTML gibi) bu kaynakların da gerekli işlemler sonucunda temizlenerek metin formatına dönüştürülmesi gerekmektedir. Sıklıkla başvuru alan sosyal medya verilerinde metin formatında olmayan yapılarında temizlenmesi

gerekmektedir (duygu belirten ifadeler ve etiketler gibi). Noktalama işaretleri ve metin olmayan farklı karakterler de çıkarılarak veri koleksiyonundan temizlenmelidir (Öztürk, 2021: 49). Bu aşamada dikkat edilmesi gereken husus tüm noktalama işaretlerinin çıkarılmasıyla oluşabilecek sorunların fark edilerek bu duruma uygun karar verilmesidir. Tarih değişkenleri arasında bulunan noktalama işaretlerinin kaldırılması verinin anlaşılmasında problem yaratabilecektir.

Veri koleksiyonu eksik ve gürültülü verilerden de arındırılmalıdır. Eksik kayıt içeren veriler metin içerisinde çıkarılabilmekte, diğer veri kayıtlarının ortalaması eksik veri yerine kullanılabilir ya da makinalarca diğer veri kayıtlarına göre tahmin edilebilmektedir. Gürültülü veriler analiz sonuçlarını olumsuz anlamda etkileyebilmektedir. Bir kişinin yaş bilgisinin eksi değeri olması ya da bir kişinin doğum tarihinin babasının doğum tarihiyle aynı olması mümkün değildir. Bu gibi eksik, tutarsız ya da gürültülü veriler temizlenebilir veya düzeltilebilir (Bayer, 2011: 18) ve analiz sonuçlarının kalitesi artırılabilir.

1.4.3.2. İşaretleme

“Belgeler bölümlere, paragraflara, cümlelere, kelimelere ve hatta hecelere ya da ses birimlerine bölünebilir. Bu işleme işaretleme (tokenization) denilmektedir.” (Feldman ve Sanger, 2007: 60). “En sık rastlanan durum ise kelimelere ayrılmasıdır.” (Haltaş ve diğerleri, 2015: 419). Bu aşamada metin verisi içerisindeki terimlerin ayıklanması ve tüm kelimelere erişilmesi amaçlanmaktadır. Binici (2016: 21), metin verisinin işaretlemede çeşitli yaklaşımlar olduğunu ve bu yaklaşımlar ile ele alınan metinlerden elde edilen kelime yapılarının farklılık gösterdiğini belirtmektedir. Binici’nin (2016: 21-23) bahsettiği söz konusu adımlar;

- **Sözlük ile işaretleme:** Metin verilerinden daha önceden elde edilmiş sözlükler vasıtasıyla işaretleme yapılması,
- **Kelime türü etiketleme:** Gramer olarak benzerlik gösteren kelimelerin türlerine göre gruplandırılması,

- **Çoklu kelime (n-gram):** Metin içerisinde kelimelerin ikili, üçlü ya da dörtlü olarak bir arada geçme sıklıklarının belirlenmesi, şeklinde açıklanmaktadır.

1.4.3.3. Durak Kelimelerinin Çıkarılması

Metin verileri içerisinde sıklıkla kullanılan fakat tek başına bir anlam ifade etmeyen, verilerin sınıflandırılmasında bir işlevi olmayan kelimelerin metinden çıkarılması sürecidir. “Anlama katkısı olmayan kelimelerin vektör uzayında yer alması, hem vektör uzayının boyutlarını artırıp, işlem maliyetini artırır hem de kirlilik yaratarak belge madenciliği çalışmalarının kalitesini düşürür” (Güven, 2007: 78). Metin içerisinde bulunan bağlaçlar (ve, veya, ama, çünkü vb.), zamirler (ben, sen, o vb.), edatlar (ancak, gibi, sadece vb.) tek başlarına bir anlam ifade etmemektedir. Durak kelimeleri (stop words) dillere göre farklılık göstermektedir. “Durak kelimelerin çıkarılması işlemi genellikle işaretleme işleminden sonra gelmektedir” (Kotu ve Deshpande, 2015: 280).

1.4.3.4. Gövdeleme (Kök Bulma)

Metin madenciliği ve veri madenciliği çalışmaları incelendiğinde gövdeleme işleminin kök bulma olarak da adlandırıldığı belirlenmiştir. “Aynı kelime kökeninde olan terimleri bir formatta birleştirmek için yapılan işleme gövdeleme (stemming) denir” (Binici, 2016: 24). Türkçe sondan eklemeli bir dildir. Kelimelere eklenerek mevcut kelimeden yeni bir anlam meydana getiren yapım ekleri ile kelimelere eklendiğinde anlamı değiştirmeyen fakat kelimenin türünü değiştiren çekim ekleri mevcuttur. Kelimeler almış oldukları eklere göre anlam değişikliğine uğramış olması ya da uğramamış olması fark etmeksizin yazılışı değiştiği için farklı bir kelime olarak algılanacaktır. Örneğin, kulak, kulağı ve kulaklık kelimeleri kök bakımından aynıdır fakat metin içerisinde bir arada kullanıldıkları durumlarda ayrı kelimeler olarak değerlendirilecektir. “Söz gelimi, “başarı” kelimesinden türeyen farklı biçimlerdeki kelimeler sıklık listelerinde görülmeyebilir ancak bunların farklı formlarının ayrı ayrı toplanıp bir araya getirildiği durumda en sık kelime “başarmak” fiili olabilir” (Atan, 2020: 231). Bir metinde geçen kelimelerin sıklığı o metnin içeriği hakkında bilgi

vereabilmektedir. Yukarıda bahsedilen durum, yapılacak olan analizlerin dolayısıyla elde edilecek bilgilerin sonucuna etki etmektedir.

Metin içerisinde en az bir kere geçen her farklı kelimenin sayısı, metinleri gösteren vektörlerin boyutunu belirlemektedir (Seçkin, 2011: 18). Kelimelerin gövdeleme işlemi yapılmadan işleme alınması, her bir kelimenin farklı değer olarak algılanmasından dolayı oluşturulacak vektör boyutunun çok büyük olmasına neden olacaktır. Gövdeleme işleminin zorluğu dillere göre değişmektedir. İngilizce kelimeler zaman yapılarına göre farklı kelimeler olarak meydana gelmektedir (begin, began, begun gibi). Türkçe kelimeler ise sonlarına alacakları ekler ile sayısız yeni kelimelerin türemesine olanak sağlamaktadır. Bu sebeple Türkçe metin verilerinde gövdeleme işleminin yapılması zor bir süreçtir denilebilir.

Gövdeleme işlemi iki farklı şekilde yapılabilmektedir Birincisi Joker (Wildkart) yöntemidir. Bu yöntem elektronik kaynaklara erişilirken veri tabanlarında ya da arama motorlarında sıklıkla kullanılmaktadır. Kelimenin köküne (*) işareti koyularak kökten türetilmiş kelimeleri tek bir grup altında toplamayı hedeflemektedir. İkinci adım ise joker yöntemine göre daha katı bir gövdeleme türü olarak kabul edilen köke kadar gövdeleme yöntemidir. Bu yöntem ile kelimenin en yalın halinin elde edilmesi amaçlanmaktadır (Oğuzlar, 2011: 35-37).

Türkçe metinlerde gövdeleme yapılmasına imkân sağlayan, açık kaynak kodlu Zemberek Kütüphanesi mevcuttur. İngilizce metinler için ise başlıca kullanılan iki algoritma mevcuttur. Bunlar “Kuhlen Stemmer” ve “Porter Stemmer” algoritmalarıdır (Thiel, 2009: 6; aktaran Binici, 2016: 25).

1.4.3.5. Terimlerin Ağırlıklandırılması

Terim ağırlıklandırma metin koleksiyonu içerisinde bulunan kelimelerin geçiş sıklığının sayısal değerler ile temsil edilmesi işlemidir. “Yapısal olmayan metinler ağırlıklandırma ile tam olarak yapısal hale dönüştürülmüş olur” (Karaca, Günel ve Taştan, 2015: 809). Kelimelerin metin içerisinde kullanılma sıklığı metin içeriği hakkında bilgi verebilmektedir ve bu durum metinlerin kategorilere ayrılması işleminde kolaylık sağlamaktadır. Literatürde ağırlıklandırma işlemleri için veri

türlerine göre farklı tekniklerin kullanıldığı gözlemlenmiştir². Kullanılan teknikler incelendiğinde genellikle aşağıdaki altı teknik ile ağırlıklandırma işleminin gerçekleştirildiği görülmektedir. Bu yöntemler;

- Bitsel ağırlıklandırma (Binary Weighting),
- Frekansa göre ağırlıklandırma (Weighting by Frequency),
- Terim frekansı (Term Frequency-TF) ve Ters doküman frekansı (Inverse Document Frequency-IDF),
- Ters sınıf frekansı (Inverse Class Frequency-ICF),
- Normalizasyon (Term Frequency Component-TFC),
- Logaritmik terim bileşenidir (Logarithmic Term Component-LTC).

1.4.3.5.1. Bitsel Ağırlıklandırma

Literatürde ikili (binary) ağırlıklandırma metodu şeklinde de ele alındığı görülmektedir. Kelimenin, metin verisi içerisinde geçmiş olması ya da olmaması durumuna bağlı olarak ağırlıklandırmanın yapıldığı yöntemdir (Canhası, 2007: 53). Bu yöntem vasıtasıyla metin verileri 1 ve 0 değerleri ile ifade edilmektedir (Çalış, Gazdağı ve Yıldız, 2013: 3). Eğer kelime metin verisi içerisinde geçiyorsa 1, geçmiyor ise 0 değerini almaktadır. Bahsedilen durum örneklendirilecek olunursa;

Veri: “Bu yıl Erzurum şehri boşanma davalarının az görüldüğü ildir.”

Sözlük: {Yıl, Ay, Boşanma, Evlilik}

Bitsel ağırlıklandırma: {1, 0, 1, 0}

² Söz konusu yargıya Miner ve diğerlerinin (2012: 387-388), Lertnattee ve Theeramunkong’un, (2008: 1146), Toplu’nun (2019: 28), Pilavcılar’nın (2007: 11) ve Seçkin’in (2011: 28) çalışmalarında ele aldıkları ağırlıklandırma tekniklerinin incelenmesi sonucunda ulaşılmıştır.

1.4.3.5.2. Frekansa Göre Ağırlıklandırma (Weighting by Frequency)

Bitsel ağırlıklandırma metodu yukarıda bahsedildiği üzere kelimenin metin içerisinde geçmiş olması ya da geçmemiş olması durumuyla ilgilenmektedir. Fakat frekansa göre ağırlıklandırma söz konusu metin verisi içerisinde kelimelerin kaç defa yer aldığını gösteren bir yöntemdir (İlhan, Duru, Karagöz ve Sağır, 2008: 27). Aşağıda frekansa göre ağırlıklandırma örneği verilmiştir;

Veri: “Erzurum kış aylarında çok soğuktur. Kış şartları şehirde çok ağır geçmektedir”

Sözlük: {Erzurum, Yaz, Soğuk, Kış}

Frekansa göre ağırlıklandırma: {1, 0, 1, 2}

1.4.3.5.3. Terim Frekansı (Term Frequency-TF) ve Ters Doküman Frekansı (Inverse Document Frequency-IDF)

Terim sıklığı (TF) ya da devrik doküman sıklığı (IDF) olarak da literatürde ele alınan bu ağırlıklandırma yöntemi metin madenciliğinde sıklıkla kullanılan sayısal bir temsildir. Örneğin söz konusu terimleri Binici (2016: 25) bu şekilde kullanmıştır. TF değeri bir terimin metin verisi içerisinde kullanılma sayısını ifade etmektedir. DF değeri bir terimin dokümanlar arasında kullanılma sayısını ifade etmektedir. IDF değeri ise bir terimin geçtiği doküman sayısının tespit edilerek terimin ağırlığının tersine çevrilmesini ifade etmektedir (Binici, 2016: 25-26). TF değeri, belirli bir belgede bir anahtar kelimenin görünme sayısının belgedeki toplam terim sayısına oranlanmasıyla elde edilmektedir (Kotu ve Deshpande, 2015: 278). $TF \times IDF$ ağırlıklandırma değeri ise TF değeri ile IDF değerinin çarpılması ile elde edilmiş olur. $TF \times IDF$ bir belgedeki terimin sıklığını (fazla olması iyidir) ve söz konusu terimin tüm koleksiyon içerisindeki nadirliğini (az olması iyidir) vurgulayan sayısal bir değeri hesaplamaktadır (Miner ve diğerleri, 2012: 388).

TFxIDF ağırlıklandırma temel olarak şu düşünceye dayanmaktadır: Bir kelime az sayıda belgede yer alıyorsa, söz konusu kelimenin ayırt edici özelliği yüksektir ve dokümanın kategorisinin belirlenmesinde önemli olduğu kabul edilmektedir. Bir kelime çok sayıda belge yer alıyorsa ayırt edici özelliği azdır ve önemsiz olarak görülmektedir. Örneğin “ve, veya” gibi sözcükler her dokümanda olduğundan ve her bir dokümanda sıklıkla geçtiğinden dolayı TFxIDF değerleri düşük çıkacaktır fakat belgelerde daha sık görülen anahtar kelimelerin TFxIDF değerleri daha yüksek çıkacaktır (Atan, 2020: 235).

1.4.3.5.4. Ters Sınıf Frekansı (Inverse Class Frequency-ICF)

Belirli bir kategori için önemli olan kelime ya da terimlerin diğer sınıflara da ait olabilmesi durumu sebebiyle, söz konusu kelime ya da terimlerin önemini temsil etmek için geçiş sıklığının hesaplandığı ağırlıklandırma yöntemidir (Lertnattee ve Theeramunkong, 2008: 1146). Bu yöntem çok miktarda sınıf ile yapılabilecek çalışmalar için oldukça elverişlidir.

1.4.3.5.5. Normalizasyon (Term Frequency Component-TFC)

Veri koleksiyonu içerisinde bulunan metinler uzunluk ya da kısalık bakımından farklılık gösterebilmektedir. Uzun bir metinde kullanılan kelimelerin sıklıkları kısa olan metinlere göre daha fazla olabilmektedir. Bir metnin uzun ya da kısa olması söz konusu metnin öneminin ya da konuyla alakalı olma durumunun bir göstergesi olduğu sonucuna varılamaz. Bu sebep doğrultusunda uzun ve kısa metinler ağırlıklandırılırken kelimelerin frekans değerlerine ek olarak metin içerisinde geçen toplam kelime sayısı da göz önüne alınmaktadır (Kesgin, 2007: 32; Adsız, 2006: 29; aktaran Seçkin, 2011: 29-30). Bu durum TFxIDF ağırlıklandırılması ile normalizasyon yöntemi arasındaki farkı oluşturmaktadır.

1.4.3.5.6. Logaritmik Terim Bileşeni (Logarithmic Term Component-LTC)

Normalizasyon yönteminin farklı bir versiyonu olarak kabul edilmektedir. Frekanslardaki etki seviyesi yüksek değişikliklerin etkisinin azaltılması için ham

frekans yerine logaritma kullanılması esasına dayanmaktadır (Adsız, 2006: 29; aktaran Seçkin, 2011: 30).

1.4.4. Özellik Seçimi

“Özellik seçimi (Feature Selection), büyük boyutlu veri kümesini, daha küçük boyutta temsil edebilecek nitelikte olan bir alt küme olarak tanımlanmaktadır” (Haltaş ve diğerleri, 2015: 420). Bu aşamada metin verileri içerisinde bulunan önemli kelimeler belirlenmektedir. Veri kaynağı içerisinde birçok özellik bulunabilmektedir. “Özellik seçimi, ilgilenilen problem için en faydalı ve en önemli özellikleri seçerek veri kümesindeki özellik sayısının azaltılmasını amaçlamaktadır” (Budak, 2018: 22). Ladha ve Deepa (2011: 1788), özellik seçiminin avantajlarını şu şekilde açıklamaktadır:

- Depolama gereksinimlerini sınırlamak ve algoritma hızını artırmak için özellik alanının boyutunu azaltmaktadır.
- Gereksiz ya da ilgisiz veriler ile gürültülü verileri kaldırmaktadır.
- Veri analizi görevlerini anlık uygulaması, öğrenme algoritmalarının çalışma süresini hızlandırmaktadır.
- Veri kalitesini artırmaktadır.
- Ortaya çıkan modelin başarısını artırmaktadır.
- Tekrar veri kümesi oluşturulmak istendiğinde kaynaklardan tasarruf edilebilmesi için özellik kümesini azaltmaktadır.
- Doğru tahminler elde edebilmek için performans iyileştirmektedir.
- Veriyi anlamayı, veriyi oluşturan süreç hakkında bilgi edinilmesini ve basit bir şekilde veriyi görselleştirmeyi mümkün kılmaktadır.

Özellik seçme algoritmaları üç alt sınıfta gruplandırılabilir. Bunlar filtre yöntemi, örtü yöntemi ve gömülü yöntemdir (Hamde, 2018: 34).

1.4.5. Veri Madenciliği

Metin madenciliği yapılandırılmamış ya da yarı yapılandırılmış veriler ile ilgili iken veri madenciliği yapılandırılmış veriler ile ilgilenmektedir. Yukarıda bahsedilen süreçler verinin yapılandırılmış bir hale getirilmesi için uygulanan çeşitli adımlardan oluşmaktadır. Bu adım, veri koleksiyonuna uygulanan söz konusu işlemler ile yapılandırılmış hale getirilen metin kümeleri üzerinde veri madenciliği yapılması sürecini kapsamaktadır. Veri madenciliği gerçekleştirilirken birçok teknikten (karar ağaçları, yapay sinir ağları, kümeleme vb.) faydalanılmaktadır (Oğuz ve diğerleri, 2007: 110).

1.4.6. Görselleştirme

Bu aşamada temel amaç verinin daha anlaşılabilir kılınmasıdır denilebilir. Görselleştirme aşamasından önce yukarıda bahsedilmiş olan adımlar veri kümesine uygulandıktan sonra birçok sonuç elde edilebilmektedir. İncelenecek sonuçların miktar olarak fazla olması ve inceleyecek kişilerin bu sonuçların ne anlam ifade ettiğini anlayamaması durumu mümkün olabilmektedir. Görselleştirme adımı, veri kümesinden elde edilen sonuçların daha etkili ve anlaşılır bir biçimde sunulması işlemi olarak açıklanabilmektedir. Metnin hacmi ne olursa olsun görselleştirme işlemi, elde edilmiş olan sonuçlardan tek bakışta kolaylıkla anlam çıkarılabilmesine imkân sağlamaktadır. Kullanıcıların gereksinimlerine, zaman metaforuna ve gerçekleştirilen çalışmanın amacına göre farklı görselleştirme teknikleri bulunmaktadır (Binici, 2016: 35). Uygun teknik ile görselleştirme yapılması önemlidir. Belirli bir zaman içerisinde gerçekleşmiş olan değişimler ile ilgili veriler görselleştirilirken dairesel diyagramların kullanılması verinin anlamlandırılması için yardımcı olmayacaktır. Bu verilerin sütun grafiği ile (histogram) görselleştirilmesi verileri daha anlamlı kılacaktır.

1.4.7. Değerlendirme

Değerlendirme aşaması metin madenciliği süreçlerinin son aşaması olarak kabul edilebilir. Bu aşama, veri üzerinde yapılan işlemler sonucunda ortaya çıkan örüntülerin (pattern) yorumlandığı ve bilgiye dönüştürüldüğü, elde edilen çıktıların

değerlendirildiği aşama olarak kabul edilebilir (Şeker, 2018: 16). Süreç adımlarının başarısı da bu aşamada ölçülmektedir denilebilir. Eğer istenilen düzeyde bilgi çıkarımı elde edilememişse süreç tekrar yeniden uygulanabilir. Süreç adımlarının tekrarlanması zaman kaybına neden olacaktır. Söz konusu problem ve meydana gelebilecek farklı diğer problemlerin yaşanmaması adına yukarıda bahsedilen adımların titizlikle ve konu uzmanları ile birlikte gerçekleştirilmesi gerekmektedir.



İKİNCİ BÖLÜM

AHKÂM DEFTERLERİ

2.1. Tanımı ve Tarihsel Gelişimi

Literatürde hüküm kelimesinin değişik manalarda tarifi yapılmaktadır. Şemseddin Sami'nin Kâmûs-ı Türkî (2011: 436) adlı sözlüğünde söz konusu kelimenin yedi farkı anlamda kullanıldığı belirtilmiştir. Bu anlamlar kelimenin hem masdar (hüküm) hem de isim (hüküm) halinde kullanımlarını içermektedir. Kâmûs-ı Türkî'de hüküm kavramının ilk olarak hâkimlik, hükûmet ve âmiriyyet anlamlarına karşılık geldiği “filânın hükmü altındadır.” örnek cümlesi ile verilmektedir. Ardından ikinci olarak emir, irâde ve kumanda anlamlarına geldiği “hükmetmek, hükmü geçmek, hükmü nâfiz olmak, filânın hükmüyle hareket ediyor” örnek ifadeleri ile belirtilmektedir. Ayrıca bir dava ya da meselenin iyice anlaşılması ve değerlendirilmesi sonucu nihaî olarak verilen “karar” anlamında da kullanıldığına yer verilmektedir. Kuvvet, ehemmiyet ve nüfûz anlamlarını ifade ederken kullanılabildiği “onun hükmü bu kadar” örnek cümlesi ile açıklanmaktadır. Te'sîr manasında kullanılabildiği ise “kışın hükmü geçti” cümlesi ile örneklendirilerek belirtilmektedir. Menzile, mesâbe ve makâm anlamlarını barındırması durumu “kayınpeder dahi peder hükmündedir” cümlesi ile örneklendirilerek açıklanırken iktizâ ve îcab anlamlarına geldiği de ifade edilmektedir.

Hüküm kavramı hakkında yapılan diğer tanımlamalar incelendiğinde hüküm kelimesinin hususi anlamda devletin ve hükümetin otoritesini ve gücünü ifade ettiğini belirten Özel (1988: 550), hâkimin belli bir konu hakkında verdiği karar anlamını barındıran, iki nesne ya da fikir arasında kurulan bağlantı ve bir konu hakkında mantık çerçevesinde verilmiş olan karar anlamlarında kullanılan kapsamlı bir kavram şeklinde tanımlamıştır. Üzüm (1998: 464) ise kavramın kelâm, mantık ve fıkıh ilimlerinde kullanılan bir terim olduğunu belirtmiştir. “İyileştirmek amacıyla menetmek, düzeltmek: karar vermek” manalarında kullanıldığını ifade ederek “ilim, derin anlayış; siyâsî hâkimiyet, karar ve yargı” anlamlarına geldiğini de açıklamıştır. Üzüm (1998:

464), bu kavramı, Özel (1988: 550)'in de yukarıda tanımladığı gibi, “iki şey, iki düşünce arasında olumlu veya olumsuz bir bağlantı kurmaktır” şeklinde tanımlamış ve hatta söz konusu bağın idrak edilmesine hüküm denildiğini belirtmiştir. Bu tanım, kavramın mantık ilmine göre tanımlamasını ifade etmektedir. Üzüm (1998: 464), hüküm kavramının Kelâm ilminde “İslâm dininin inanç, ibadet, muamelât ve ahlâka dair temel ilkelerini” ifade ettiğini belirtmiştir. Fıkıh ilminde ise “mükelleflerin fiilleriyle ilgili ilâhî hitaplara” hüküm denildiğini açıklamıştır.

Uzunçarşılı (1998: 280) ise hüküm kavramının kat'î emir ve karar anlamına geldiğini belirterek “her hangi bir iş ve vazife için pâdişah tarafından verilen tahrirî emre Hükûm-i hümayûn denilir; Dîvân-ı Hümayûndan ve maliyeden, darphane ve saireden hükümdar namına yazılan emir ve tenbihâtın hepsi” şeklinde açıklamıştır. Ahkâm kelimesi ise hükümün çoğuludur ve “emirler, hükümler” anlamına gelmektedir (Devellioğlu, 2008: 24). Bir başka tanımda “Ahkâm, daha geniş mânada, belli bir konu hakkında konulmuş bulunan kuralların bütününe ifade etmek için kullanılır” (Özel, 1988: 550) şeklinde açıklanmaktadır. Ahkâm defterleri ise, isminden anlaşılabacağı üzere, ahkâmın bir araya getirilmesi sonucu oluşan defter formlarına verilen addır. Pakalın'a (2004: 30) göre “Ahkâm defteri, Kanunnamelerle hükümlerin ve nizam mahiyetinde olan kararların aynen kaydına mahsus defterin adı”dır. Ahkâm defterleri hakkında yapılan bir diğer tanımda ise kavram “Osmanlı Devleti'nde genellikle maliyeye ait hükümlerin toplandığı defterlere verilen ad” şeklinde ifade edilmiştir (Sahillioğlu, 1988: 551).

Osmanlı İmparatorluğunda halk yaşamış olduğu sorunların çözüme kavuşturulması için padişaha ya da padişahın görevlendirdiği kişi ve kurumlara başvurabilme hakkına sahip olmuştur. Bu başvurular Dîvân-ı Hümayûn'da görüşülerek sorunun çözümüne dair sonuca bağlanmıştır ve verilen kararların birer suretleri eyaletlere göre oluşturulan Ahkâm defterlerine kayıt edilmiştir (Günay, 2013a: 9). Hükümler, muhataplarına bizzat padişah tarafından iletilmemiş bunun yerine padişahın yetki verdiği makamlar ve daireler tarafından işleme koyulmuştur. Sadâret makamından ya da defterdarlıktan çıkan hükümler (Sahillioğlu, 1988: 551), oluşmalarına neden olan durumlara göre ferman, nişan berat, tevki ve menşur şeklinde adlandırılmıştır (Uzunçarşılı, 1998: 280). İdarî ve hukukî açıdan resmî yazışmaların

birer suretlerinin saklanması gerekmektedir. Bu gereklilik sonucu resmî devlet arşivlerinin oluştuğu söylenebilir (Aydın, 2003: 22). Hükümlerin bir kopyasının tutulması ve söz konusu hükümlerin bir defter formunu oluşturması bahsedilen gereklilik ile açıklanabilmektedir. Oluşan bu defterler isimlerini, hükümlerin alakalı olduğu konulara göre almışlardır. Örneğin malî işler ile ilgili bir hüküm verilmiş ise söz konusu hükmün kayıt edildiği deftere maliye ahkâm defterleri denilmiştir (Uzunçarşılı, 1998: 280).

Dîvân-ı Hümâyûn, Orhan Gazi döneminden itibaren (1326-1359) varlığını sürdüren, Fatih Sultan Mehmed döneminde (1451-1481) ise müessese haline geldiği kabul edilen ve Osmanlı İmparatorluğu'nun klasik çağını¹ kapsayan dönemlerde yetkileri itibariyle padişahattan sonraki en yetkili makamdır (Sarıcaoğlu, 2007: 87). Dîvân-ı Hümâyûn, kişilerin kadın ya da erkek olması, mensup olduğu dini, dili, yaşı ve sosyal statüsü fark etmeksizin başvurabildiği ve devletin tüm önemli makamlarında bulunan kişilerin katılarak padişah adına karar verdikleri, devletin idarî, siyasî ve adlî konuların görüşüldüğü karar organıdır (Mumcu, 1994: 430-431). Dîvân-ı Hümâyûn'un temel görevi adaleti sağlanmasıdır ve söz konusu adalet kavramı, halkın şikâyetlerini padişaha iletebilmesi ve var olan haksızlıkların padişahın emri sonucunda giderilmesidir (İnalcık, 1988: 33). Bu bilgiler çerçevesinde Dîvân-ı Hümâyûn yetki ve sorumlulukları sebebi ile devletin gücünü ve otoritesini temsil eden bir kurumdur sonucuna varılabilmektedir. Dîvân-ı Hümâyûn toplantılarında ele alınan siyasî, askerî, malî ve idarî konular hakkında verilen kararların kronolojik bir düzen içerisinde kaydedildiği defterlere Mühimme Defteri adı verilmiştir (Yıldırım, Atik, Cebecioğlu, Demirbaş, Karazeybek, Safi, Serin, Uslu, Yekeler ve Yıldırım, 1994: ix-x). Fatih Kanunnâmesi ile gerçekleştirilen anayasal ve bürokratik düzenlemeler sonucu tam anlamıyla gelişmeye başlayan Dîvân-ı Hümâyûn'da (Mumcu, 1994: 430) hükümler,

¹ Literatürde Osmanlı İmparatorluğu'nun Klasik Çağ'ına doğrudan ve ismen işaret eden eserlere tesadüf etmek mümkündür. Merhum Halil İnalcık tarafından kaleme alınan *Osmanlı İmparatorluğu Klasik Çağ (1300-1600)* (2019) isimli kitap buna önemli bir örnektir. Bir başka örnek vermek gerekirse Feridun M. Emecen'in *Osmanlı Klasik Çağında Siyaset* (2009), *Osmanlı Klasik Çağında Savaş* (2010), *Osmanlı Klasik Çağında Hanedan, Devlet ve Toplum* (2011) ve *Osmanlı Klasik Çağında Hilafet ve Saltanat, Aktörler, Yazım ve İmaj* (2020) isimli kitapları belirtilebilir. Klasik çağ kavramı Osmanlı İmparatorluğunun kuruluş ve yükseliş dönemleri olarak adlandırılan, XIV. yüzyıldan başlayıp XVI. yüzyılın sonlarına kadar devam eden süreci kapsayan döneme verilen bir diğer addır (Başar, 2003: 25).

Beylikçi kalemi tarafından düzenlenen defterlere geçirilmiş ve bu işlemi yapan kişilere ise Mühimme Nüvisan denilmiştir (Yıldırım ve Hayırlıoğlu, 2019: 208).

Şikâyet Defterleri adıyla oluşturulan yeni defter türü, Osmanlı bürokrasisinden çıkan bütün ferman kayıtlarının H.1059/M.1649 senesinden itibaren Mühimme Defterlerine kaydedilmemesi sonucunda meydana gelmiştir. (Şimşir, 1994: 359). Mühimme Defterlerinin yanı sıra Ecnebî Defterleri, Şikâyet Defterleri, Nâme Defterleri, Ahkâm Defterleri ve sonraları Nişân Defterleri gibi konularına göre düzenlenen farklı yeni defter serilerinin de oluştuğunu belirten İnalçık (1988: 34), bu durumu farklı konulara sahip söz konusu defterleri meydana getiren kâtiplerin uzmanlaşmasıyla ve Osmanlı İmparatorluğu'nda meydana gelen bürokratik gelişmelerle ilişkilendirmiştir. Emecen (2005: 115-117), Mühimme Defterlerinin ilk başta bu isimle anılmadığını belirtmektedir. Söz konusu defterlerin başında bulunan ifadelerle göre adlandırıldığını, 17. yüzyılın ortalarına doğru şikâyet ahkâmlarının ayrı bir tür olarak ele alınması sonucunda Mühimme adının oluştuğunu ifade etmiştir. İçerik ve şekil bakımından farklı olan bu iki defter türünün Osmanlı kalemiye teşkilatındaki anlayış değişikliklerinin bir yansıması olduğunu belirtmiştir. Dîvân kaleminde bulunan kişilerin işlerini kolaylaştırmak ve daha pratik hale getirmek için başta bulunan bürokratin kişisel tercihleri sonucunda oluştuğuna da değinmiştir. Şikâyet Defterleri de aynı şekilde Mühimme Defterleri gibi Beşikçi Kalemi tarafından yazıldığından, tür, biçim, ebat ve defterlere düşülen kayıtların iç düzenleri konusunda benzerlik göstermektedir. Mühimme ve Şikâyet Defterlerinin ayırt edilebilmesi için içeriklerinin değerlendirilmesi daha belirleyici olacaktır. Mühimme Defterlerinde devletin askerî, siyasî ve idarî konularıyla ilgili hükümler yer alırken, Şikâyet Defterlerinde toprak anlaşmazlığı ve alacak verecek gibi sosyal konular yer almaktadır (Şimşir, 1994: 359). Şikâyet Defterlerinde yer alan hükümler kamuyu ilgilendiren şikâyetlerden ziyade kişisel sorunlar sonucunda başvuru konularının çözümlerini içermektedir. Kamuya ait işlemler ise Mühimme Defterlerine kayıt edilmiştir (İnalçık, 1988: 36).

Osmanlı İmparatorluğu'nun klasik çağı olarak adlandırılan kuruluş ve yükseliş dönemleri incelendiğinde, söz konusu dönemde meydana gelmiş önemli siyasî konular olarak kabul edilen İstanbul'un fethi, İran ve Mısır seferleri ve Akdeniz'in Türk gölü

haline getirilmesi (Başar, 2003: 25), İmparatorluk sınırlarının genişlediğine ve bu duruma bağlı olarak nüfusun arttığına yönelik çıkarımların yapılmasına olanak sağlamaktadır. Kanuni Sultan Süleyman'dan sonraki dönemlerde Osmanlı İmparatorluğu'nun "merkezi otoritesi zayıflamaya başlamıştır." (Şimşir, 1995: 329). Bu duruma neden olan etkenleri Kodaman (2007: 13) şu şekilde ifade etmiştir: İmparatorluğun fütuhât döneminin sona ermesi, ticaretin yeni güzergâhlar üzerinden sağlanması ve dolayısıyla ipek yolunun önemini yitirmesi ayrıca Amerika'dan ithal edilen altın ve gümüş gibi değerli madenlerin Avrupa üzerinden gelmesi, bu yüzden maliyetlerin artması ve ihraç edilen ürünlerin fiyatlarının düşmesi gelir kaynaklarını, dolayısıyla da hazineyi olumsuz etkilenmiştir. Tüm bu durumların yanı sıra fetihler sonucu artan bürokrasinin maaş yükü, güvenliğin sağlanması adına daimî olarak düzenli ordu bulundurma gerekliliği ve harp masrafları da hazineyi olumsuz etkileyen sebeplerdendir. Söz konusu olumsuz etkilerden kurtulmak için vergileri artırma, paranın değerini düşürme ve maaşları azaltma yöntemine gidilmiştir. Bu durum İmparatorluk mensupları arasında huzursuzlukların meydana gelmesine neden olmuştur. Söz konusu durumların merkezi otoriteye duyulan güvenin ve sadâkatin azalmasına ve dolayısıyla doğrudan ya da dolaylı olarak İmparatorluğun adlî, idarî, siyasî ve iktisadî yapısında bozulmalara neden olduğu çıkarımı yapılabilmektedir. Yukarıda bahsedilen nüfus artışı ve devlet otoritesinde yaşanan çözümler gibi durumların etkisiyle şikâyetler artmıştır ve şikâyetlerin hızlı bir şekilde çözülmesi ihtiyacı ortaya çıkmıştır. Bürokratik gecikmelerin önlenmesi ve sorunların daha hızlı bir şekilde çözüme ulaştırılması amacıyla Reîsülküttâb Râgıp Efendi tarafından mahallî sistem getirilmiştir (Öztürk, 2001: 615). Şikâyet ve dilek niteliğinde olan konularla alakalı hükümlerin tutulduğu Şikâyet Defterleri 1742 tarihinden itibaren eyaletlere ayrılarak tutulmuştur ve böylece Ahkâm-ı Şikâyet ya da Ahkâm Defterleri adı verilen yeni bir seri meydana gelmiştir (Temelkuran, 1975: 156). Oluşan yeni seriyi "Eyâlet Ahkâm Defteri" olarak adlandıran Emecen (2005: 125), Râgıp Efendinin mahalli sistemi getirme sebebini "mahallinde halledilemeyip merkeze aksettirilen problemlerin çoğalması ve artan bürokratik muameleatı kolaylaştırmak için ihdas edilmiş olmalı" şeklinde yorumlamıştır. Literatürde Ahkâm Defterlerinin belirli bir bölgeyi kapsamına karşılık Şikâyet Defterlerinin ülkenin tamamını kapsadığı, bu sebeple söz konusu defterlerin birbirlerinden farklı iki tür oldukları belirtilmiştir.

Ayrıca Ahkâm Defterleri şekil bakımından Şikâyet Defterlerinden daha dar ve uzun olması (Keleş, 2020: 24), bu defterleri birbirlerinden ayırt eden bir özellik olarak kabul edilmektedir. Şikâyet Defterlerinin varlığını devam ettirmesi durumu ise Ahkâm Defterleri bulunmayan bölgeleri kapsamı ile açıklanmıştır. Fakat Günay (2013a: 17) çalışmasında, Ahkâm Defterlerinin yazıldığı yerlerde, daha sonraki bir defterin Şikâyet Defteri adı altında kaydedildiği durumların tespit edildiğini belirtmiştir. Defterlerin kapak kısmına ya da içerisine o defterin nereye ait olduğunun, yılının ve türünün yazıldığını ve bazı defterlerin türlerinin Ahkâm-ı Şikâyet gibi kelimeler ile belirttiğini ifade etmiştir. Dolayısıyla Ahkâm Defteri olarak bilinen defterlerin, “Ahkâm-Şikâyet Defterleri” olarak da adlandırılabileceği sonucuna varmıştır.

2.2. Ahkâm Defterlerinin Önemi

Ahkâm Defterlerinde yer alan hükümler, genel itibariyle defterlerin tutulduğu bölge halkının şikâyetçi oldukları konular hakkında, padişaha ya da yetkililere başvuruları sonucunda görüşülen konuları içermektedir (Günay, 2013a: 23). Bu noktadan hareketle bölge halkının yaşadıkları sorunları içeren bu defterler söz konusu bölge hakkında başvurulabilecek bilgi kaynaklarından denilebilir. Bölge ile ilgili sosyal ve ekonomik konularda bilgilerin elde edilebileceği gibi idarecilerin kendi aralarında ya da idarecilerin halk ile aralarında var olan sorunlarının da anlaşılmasını sağlayabilmektedir (Emecen, 2005: 125). Ayrıca halkın birbirleri arasında meydana gelen problemler ile söz konusu problemlere getirilen çözümlerin neler olduğunun da incelenmesine olanak sağlamasının yanı sıra, kanunların pratikte ne kadar uygulandığı ve uygulanabilirlik durumu, hukuksal eksiklikler ve bu eksikliklerin meydana getirdiği yeni problemler hakkında bilgi vermesi bakımından oldukça önemlidirler (Öztürk, 2001: 621- 622).

Devletin bir yerde kendi gücünü hissettirmek amacıyla adaleti tesis etmesi, padişahın ve onun nezdinde bölge idarecilerinin en önemli vazifelerindendir. Osmanlı İmparatorluğu’nun kuruluşundan itibaren adalet kavramına önem verilmiş ve bu kavram, sonraki nesillere de aktarılmıştır. Adalet kavramı içerisinde dil, din ve ırk ayrımı gözetmeksizin devletin sınırları içerisinde yaşayan reayanın herhangi bir olay karşısında sıkıntılarını padişaha iletme hakkı bulunmaktaydı. Halk şikâyetlerini

padişaha Dîvân-ı Hümâyûn vasıtasıyla, padişahın camiye gidişinde ya da dönüşünde (İpşirli, 1993: 90-91), avlanmaya çıktığında veya sefere giderken ya da geri gelirken bizzat kabul etmesi sonucunda iletebilmekteydi (İnalcık, 1988: 33). Böylelikle devlet, sınırları içerisindeki taşrada yaşayan reayanın sıkıntısını tespit edebilme ve adalet sisteminin işlerliğini de kontrol edebilme imkanına kavuşmuş olmaktaydı. Bu durumun önemli bir diğer yanı da reayanın sıkıntılar karşısında kendilerini çaresiz hissetmelerinin önüne geçilmesidir. Bu uygulama, hem reayanın güçlü karşısında ezilmediğini hissetmesi hem de sosyal adaletin sağlanması açısından önemliydi. Reayanın rahatlıkla şikâyetlerini merkeze iletmesinden dolayı ayrıca yöneticilerin denetlenmesi de sağlanmış oluyordu (Günay, 2013a: 23-24).

Resmî bir yazının ya da belgenin içeriği, belgenin oluşturulduğu devlet hakkında birçok bilgiyi barındırmaktadır. Bu belgelerden söz konusu ülkenin iç ve dış politikası, sosyal, kültürel, ekonomik yaşantısı, dini inançları, hukukî ve askerî yapısı gibi çoğu alanla ilgili metin içerisinde gömülü ya da gizli kalmış bilgilerin çıkarımı yapılabilmektedir (Özdemirci ve Torunlar, 2015: 17). Ahkâm Defterleri, devletin yürütme organı olan Dîvân-ı Hümâyûn kararlarını içermektedir ve dolayısıyla devletin resmî düşüncesini göstermesi hususunda önemli bir yansıtma aracı olduğu ifade edilebilir. (Günay, 2013a: 24).

Şimşir (1994: 363) çalışmasında Ahkâm Defterlerini, varlığını günümüze sürdürememiş mimari yapılar ile vakıflar hakkında yapılacak olan çalışmaların yanında iktisat tarihi ve hukuk gibi insan ve insanın yaptıklarını konu alan araştırmacıların başvurabilecekleri bir kaynak olarak göstermektedir.

Günümüzde verilerin meydana gelmesi daha kolay ve daha hızlıdır. Bu durumun en önemli sebebinin gelişen bilgi teknolojileri ve bu teknolojilerin internet ile birlikte yaygın bir şekilde kullanılması olduğu kabul edilebilmektedir. Bilgi teknolojilerindeki gelişim sonucunda üretilen araç ve yöntemler vasıtasıyla verilerden faydalı bilgiler elde edilmesi mümkün olmuştur. Bilgi ve belge merkezlerinden olan arşivler ve kütüphaneler, yeni ve faydalı bilgilerin elde edilebilmesi için büyük miktardaki veri kaynağını bünyesinde barındırmaktadır. Çalışmada ele alınan Ahkâm Defterleri, çok sayıda hüküm verisinden oluşması, muhtevası bakımından ise kayda

alındıkları bölgenin sosyal, kültürel, ekonomik ve demografik yapısı hakkında bilgi elde edilebilmesini mümkün kılmasının yanı sıra devletin hukukî, iktisadî, idarî ve siyasî yapısı hakkındaki bilgileri erişilebilir kılması açısından büyük önem arz etmektedir. Söz konusu bilgilere erişilebilmesinin yanı sıra metin madenciliği ve veri madenciliği uygulamaları ile çeşitli, doğru ve hızlı çıkarımlara varılabilmesini ve elde edilen çıktılar üzerinde disiplinlerarası çalışmalar sonucunda faydalı bilgilere ulaşma durumunu mümkün kılabilecektir. Bu durum şu şekilde örneklendirilebilir; Ahkâm Defterlerinin yazılmasından sorumlu Beylikçi kalemindeki görevlilerin üslup analizi yapılabilir, böylelikle farklı görevlilerce kırık divanî yazı biçimiyle hızlı bir şekilde yazılan ve okunması zor olan hükümlerde, okunamayan kelimelerin hangi görevli tarafından yazıldığı belirlenerek daha başarılı transkripsiyon işlemleri gerçekleştirilebilecektir. Bunun yanı sıra söz konusu görevlilerin kullandıkları kelimelerin zaman içerisinde bir değişim gösterip göstermediği incelenerek filoloji bilimi bakımından dilin gelişimi ve değişimi üzerinde yapılabilecek çalışmalara kaynaklık edebilecektir.

2.3. Ahkâm Defterlerinin Diplomatik ve Fiziki Özellikleri

Ahkâm Defterlerinin kapak kısmında, o defterin hangi bölgeye ait olduğu yazılmaktadır. Ardından defteri meydana getiren kayıtların onlu sisteme göre [(evâil (ayın başı), evâsıt (ay ortası), evâhir (ay sonu))] ay ve yıl bilgileri belirtilerek tarihleri yazılmaktadır (Günay, 2013a: 16-17). Hükümler yazılmadan önce söz konusu hükmün hangi Sadrazam ve Reîsülküttâb zamanında yazıldığına bilgisi verilmektedir. Sadrazam ve Reîsülküttâb değişmiş ise isimleri deftere kaydedilerek hükümlerin yazımına devam edilmektedir fakat bahsedilen makamlarda bir değişiklik yok ise “Müşârün-ileyh” tabiri kullanıldıktan sonra hükümler yazılmaktadır (Çiftçi, 2020: 13). Söz konusu defterin türü belirtilirken “.... hazâ defter-i kuyûd-ı ahkâm-ı şikâyet/ahkâm-şikâyet”, “... hazâ defter-i şikâyet...”, “.... hazâ defter-i kuyûd-ı ahkâmü’ş-şikâyât...” gibi farklı ifadeler kullanılmaktadır (Günay, 2013a: 17). Bu ifadeler doğrultusunda Şikâyet Defterleri ile Ahkâm Defterlerinin arasında keskin bir farkın bulunmadığı sonucu çıkmaktadır. Literatür incelendiğinde Mühimme Defterleri, Şikâyet Defterleri ve Ahkâm Defterleri temelde birbirinin devamı

niteliğindedir şeklinde yorum yapılabilir. Bu yorum, bahsedilen üç defter türünün bürokraside gerçekleşen ihtisaslaşma, konuyu ele alma ve çözüme ulaştırma ve hatta gerekirse tekrar başvurma gibi işlemleri hızlı ve kolay kılması bakımından öncelikle konularına göre ve ardından mahalli sisteme göre ayrılarak kayıt edilmesi durumuna dayandırılmaktadır. Yukarıda defterin türü belirtilirken Ahkâm ve Şikâyet isimlerinin birlikte kullanılmasının nedeninin bu şekilde açıklanabileceği düşünülmektedir.

Ahkâm Defterlerini oluşturan kayıtlara gerek görüldüğü takdirde tekrar müracaat edilebildiğine daha önce de değinilmiştir. Dîvân-ı Hümâyûn'da görüşülen kararların kaydedilmesi sonucu meydana gelen bu defterler dîvan tutanağı niteliğine sahiptir (Çelik, 2020: 30). Kanıt değeri taşıyan ve kurumsal bir başvuru kaynağı olarak da arşivsel değere sahip olan bu defterlerin muhafaza edilmesi önem arz etmektedir. Defterlerde geçen “.... Hazîne-i Âmire’de mahfûz...” veya “... Defterhâne-i Âmire’de mahfûz ... defterine müracaat olundukda ...” (Günay, 2013a: 18) şeklindeki ifadeler yukarıda bahsedilen durumu destekler nitelikte olmakla beraber, Ahkâm Defterlerinin Hazîne-i Âmire ve Defterhâne-i Âmire’de muhafaza edildiği ve defterlere tekrar başvurulduğu sonuçlarına ulaşılmasını da mümkün kılmaktadır.

Dîvân-ı Hümâyûn'da ele alınan meselelerin kayıt edilmesinde kullanılan yazı çeşidine “divanî” denilmektedir ve bu yazı şekli süslü, gösterişli ve resmî bir niteliğe sahiptir (Gökbilgin, 1992: 44). Ahkâm Defterlerini oluşturan şikâyetlerin Dîvân-ı Hümâyûn'da görüşüldüğü düşünüldüğünde, bu defterlerin de divanî yazı türüyle yazıldığı sonucuna ulaşılması mümkün olmaktadır. Divanî yazı türü, siyakat ve rik’a yazı türleri ile birlikte Osmanlı arşiv belgelerinde en sık karşılaşılan yazı türüdür ve söz konusu yazı türleri Osmanlılar tarafından geliştirilmiştir (Kütükoğlu, 1998: 68). Fakat Kütükoğlu (1998: 63), arşiv belgelerinde kullanılan yazıların tam anlamıyla divanî olmadığını ve “kıрма” ya da “hurda” adı verilen yazı şekilleriyle yazıldığını ifade etmektedir. Ayrıca süratli bir şekilde yazma isteğinin, yazının noktalama işaretlerinin kullanılmamasına sebep olduğunu ve bu durumun söz konusu yazının okunulmasını zorlaştırdığını belirtmektedir.

Şahin ve Emecen (1994: v), II. Bâyezid dönemine ait olan ve Haziran- Temmuz 906/1501 tarihli divan kayıtlarını kapsayan Ahkâm Defterini inceledikleri çalışmalarında, çeşitli bürokratik süreçler sonucunda hazırlanmış olan Dîvân-ı Hümâyûn kararlarının ve çıkan hükümlerin ilgililere gönderilmeden önce bir sûretlerinin defter halinde kronolojik olarak kaydedildiğine değinmiştir. Ahkâm Defterlerine yazılmış olan hükümler, bir düzen oluşturması ve oluşturulan düzenin korunması amacıyla tarihi sıralamaya yani olayların kronolojik sıralamasına önem verilerek kaydedilmiştir. Nüfusun yoğun olduğu bölgelerde şikâyetlerin ve dolayısıyla söz konusu şikâyetlere gelen cevapların sayısının yüksek, nüfusun az olduğu bölgelerde ise tam tersi bir durumun olması muhtemeldir. Günay (2013b: 293) çalışmasında iş yükünün fazla olduğu yerlerde tutulan defterlerin daha kısa tarih aralığını kapsadığını, iş yoğunluğunun az olduğu bölgelerde ise farklı senelere ait daha fazla hükmün bir araya getirildiğini belirtmiştir. Dolayısıyla bu bölgelere ait defterlerin kapsadığı tarih aralığının daha uzun olacağı sonucuna varılmaktadır. Söz konusu defterlerin dolduğuna ve yeni deftere geçilmesi gerektiğine dair verilen kararda belirleyici olan seviyenin ne olduğu hakkında bir bilgi elde edilememiştir fakat Günay (2013b: 293) “belirli bir yekûna ulaştıktan sonra” ifadesiyle bu seviyeye değinmiştir. Her sene için ayrı ayrı oluşturulan defterlerin sayfa sayıları hakkında bir standart bulunmadığı, defterin sayfa sayısının ilgili sene içerisinde meydana gelen şikâyetlerin miktarı ile ilişkili olduğu ifade edilebilir.

Ayrıca çalışmada Şikâyet Defterleri ile Ahkâm Defterleri arasındaki farklardan bahsedilirken, değinildiği üzere, Ahkâm Defterleri Şikâyet Defterlerine göre dar ve uzun bir yapıya sahiptir.

1716 yılında Mora Ahkâm Defteri, daha sonra 1742 tarihinden itibaren ise diğer eyaletlerde tutulmaya başlayan Ahkâm Defterleri, II. Meşrutiyet dönemine kadar devam etmiştir (Başbakanlık Osmanlı Arşivi Rehberi, 2017: 30-41).

Tablo 2. 1: Ahkâm Defterlerinin Eyaletlere Göre Listesi.

	Defterin Adı	Hicri Takvim	Miladi Tarih	Defter Sayısı
1	Adana	C.1155-S.1295	1742-1878	9
2	Anadolu	C.1155-R.1306	1742-1889	185
3	Bosna	C.1155-1285	1742-1867	9
4	Cezayir ve Rakka	B.1159-Z.1308	1746-1891	25
5	Diyarbakir	C.1155-L.1292	1742-1875	9
6	Erzurum	C.1155-Ş.1296	1742-1879	19
7	Halep	C.1155-B.1266	1742-1850	9
8	İstanbul	C.1155-R.1328	1742-1910	26
9	Karaman	C.1155-Z.1255	1742-1867	39
10	Maraş	C.1155-C.1294	1742-1867	6
11	Mora	L.1128-Z.1255	1716-1840	21
12	Özi ve Silistre	C.1155-ZA.1294	1742-1877	49
13	Rumeli	C.1155-N.1326	1742-1908	85
14	Sivas	C.1155-L.1326	1742-1908	36
15	Şam-ı Şerif	C.1155-N.1326	1742-1908	9
16	Trabzon	C.1155-1329	1742-1911	8

Kaynak: Başbakanlık Osmanlı Arşivi Rehberi, Ankara 2017, s. 30-41.

2.4. Ahkâm Defterlerinin İçeriği

Şer'iyye sicilleri, bulundukları yerde çözüme kavuşturulan davaları içerirken, Ahkâm Defterleri kendi mahallinde çözülememiş ve merkeze intikal etmiş davalardan oluşmaktadır. Bu sebeple şehir tarihi çalışan bir bilim insanı için şer'iyye sicillerinden sonra en önemli kaynaklardan biri de Ahkâm Defterleridir. Bu iki kaynak birlikte ve karşılaştırmalı bir şekilde incelendiğinde, araştırılan konular hakkında farklı detayların elde edilebilmesi mümkün olabilecektir. Örneğin, bir kişi ya da yer hakkında bilinmeyen bir ayrıntı Ahkâm Defterlerinde bulunmuyorken Şer'iyye Sicillerinde bulunuyor olabilir. Ayrıca araştırılan konunun teyit edilebilmesi için de olanak sağlayabilecektir. Şer'iyye sicillerinin varlığının günümüze kadar korunamadığı yerlerde bulunan Ahkâm Defterleri, o bölgenin Şer'iyye Sicilleri niteliği taşıyan arşiv malzemesi olarak kabul edilebilir (Şimşir, 1994: 363). Bu duruma Erzurum'da (bununla birlikte Kars'a ait olanlar da) Şer'iyye Sicillerinin bulunamamasından (Gedikli, 2005: 196) dolayı, Erzurum Ahkâm Defterleri örnek olarak gösterilebilir. Ahkâm Defterleri kayıtlarında, Şer'iyye Sicillerinde olduğu gibi, yer adları, taraflar, şahitler ve hükümde yer alan adres bilgileri hakkında ayrıntıya yer verilmemektedir. Kayıtlar alınırken ufak mahalle vs. adları yerine, defterlerde daha kapsamlı yerleşim

adları kullanılmaktadır (Günay, 2013a: 2018). Bu farkın sebebi Uğur'un (2010: 8-9) belirttiği bilgiler doğrultusunda şu şekilde açıklanabilir; Şer'iyye sicilleri mahkemelerde verilen kararlardan meydana gelmektedir ve kadı ya da naibi tarafından kaleme alınmaktadır. Ahkâm Defterlerinin ise Beylikçi kalemi tarafından yazıldığı bilinmektedir. Şer'iyye sicillerinin, günlük yaşamda karşılaşılan sıradan konulara ait kayıtların yanı sıra cinayet, nikâh, vergi, tayin, lonca davaları ve vakıf konuları ile ilgili kayıtlardan oluştuğu görülmektedir. Bu konuların yanında tereke kayıtlarının tutulduğu, bu kayıtlarda tarafların mal varlığı, alacak verecek durumları ve mirasın dağıtılma şekli ile ilgili bilgilerin bulunduğu da bilinmektedir. Ayrıca sicillerde narh kayıtlarının tutulduğu ve bu kayıtların ürün ile fiyat bilgilerini, nikâh ve boşanma ile ilgili kayıtların ise mehir ve nafaka bilgilerini içerdiği bilinmektedir. Buna göre bahsedilen kayıtların oldukça detay içeren bilgilerden oluştuğu sonucuna varılmaktadır. Farklı konularda detaylı kayıt işlemi gerektiren bu sicillerin konularına göre kayıt edilebilmesi için yazım kurallarını ya da kayıt şablonunu belirten ve danışma kaynağı olarak da nitelendirilebilmesi mümkün olan "ilm-i şürût ve ilm-i sak" adında kitaplar vardır (Uğur, 2010: 9). Şer'iyye sicilleri detaylı kayıtlardan oluşuyorken, Ahkâm Defterlerinin daha genel konulu kayıtlardan oluşması yukarıda bahsedilen durumun en temel sebebi olarak gösterilebilir. Ayrıca defterlerdeki kayıtların farklı kalemlerce ve kişilerce oluşturulma biçimleri de bu durumu etkileyen sebepler arasında gösterilebilir.

Ahkâm kayıtlarında yer alan hükümler üç kısımdan oluşmaktadır:

- İlk kısımda hüküm metnine, bölgede bulunan örfi ve şer'i yetkiliye ya da her ikisine hitap edilerek giriş yapılmaktadır.
- Hükümün ikinci kısmında şikâyetçi kişi ya da kişiler ile şikâyetin nedeni açıklanmaktadır.
- Son kısımda ise söz konusu şikâyetin nasıl çözüleceği hakkında verilen kararlar ifade edilmektedir (Cevger, 2012: 10).

Dîvan-ı Hümâyûn'a insanların şikâyetlerini iletmesinin iki yolu bulunmaktaydı: Birinci yol, şikâyetçilerin bulundukları yerden merkeze giderek burada bizzat şikâyetle bulunmalarıdır. Merkeze gelen insanlar dertlerini Dîvan-ı

Hümâyûn'a ya da bizzat padişaha ilettikten sonra haklarında yazılan kararları beklmeleri gerekmektedir. Merkeze ulaşmanın, dönemin koşulları göz önünde bulundurulduğunda oldukça zorlu ve zaman alıcı bir süreç olduğu tahmin edilebilir. Ayrıca şikâyet başvurularının zamanla artması, insanların şikâyetlerinin görüşülmesinin ve hükümlerin yazılma süresinin uzamasına neden olmuştur. Bahsedilen sebeplerden dolayı merkeze bizzat gidilerek şikâyetlerin iletilmesi yöntemi halk açısından zorlu bir uygulama olmuştur. Belge içerisinde “südde-i sa'âdet/âsîâne-i aliyâneme gelüb/der-i sa'adetime gelüb ...” gibi ifadelerin bulunması kişinin merkeze gelerek şikâyetini arzuhal (dilekçe) ile ilettiğini ifade etmektedir. İnsanların şikâyetlerini iletmek için başvurdukları ikinci yol ise şikâyetçilerin merkeze şahsen gitmeden, merkeze dilekçe yazarak dertlerini anlatmalarıdır. Bu tarz başvuru yapan kişilerin belgelerinde “arzuhal/bâ-arzuhal” ifadesi görülmüştür. Halk, bu iki yoldan genellikle birincisini tercih etmekteydi. (Günay, 2013a: 21). Birinci yolun seçilmesinde dilekçelerin merkeze ulaştırılamaması gibi endişelerin etkili olduğu düşünülebilir.

Ahkâm Defterlerinde yer alan kayıtlar incelendiğinde genellikle yerel mahkemelerin çözüme kavuşturamadığı meselelerin dîvana geldiği ve dîvanda görülen meselelerin ise çözümsüz bırakılmadığı görülmektedir. Dava konularına bakıldığında yazlık, kışlık, bağ, bahçe, mera yeri anlaşmazlığı, arazi meseleleri, arazi sulama, su yolları, tevliyet anlaşmazlıkları, tevcihât ile ilgili meseleler, gayrimüslimlerin haklarının korunması ve güvenliklerinin sağlanması ile ilgili tembihler, faizle ilgili konular, haksız bir şekilde vergi talep edilmesinin önlenmesi, miras ile ilgili konular ve mirasçıların birbirleri arasındaki anlaşmazlıkları, mal ve eşyanın haksız gaspı, eşkıyalık faaliyetleri ve bu faaliyetler ile ilişkisi olanların ahvali, asker sevkıyatı ve yollardan geçecek olan önemli kişi ya da malzemenin korunması ve oluşabilecek ihtiyaçların giderilmesi, halka zulüm edilmesinin önüne geçilmesi, esnaf şikâyetleri, tımar tevcihleri, iltizam verilmesiyle alakalı meseleler, cizye konusu ile alakalı hükümler vb. örfi ve şer'i birçok konu hakkında şikâyetler ve hükümler yer almaktadır (Günay, 2013a: 21-22).

2.5. Erzurum 1 Numaralı Ahkâm Defteri

2.5.1. Defterin Fiziki Yapısı

Dört ayrı Yüksek Lisans tez çalışmasına konu olan 1 Numaralı Erzurum Ahkâm Defteri,² BOA'da {A.DVNS.AHK.ER.d kodu ile kayıtlı bulunmaktadır. Defterin kimlik bilgileri içerisinde, 48 x 17 cm ebadının ve ciltli olduğu ayrıca cildinin de ebrulu olduğu ifade edilmiştir. Defterin baştan bir sayfasının, sondan ise iki sayfasının boş olduğu ve toplamda 293 sayfadan meydana geldiği belirtilmektedir. Kapak üzerinde ise “Erzurum Defteri min evâil-i Cemâziye'l-âhire sene 1155 ilâ evâhir-i Şevval sene 1162” yazısı bulunmaktadır (Bkz. Resim 2.1). Dolayısıyla söz konusu defterin Ağustos 1742 ile Eylül/Ekim 1749 tarihleri arasındaki hüküm kayıtlarını kapsadığı sonucuna ulaşılmaktadır.

Resim 2. 1: Erzurum 1 Numaralı Ahkâm Defteri'nin Kapağı.



Kaynak: BOA, DVNS. AHK. ER. d., No 1.

² 1 Numaralı Erzurum Ahkâm Defterinin 1 ile 65. sayfa aralığı Muhittin Kul, 65 ile 130. sayfa aralığı Hikmet Çiçek, 131 ile 212. sayfa aralığı Feyzanur Susam ve 212 ile 293. sayfa aralığı ise Nazmiye Gül Kantaroğlu tarafından hazırlanan dört ayrı Yüksek Lisans tezinde ele alınmıştır.

293 sayfalık defter içerisinde toplamda 1212 hüküm bulunmaktadır. Ayrıca defterde toplamda 4 adet derkenâr da bulunmaktadır. Defterde bulunan;

- 128 hükmün Muharrem ayında,
- 124 hükmün Safer ayında,
- 126 hükmün Rebiülevvel ayında,
- 135 hükmün Rebiülâhir ayında,
- 130 hükmün Cemaziyelevvel ayında,
- 120 hükmün Cemaziyelâhir ayında,
- 98 hükmün Receb ayında,
- 88 hükmün Şaban ayında,
- 50 hükmün Ramazan ayında,
- 72 hükmün Şevval ayında,
- 79 hükmün Zilkade ayında,
- 62 hükmün ise Zilhicce ayında kaleme alınmıştır (Kul, 2011: 23-46; Çiçek, 2011: 28-55; Susam, 2019: 19-79; Kantaroğlu, 2019: 26-83).

Yukarıdaki bilgilere dayanarak defterde bulunan hükümlerin aylara göre yüzdelik dağılım oranı da belirtilebilir. Bu oranlar Tablo 2.2.'deki gibidir;

Tablo 2. 2: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Aylara Göre Yüzdesel Dağılımı.

Aylar	Toplam	Yüzde
M	128	11%
S	124	10%
Ra	126	10%
R	135	11%
Ca	130	11%
C	120	10%
B	98	8%
Ş	88	7%
N	50	4%

L	72	6%
Za	79	7%
Z	62	5%
Toplam	1212	100%

Kaynak: Kul, 2011: 23-46; Çiçek, 2011: 28-55; Susam, 2019: 19-79; Kantaroğlu, 2019: 26-83.

1 numaralı Erzurum Ahkâm Defterinde bulunan hükümlerin ay ve yıl bakımından dağılımları ve bu değerlerin toplamı Tablo 2.3'te verilmiştir:

Tablo 2. 3: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Aylara ve Yıllara Göre Dağılımı.

	1155	1156	1157	1158	1159	1160	1161	1162	Genel Toplam
M	-	31	20	11	24	16	10	16	128
S	-	14	24	20	10	28	18	10	124
Ra	-	20	9	21	25	19	20	12	126
R	-	30	16	13	19	30	16	11	135
Ca	-	13	17	19	25	26	23	7	130
C	21	10	7	7	19	18	25	13	120
B	19	13	5	8	12	13	8	20	98
Ş	8	14	4	11	10	9	18	14	88
N	13	5	3	3	7	11	6	2	50
L	15	8	15	5	6	7	5	11	72
Za	13	5	7	6	16	18	13	1	79
Z	14	8	10	10	9	7	4	-	62
Genel Toplam	103	171	136	134	182	202	166	118	1212

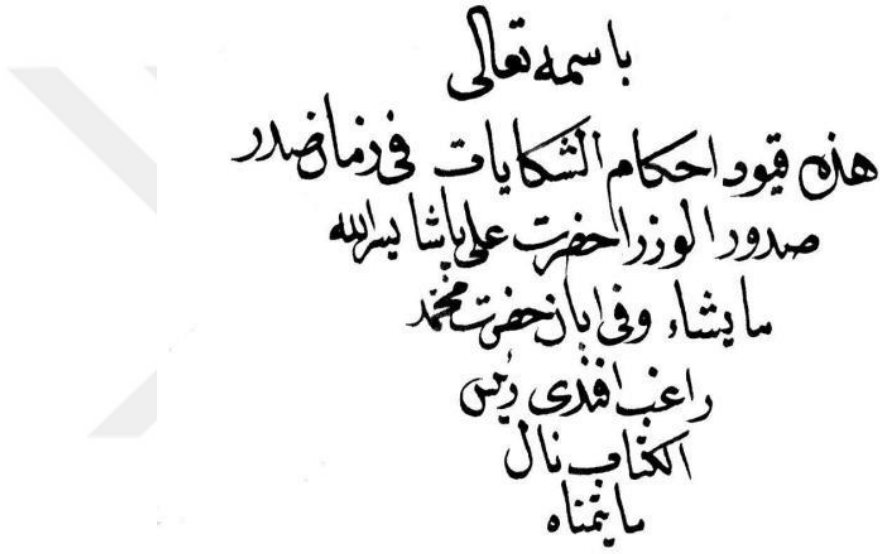
Kaynak: Kul, 2011: 23-46; Çiçek, 2011: 28-55; Susam, 2019: 19-79; Kantaroğlu, 2019: 26-83.

Yukarıdaki tablolar incelendiğinde şu çıkarımlar yapılabilmektedir;

H. 1155/M. 1742 yılında 103 adet hüküm kaleme alınmıştır ve bu hükümler defterde bulunan hükümlerin %8'lik kısmını oluşturmaktadır. H. 1156/M. 1743 yılına ait 171 hüküm bulunmaktadır ve bu hükümler defterin %14'lük kısmını meydana getirmektedir. H. 1157/M.1744 yılında 136 adet hüküm mevcuttur. Bu hükümler defterdeki hükümlerin %11'ine denk gelmektedir. H. 1158/M. 1745 yılında 134

hüküm bulunmakta ve bu da defterin %11'ini oluşturmaktadır. 182 hüküm H. 1159/M. (1746) yılına aittir ve %15'lik bir oranı ifade etmektedir. H. 1160/M. 1747 yılına ait 202 hüküm söz konusudur ve defterin %17'sini oluşturmaktadır. H. 1161/M. 1748 yılında 166 hüküm kayda alınmıştır. Bu hükümler defterin %14'ünü temsil etmektedir. H. 1162/M. 1749 yılında ise 118 adet hüküm yazılmıştır ve defterin %10'luk kısmını meydana getirmektedir.

Resim 2. 2: 1 Numaralı Erzurum Ahkâm Defterinin Girişi.



Kaynak: BOA, DVNS. AHK. ER. d., No 1: 1.

1 numaralı Erzurum Ahkâm Defterine hükümlerin yazılmasına başlamadan önce Resim 2.2.'deki, “Bi-ismihû te’âlâ hâza kuyûd-ı ahkâmî’ş-şikâyât fî-zamân-ı sadr-ı sudûri’l-vüzerâ hazret-i Ali Paşa yesserallahu mâ-yeşâü ve fî âbân-ı hazret-i Mehmed Ragîb Efendi reîsü’l-küttâb nâle-mâ yetemennahu” ibaresi yer almaktadır ve bu ibarenin ardından hükümler yazılmaya başlanmıştır. Defterin her bir sayfasının sağ ve sol üst köşelerine sonradan Latin rakamları eklenerek sayfaların numaralandırılması yapılmıştır.

Defterlerdeki hükümlerin bazen tek kişiye bazen de birden fazla yöneticiye yazıldığı görülmektedir. Verilen hükümlerde hiyerarşik sisteme önem verilmektedir. Hükümün birden fazla yöneticiye yazıldığı durumlarda ilgili yerin en üst makamından başlamak kaydıyla hüküm, hiyerarşik düzene sadık kalınarak kaleme alınmıştır.

Hükümlerin birden fazla yöneticiye yazılması durumu hakkında şu çıkarımlarda bulunulabilir;

- Eğer hükümlerin konuları başka yöneticilerle alakalı ise söz konusu hüküm ilgili yöneticilere gönderilebilir.
- Bir şikâyet sonucunda verilen hüküm, benzer durumların yaşanmaması için yerel önlemlerin alınması, yaşandığı takdirde ise örnek teşkil etmesi sebebiyle sonucunun belirlendiğine vurgu yapılması bakımından ilgililerin hem bilgilendirilmesi hem de uyarılması için diğer yöneticilere de gönderilebilir.
- Taşrada bulunan halkın yerel yöneticilerle sorun yaşaması muhtemeldir. Söz konusu durumlarda halkın yöneticileri şikâyet edebilmesinin önü, bölümün başında da bahsedildiği üzere Dîvan-ı Hümayun'a ya da padişaha doğrudan ulaşılabilmesi açısından, açıktır. Sorunun çözümü hakkında verilen hükmün tarafsızca uygulanabilmesi için konunun takibinin yapılabilmesi ve ayrıca halkın mağdur olmaması sebebiyle farklı yöneticilere iletilebilir.

Hükümler, tarih sıralamasına önem verilerek kayda alınmıştır. Defterin alt kısımlarında boşluk bırakıldığı ve sonradan ekleme yapıldığı görülmektedir. Sonradan ekleme yapılan hükümlerin tarih bütünlüğünü bozduğu görülmüştür. Defterlerde sehven mükerrer yazılan hükümler ile verilen emirlerin iptal edilmesi gibi durumların da var olduğu bilinmektedir (Çiçek, 2011: 19).

2.5.2. Defterin Diplomatik ve Teknik Özellikleri

Defterdeki hükümlerde hüküm, emir ve ferman gibi dîvanda alınan kararların suretleri bulunmaktadır. Bu kararların tamamında hitap makamı yer almaktadır. İlk olarak hükmün gönderildiği kişiye göre kişinin elkabı sonrasında elkabı yazılan kişinin adı zikredilerek dua cümlesi yazılmaktadır. Fakat bu kararların çoğunda elkab ve dua kısmının olmadığı bilinmektedir. Elkab ve dua kısmından sonra standart ifadeler yer almaktadır. Defterlerdeki hükümler üç bölüme ayrılmaktadır (Çiçek, 2011: 19-20):

- Birinci bölümde, yörenin şer'i ya da örfi amirine, idarecisine hitap kısmı bulunmaktadır.
- İkinci bölümde, hükmün yazılmasına sebep olan olaylar ve gelişmeler hakkında bilgi verilmektedir. Burada davayla alakalı sunulan arzuhaller, mektuplar ve sözlü ifadeler de yer almaktadır.
- Üçüncü ve son bölümde ise ikinci bölümde bahsedilen olaylar ve gelişmeler ile ilgili bürokrasinin aldığı kararlar anlatılmaktadır.

2.5.3. Defterin Dil Özellikleri

Defterlerin dili oldukça sade ve anlaşılır bir şekilde Türkçe olarak yazılmıştır. Hükümlerde Arapça ve Farsça kelimeler ve kalıplar da yer almaktadır. Metinlerin nakil ve iblâğ kısmında bulunan müracaatlar anlaşılır ve kısa yazılmıştır (Çiçek, 2011: 20). İmla ve kelimelerin yazılış biçimlerinde bazı tutarsızlıklar bulunabilmektedir. Bu duruma aynı belgenin içerisinde ya da tüm belgede rastlamak mümkün olabilmektedir (Kul, 2011: 12-13). Defter “divanî kırma” yazı biçimi ile kaleme alınmıştır (Kantaroglu, 2019: 16). Divanî kırma yazı biçiminde yazının hızlı bir şekilde yazılması sonucu noktalama işaretlerinin tam anlamıyla kullanılmaması, ayrıca söz konusu defterlerin Beylikçi kaleminde bulunan farklı görevlilerce yazılması yukarıda bahsedilen tutarsızlıkların nedenleri olarak gösterilebilir.

2.5.4. Defterde Ele Alınan Konular

Erzurum 1 Numaralı Ahkâm Defterlerinin 1-293 sayfalarında toplam 1216 adet hüküm yer almaktadır. Söz konusu hükümler içerisinde 3 adet iptal edilmiş hüküm ve 4 adet derkenâr bulunmaktadır. Bu hükümleri içerikleri bakımından belirli kategorilere ayırmak mümkündür. Hükümler konu bazında incelendiğinde dağılımları genel olarak şu şekilde ifade edilebilir;

- 268 adet asayiş,
- 432 adet sosyal, ekonomik ve idarî konular,

- 509 adet ise arazi meselesi ve vergi ile ilgili hüküm bulunmaktadır (Kul, 2011: 23-46; Çiçek, 2011: 28-55; Susam, 2019: 19-79; Kantaroğlu, 2019: 26-83).

Hükümlerin bir kısmının birden çok konuyu ilgilendirmesi söz konusudur. Bu duruma örnek olarak şu hüküm verilebilir; Kars-ı Erzurum'da Şeyh Ebü'l-Hasan El-Harkani zaviyesi vakfına ait olan mala bazı kişiler tarafından müdahale ve te'addi edilmiştir. Bu durumun çözülmesi için Kars beylerbeyine ve Kars-ı Erzurum kadısına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 238/4; aktaran Kantaroğlu, 2019: 169-170). Bu hüküm doğrultusunda ismi belirtilmiş olan vakıf malının (arazi, bina, para, mahsul gibi) zorla alınması/gasp edilmesi durumunun söz konusu olduğu görülmektedir. Bahsedilen hükmün içerisinde geçen zorla alma/gasp etme durumu, ilgili hükmün asayiş konu başlığı altında değerlendirilebilmesini olağan kılmaktadır. Fakat zorla alma/gasp etme eyleminin vakıf malı üzerinde gerçekleştirilmesi esasında hükmün esas konusunun vakıfla alakalı olduğu ve zorla alma/gasp etme durumunun tali konu olduğu sonucunu meydana getirmektedir. Burada hükmün esas konusu temel alınmış, tali konu hüküm bazında dikkate alınmamıştır. Ayrıca hükümlerin konu başlıkları belirlenirken her bir hükmün konusunu ifade eden spesifik başlıkların yanı sıra daha kapsayıcı genel konu başlıkları belirlenmiştir.

Tablo 2. 4: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Genel Konulara Göre Dağılımı.

Hükümlerin Konusu	Adet
Asayiş	268
Sosyal, Ekonomik ve İdarî Konular	432
Arazi Meselesi ve Vergi	509
Toplam	1209
(3 adet iptal edilen hüküm ve 4 adet derkenâr bulunmaktadır)	

Kaynak: Kul, 2011: 23-46; Çiçek, 2011: 28-55; Susam, 2019: 19-79; Kantaroğlu, 2019: 26-83.

2.5.4.1. Asayiş

Erzurum 1 Numaralı Ahkâm Defteri içerisinde bulunan hükümler incelendiğinde asayiş ile ilgili 268 hüküm bulunmaktadır. Asayiş konu başlığı toplamda 7 alt konu başlığından oluşmaktadır. Bu alt konu başlıkları; cana ve mala müdahale ve tecavüz, fuhuş, iftira, cinayet, güvenliğin sağlanması, diyet talebi ve zimmete mal geçirmedir. Eşkîyalık, mal gaspı, para gaspı, tecavüz, darp ve hırsızlık ile ilgili hükümler daha kapsayıcı bir başlık olan “cana ve mala müdahale ve tecavüz” alt konu başlığı altında, can güvenliğinin sağlanması, resmî evrak ve değerli malların korunması ve şahıs güvenliğinin sağlanması ile ilgili hükümler ise daha kapsayıcı bir başlık olan “güvenliğin sağlanması” alt konu başlığında bir araya getirilmiştir. Yaşanılan asayiş problemlerinin yanı sıra deftere konu olan ve halkın günlük yaşantısında karşılaşmış olduğu problemleri, problemin muhataplarının devlet adamı, yönetici ve din görevlisi olması fark etmeksizin, doğrudan padişaha ya da Dîvân-ı Hümâyûn’a iletebilmeleri hukuk devleti anlayışının hâkim olduğunu göstermektedir.

Asayiş ile ilgili hükümlerin içerisinde en çok cana ve mala müdahale ve tecavüz konulu hükümler dikkat çekmektedir. Bu hükümlere örnek olarak aşağıdaki hüküm gösterilebilir; İbrahim adlı bir kişi kardeşi Abdullah ile beraber ticaret yapmak için yolculuk ederlerken eşkıyalar tarafından yolları kesilmiştir. İbrahim ve Abdullah öldürülmüştür. Eşyaları, malları ve paraları gasp edilmiştir. Konunun çözülmesi için Erzurum ve Trabzon valilerine, Pazarsuyu, Akköy, Karahisar-ı Şarki, Giresun vb. kazaların kadılarına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 13/2; aktaran Kul, 2011: 85). Bir başka örnekte ise; Eşkîya Şehzuvarzâde İsmail Bey haksız bir şekilde Seyyid Ömer ve Seyyid Süleyman adlı kişilerin mahsullerini ellerinden almıştır. Söz konusu mahsullerin hak sahiplerine iade edilmesi için Erzurum valisine ve Kuruçay kadısına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 143/2; aktaran Susam, 2019: 118).

2.5.4.2. Sosyal, Ekonomik ve İdarî Konular

1 numaralı Erzurum Ahkâm Defterinde yer alan 432 hüküm sosyal, ekonomik ve idarî konular başlığı altında toplanmıştır. Sosyal, ekonomik ve idarî konular başlığı

altında bir araya getirilen hükümler 17 alt konu başlığı ile ifade edilebilmektedir. Bu alt konu başlıkları; vakıf, borç meselesi, anlaşmazlık, gündelik yardımı, yol kapama, göç, faiz, haksız para talebi, haksız mal talebi, ihtida, mehir, miras, nikâh akdi, nafaka, berat tevcihi ve memuriyet ve tedariktir. Vakıf tevcihi, vakıf yönetimi tevcihi ve vakfa müdahale ile ilgili hükümler “vakıf” konu başlığı altında, mülk anlaşmazlığı, satış anlaşmazlığı ve su anlaşmazlığı ile ilgili hükümler “anlaşmazlık” konu başlığı altında, haksız hisse talebi, miras ihtilafı, tereke taksimi ve mirasa müdahale ile ilgili hükümler “miras” konu başlığı altında, sâ’ilik beratı ve müderrislik beratı ile ilgili hükümler “berat tevcihi” konu başlığı altında ve son olarak sorumluluklarını yapmama, göreve müdahale, vazifeyi kötüye kullanma, soruşturma, vazife tevcihi ve tayin talebi ile ilgili hükümler ise “memuriyet” konu başlığı altında bir araya getirilmiştir. Sosyal, ekonomik ve idarî konular ile ilgili verilmiş hükümlere örnek olarak aşağıdaki hükümler gösterilebilir; İskefsir kazasına bağlı Ermeni köyü sakinlerinden Salih ve Gevher adlı kişilerin ortak olarak sahip oldukları tarlaları bulunmaktadır. Ortaklardan Gevher adlı kadın vefat etmiştir. Gevher hatunun oğlu olmadığından tarlaların tapu hakkı diğer ortak olan Salih’e geçmektedir. Hatunun iki damadı söz konusu tarlalardan, ellerinde fetva olduğu iddiasını ileri sürerek kanun dışı bir şekilde hak talep etmişlerdir. Kanun dışı bu talebin reddedildiği belirtilerek kanunların gerektirdiği şekilde konunun çözülmesi için Erzurum valisi ve İskefsir kadısına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 197/5; aktaran Susam, 2019: 290-291).

Karahisar-ı Şarkî sancağına bağlı Çavuşlar köyü sakinlerinden Meryem adlı kişi vefat etmiştir ve malları miras olarak kardeşi Mustafa’ya kalmıştır. Fakat Hüseyin, İbrahim Ağa, Ali, İbrahim Beşe ve Kel Hasan adlı kişiler söz konusu mallara müdahale etmiştir. Meselenin çözümü için Karahisar-ı Şarkî kadısı ve voyvodasına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 160/1; aktaran Susam, 2019: 171).

2.5.4.3. Arazi Meselesi ve Vergi

Osmanlı Devleti ekonomisinin önemli bir kısmını halktan alınan vergiler oluşturmaktadır. Defterde yer alan hükümler incelendiğinde arazi meselesi ve vergi ile ilgili hükümlerin fazla olduğu gözlemlenmektedir. Arazi meselesi ve vergi konu başlığı altında toplamda 509 adet hüküm bulunmaktadır. Bu hükümler; arazi ve sınır

meselesi, zeamet, malikâne, timar, öşr, haksız vergi talebi, vergi vermeme, cebelü bedeliyesi ve ulûfeye müdahale ile ilgili alt konu başlıklarıyla ifade edilebilmektedirler. Mülke müdahale, sınır ihtilafı, meraya müdahale, yurtluğa ve ocaklığa müdahale, koruya müdahale ve yaylağa müdahale ile ilgili hükümler “arazi ve sınır meselesi” konu başlığı altında, zeamete müdahale, zeamet tahsili ve zeamet mutasarrıflığı tevcihi ile ilgili hükümler “zeamet” konu başlığı altında, malikâneye müdahale ve malikâne tahsili ile ilgili hükümler “malikâne” konu başlığı altında, timar tevcihi, timara müdahale, timar mutasarrıflığı, timar tahsili ve timar anlaşmazlığı ile ilgili hükümler “timar” konu başlığı altında ve son olarak öşr tahsili ve öşr’e müdahale ile ilgili hükümler ise “öşr” konu başlığı altında bir araya getirilmiştir. Defter içerisinde bulunan çok sayıdaki arazi meselesi ve vergi ile ilgili hükümlere aşağıdaki hükümler örnek olarak verilebilir;

Gercanis (Refâhiye) kazasına bağlı Manek Deresi nahiyesinde vergi alınmasını gerektirecek mülkü bulunmayan Ahmed isimli kişiden haksız bir şekilde vergi talep edilmiştir. Konunun çözülmesi için Gercanis kadısına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 157/3; aktaran Susam, 2019: 162).

Vergi konusu ile alakalı hükümlere verilebilecek bir diğer örnek hükümde ise yeniçeri gibi çeşitli askerî birliklerden olduklarını belirten bazı kişilerin vergilerini ödemek istemedikleri belirtilmiştir. Söz konunun durumun çözülmesi için Erzurum valisine, Kelkid, Bayburd ve Koğans (Kale) kadılarına hüküm yazılmıştır (BOA, DVNS. AHK. ER. d., No 1, 182/4; aktaran Susam, 2019: 241-242).

Aşağıda verilen tabloda 3 ana konu başlığını meydana getiren alt konu başlıkları ve hüküm sayıları (parantez içerisinde) görülmektedir. Tablo incelendiğinde 218 hüküm ile cana ve mala müdahale ve tecavüz alt konu başlığı içerisinde bulunan hükümlerin sayısı dikkat çekmektedir. Bir diğer dikkat çekici alt konu başlığı ise 129 adet hükümle haksız vergi talebidir. Diğer konu başlıkları ve hüküm sayıları Tablo 2.5.’te detaylı olarak verilmiştir.

Tablo 2. 5: 1 Numaralı Erzurum Ahkâm Defteri'ndeki Hükümlerin Konulara Göre Detaylı Dağılımı.

1. ASAYİŞ	2. SOSYAL, EKONOMİK VE İDARİ KONULAR	3. ARAZİ MESELESİ VE VERGİ
<u>1.1. Cana ve Mala Müdahale ve Tecavüz (218)</u> <i>1.1.1. Eşkıyalık (142)</i> <i>1.1.2. Mal Gaspı (35)</i> <i>1.1.3. Para Gaspı (29)</i> <i>1.1.4. Tecavüz (4)</i> <i>1.1.5. Darp (2)</i> <i>1.1.6. Hırsızlık (6)</i> <u>1.2. Fuhuş (1)</u> <u>1.3. İftira (11)</u> <u>1.4. Cinayet (24)</u> <u>1.5. Güvenliğin Sağlanması (5)</u> <i>1.5.1. Can Güvenliğinin Sağlanması (2)</i> <i>1.5.2. Resmî Evrak ve Değerli Malların Korunması (1)</i> <i>1.5.3. Şahıs Güvenliğinin Sağlanması (2)</i> <u>1.6. Diyet Talebi (8)</u> <u>1.7. Zimmete Mal Geçirme (1)</u>	<u>2.1. Vakıf (76)</u> <i>2.1.1. Vakıf Tevcihi (2)</i> <i>2.1.2. Vakıf Yönetimi Tevcihi (3)</i> <i>2.1.3. Vakfa Müdahale (71)</i> <u>2.2. Borç Meselesi (162)</u> <u>2.3. Anlaşmazlık (34)</u> <i>2.3.1. Mülk Anlaşmazlığı (4)</i> <i>2.3.2. Satış Anlaşmazlığı (10)</i> <i>2.3.3. Su Anlaşmazlığı (20)</i> <u>2.4. Gündelik Yardımı (1)</u> <u>2.5. Göç (14)</u> <u>2.6. Faiz (4)</u> <u>2.7. Haksız Para Talebi (6)</u> <u>2.8. Haksız Mal Talebi (2)</u> <u>2.9. İhtida (2)</u> <u>2.10. Mehir (1)</u> <u>2.11. Miras (74)</u> <i>2.11.1. Haksız Hisse Talebi (4)</i> <i>2.11.2. Miras İhtilafları (13)</i> <i>2.11.3. Tereke Taksimi (16)</i> <i>2.11.4. Mirasa Müdahale (41)</i> <u>2.12. Nikâh Akdi (9)</u> <u>2.13. Nafaka (1)</u> <u>2.14. Berat Tevcihi (2)</u> <i>2.14.1. Sâ'ilik Beratı (1)</i> <i>2.14.2. Müderrislik Beratı (1)</i>	<u>3.1. Arazi ve Sınır Meselesi (112)</u> <i>3.1.1. Mülke Müdahale (94)</i> <i>3.1.2. Sınır İhtilafları (1)</i> <i>3.1.3. Meraya Müdahale (8)</i> <i>3.1.4. Yurtluğa ve Ocaklığa Müdahale (1)</i> <i>3.1.5. Koruya Müdahale (2)</i> <i>3.1.6. Yaylağa Müdahale (6)</i> <u>3.2. Zeamet (73)</u> <i>3.2.1. Zeamete Müdahale (66)</i> <i>3.2.2. Zeamet Tahsili (5)</i> <i>3.2.3. Zeamet Mutasarrıflığı Tevcihi (2)</i> <u>3.3. Malikâne (18)</u> <i>3.3.1. Malikâneye Müdahale (16)</i> <i>3.3.2. Malikâne Tahsili (2)</i> <u>3.4. Timar (126)</u> <i>3.4.1. Timar Tevcihi (19)</i> <i>3.4.2. Timara Müdahale (100)</i> <i>3.4.3. Timar Mutasarrıflığı (1)</i> <i>3.4.4. Timar Tahsili (4)</i> <i>3.4.5. Timar Anlaşmazlığı (2)</i> <u>3.5. Öşr (11)</u> <i>3.5.1. Öşr Tahsili (4)</i> <i>3.5.2. Öşr'e Müdahale (7)</i> <u>3.6. Haksız Vergi Talebi (129)</u> <u>3.7. Vergi Vermeme (32)</u> <u>3.8. Cebelü Bedeliyesi (7)</u>

<u>2.15. Memuriyet (42)</u> <i>2.15.1. Sorumluluklarını Yapmama (2)</i> <i>2.15.2. Göreve Müdahale (16)</i> <i>2.15.3. Vazifeyi Kötüye Kullanma (18)</i> <i>2.15.4. Soruşturma (2)</i> <i>2.15.5. Vazife Tevcihi (3)</i> <i>2.15.6. Tayin Talebi (1)</i> <u>2.16. Yol Kapama (1)</u> <u>2.17. Tedarik (1)</u>	<u>3.9. Ulûfeye Müdahale (1)</u>
---	--------------------------------------

ÜÇÜNCÜ BÖLÜM

METİN MADENCİLİĞİ UYGULAMASI

3.1. Uygulamada Kullanılan Programlar

1 numaralı Erzurum Ahkâm Defteri üzerinde metin madenciliği uygulamaları yapılırken Microsoft Excel ve KNIME (Konstanz Information Miner) programlarından faydalanılmıştır. Microsoft Excel programı vasıtasıyla, dağınık ve farklı formatta olan veriler tablolar halinde bir araya getirilmiştir. Excel programı çalışmanın veri tabanını oluşturmuştur. KNIME programı ile daha öncesinde Excel programı sayesinde oluşturulmuş olan tablolar içerisindeki veriler belirli ön işleme aşamalarından geçirilerek sınıflandırma işlemine hazır hale getirilmiştir ve ardından uygun algoritmalar ile sınıflandırma işlemleri gerçekleştirilmiştir.

3.1.1 Microsoft Excel

Bilgisayar teknolojilerinin hızla gelişmesi ile kişilerin iş ve işlemlerini düzenleyebilecekleri ve kontrol altında tutabilecekleri ayrıca bu iş ve işlemlerini kolaylıkla yönetebilecekleri program ve yazılımlar geliştirilerek süreçlerin daha hızlı ve pratik bir hale getirilmesi için çözümler üretilmiştir. Microsoft Office yazılım paketi içerisinde yüklü olarak bulunan programlar bahsedilen duruma örnek olarak gösterilebilir. Microsoft Office yazılım paketi içerisinde bulunan Word, PowerPoint, Access ve Excel sıklıkla kullanılan ve yaygın olarak bilinen programlardır. Microsoft Windows ve Apple Macintosh gibi farklı işletim sistemlerinde çalışabilen Microsoft Excel programı bilgisayarlara verilerin girilmesine imkân sağlayan, verilerin farklı koşullara göre listelenmesini, verilerin birleştirilmesini ve ayrılmasını mümkün kılan ayrıca veriler içerisinde arama yapılabilmesine olanak sağlayan işlevsel bir programdır. Bir ya da daha fazla çalışma sayfası içerebilen, kolay ulaşılabilirliği ve birçok hesaplamanın yapılabilmesine olanak sağlayan, formül yazma ve grafik çizme imkânı tanıyan etkili bir programdır (Öztek, 2001; aktaran Kutluca, 2013: 43-44).

Basit matematiksel hesaplamaların yapılmasından ileri düzeyde kâr-zarar analizlerinin yapılmasına kadar geniş bir kullanım alanına sahiptir (Bağcı, 2022: 25).

3.1.2. KNIME

Konstanz Information Miner olarak adlandırılan KNIME programı (Berthold, Cebron, Dill, Gabriel, Kötter, Meinl, Ohl, Sieb, Thiel ve Wiswedel, 2008: 319), 2004 yılında Konstanz Üniversitesinde (Universität Konstanz) bir grup yazılımcı tarafından geliştirilmeye başlanmıştır ve 2006 yılında ilk sürümü kullanıma sunulmuştur (Akyüz ve Gülten, 2022: 136). Veri boyutunda kısıtlama yapmayan, bu sebeple büyük miktarlarda ve farklı veri türleri ile çalışılmasına imkân sağlayan KNIME, Java programlama dili ile yazılmış bir programdır.¹ Açık kaynak kodlu olan program GPL (Genel Kamu Lisansı, General Public License) lisansı altında kullanılabilmektedir (Thiel, 2016: 63). Yazılımların özgür olması gerektiğini belirten Richard M. Stallman tarafından 1984 yılında (Güneş, 2007: 153) GNU işletim sistemi (Stallman, 2002: 20) oluşturulmaya başlanmıştır. “GNU (GNU Unix Değildir) ismi, “GNU’s Not Unix” ifadesinin özyinelemeli bir kısaltmasıdır (Stallman, 2002: 19). GNU projesi kullanıcıların hizmetine bir yazılımı çalıştırma, kopyalama, dağıtma, inceleme, değiştirme ve geliştirme özgürlükleri (Güneş, 2007: 153) ile çevrimiçi kaynakları sunmaktadır (Akadal, Özdemir ve Ayvaz Reis, 2013: 909). Özgür yazılım düşüncesinden meydana geldiği bilinen projeyi tanımlayan kavramın, Unix işletim sistemine benzer şekilde kapalı kaynak kodlu olmadığına vurgu yapılması için seçildiği sonucunu çıkarmak mümkündür. GNU projesinin üç farklı lisansı bulunmaktadır: 1) Genel Kamu Lisansı (GPL - General Public License), 2) Kısıtlı Genel Kamu Lisansı (LGPL - Lesser General Public License), 3) Özgür Belgeleme Lisansı (FDL - Free Document License) (Demirtel ve Sağiroğlu, 2011: 57). Güneş (2007: 154) çalışmasında, bu lisanslar arasında diğerlerine oranla daha sık kullanılan lisansın Genel Kamu Lisansı (GPL - General Public License) olduğunu belirtmektedir. Genel Kamu Lisansı’nın (GPL - General Public License) 1989 ve 1991 yıllarında yayınlanan sürümleri bulunmaktadır ve 2007 yılında yayınlanan güncel sürümü GLP

¹ <https://www.knime.com>

v3 (Gülüm ve Köksal, 2021: 195) ile KNIME programı da lisanslanmaktadır². Söz konusu lisans temel olarak var olan kod ya da ürünün paylaşılabilmesini ve üzerinde değişiklik yapılabilmesini mümkün kılmaktadır. İlgili kod ya da üründen yeni bir ürün elde edilebilir ve ticari amaçlarla kullanılarak gelir elde edilebilir fakat bu durum bir yükümlülük doğurmayacaktır. Lisans sadece elde edilen yeni kod ya da ürünün de bu lisans ile birlikte kullanılmasını gerektirir. Böylece kullanıma açılan kod ya da ürünler diğer kullanıcıların istek ve ihtiyaçlarına göre yeniden düzenlenebilir ve bu durum kümülatif bir şekilde ilerleme göstermektedir. Yeni kod ve ürünler büyük miktardaki kullanıcı toplulukları ile paylaşarak var olan sorunların hızlı bir şekilde fark edilmesine ve bu hataların kolayca çözümlenmesine olanak sağlamaktadır. KNIME programı çok sayıda kullanıcı topluluğuna sahiptir. Lisansın sağladığı avantajla paralel olarak KNIME kullanıcı topluluğu birbirine yardım edebilir ve hazır olan iş akışlarını birbiri ile paylaşarak hem programın gelişmesine hem de yeni kod ve ürünlerin elde edilmesine katkı sağlamaktadır.

KNIME programı içerisinde yapılacak tüm işlemler için bir çalışma ekranı (Workspace) bulunmaktadır. Programda veri girişlerini ya da modellerini işleyen ve çıktılarından ise sonuçlar üreten farklı Düğümler (Node) bulunmaktadır (Thiel, 2016: 63). Bu düğümler, yapılacak iş ve işlemler doğrultusunda birbiri ile bağlanır ve bir akış diyagramı oluşturulur. KNIME sürükle bırak yöntemi ile çalışmaktadır. Makine öğrenmesi, veri madenciliği ve metin madenciliği yapmaya imkân sağlayan tüm düğümler iş akışına (Workflow) Düğüm Deposu (Node Repository) bölümünden kolaylıkla eklenebilir. KNIME, farklı dosya uzantılarından (*.txt., *.csv, *.arff) veri okuyabilmektedir (Baykal ve Coşkun, 2018: 109). Söz konusu programın desteklediği diğer dosya türleri ise Excel, CSV, PDF, mdf, mol2, sdf, JSON, XML, RDF ve elektronik posta biçimleri ile görüntü, ses ve ağlardır³. Node yapıları sayesinde diğer veri analizi ve makine öğrenimi algoritmaları ile kolaylıkla entegre olabilmektedir ve Windows, Linux ile macOS gibi işletim sistemlerinde ücretsiz bir şekilde kullanılabilir (Toplu, 2019: 33). Kolay anlaşılabilir ve tek bir ara yüze sahip olması, hazır düğümleri sayesinde kod yazımını en aza indirmesi ve çalışmalarda

² <https://www.knime.com/downloads/full-license>

³ <https://www.knime.com/file-access-file-transformation>

kullanılmak istenilen düğümler hakkında hem programın içerisinde hem de forum sayfasında⁴ bilgi verilmesi, bakımından oldukça kullanışlı bir programdır. Ayrıca söz konusu forum sayfasında diğer KNIME kullanıcıları tarafından geliştirilen iş akışlarının ve düğümlerin paylaşıyor olması, oluşturulan modellerin erişilebilir olmasına ve geliştirilmesine olanak sağlamaktadır. Bunların yanı sıra forum sayfasında tanıtımı yapılan düğümler sürükleyip bırak yöntemiyle çalışmalara doğrudan eklenebilmektedir. Tekerek (2011: 167) KNIME programının farklı dosya ve veri tabanlarından veri alışverişi yapabildiğine, verilerin ön işleme aşamalarından geçirilerek anlamlı hale getirilebildiğine, veri ve sonuçlarını çok sayıda görüntüleme aracı vasıtasıyla görselleştirerek yapılacak olan çıkarımları mümkün kıldığına ve karar ağaçları, destek vektör makineleri gibi yöntemlerle madencilik işlemlerinin yapılmasına olanak sağladığına değinmiştir. Bilgin (2009: 810) ise KNIME programının geliştirilebilir olduğuna değinmiştir. Kullanıcıların program içerisinde kodlarını yazabilecekleri yazılım geliştirme kiti (software development kit) sayesinde kendi modüllerini yazabildiklerini ifade etmiştir. Diğer veri tabanı sunucuları, veri madenciliği ve istatistik uygulamaları arasında verilerin aktarılmasını mümkün kılan PMML (Predictive Model Markup Language-Öngörülü Model İşaretleme Dili) adındaki XML tabanlı dilde verilerin içe alımını sağladığına ve ayrıca Data Write bileşenleri ile verilerin dışa aktarımını da sağladığına değinmiştir. Yukarıda bahsedilen KNIME programının özellikleri ve sağlamış olduğu imkânlar, bu çalışmada bahsi geçen programın kullanılmasının belirleyici nedenlerini oluşturmaktadır.

3.2. Veri Toplama Süreci

1 numaralı Erzurum Ahkâm Defteri toplamda 293 sayfadan oluşmaktadır. Defterde bulunan hükümler Arap harflerinin yanı sıra kısmen Farsça ve Türkçe seslere karşılık gelen harfler (p, ç, j, g ve nazal n) ile yazılmıştır. Söz konusu hükümlerin üzerinde yapılacak analiz işlemleri için metinlerin Latin harflere dönüştürülmesi gerekmektedir. 1 numaralı Erzurum Ahkâm Defteri, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Tarih Anabilim Dalı Yüksek Lisans öğrencilerinden Muhittin Kul

⁴<https://hub.knime.com/>

ve Hikmet Çiçek tarafından ayrıca Amasya Üniversitesi Sosyal Bilimler Enstitüsü Tarih Anabilim Dalı Yüksek Lisans öğrencilerinden Feyzanur Susam ve Nazmiye Gül Kantaroğlu tarafından tez çalışması olarak ele alınmıştır. Defter bu dört tez çalışmasında belirli sayfa aralıklarına göre ayrılarak transkript edilmiştir. Muhittin Kul defterin ilk 65 sayfasındaki hükümleri, Hikmet Çiçek 65 ile 130. sayfa aralığındaki hükümleri, Feyzanur Susam 131 ile 212. sayfa aralığındaki hükümleri ve son olarak Nazmiye Gül Kantaroğlu ise 212 ile 293. sayfa aralığındaki hükümleri tez çalışmasında incelemiştir. Bu tezlerde hükümler Latin harflerine çevrilmiştir ayrıca hükümler özetlenerek konuları belirlenmiştir.

Yukarıda bahsedilen tezlerde transkript edilmiş olan hükümler bu çalışmanın veri kaynağını oluşturmaktadır. Bu tezlerdeki hükümlerin veri kaynağı olarak seçilmesinde; 1 numaralı Erzurum Ahkâm Defterinin tamamı transkript edildiği için hükümler arasındaki ilişkilerin ve bütünlüğün korunduğu ve bu sebeple konulara göre sınıflandırma yapıldığında konular, olaylar, kişiler ve mekân hakkında yapılacak olan çıkarımların daha anlamlı sonuçlar vereceği düşüncesi etkili olmuştur. Ayrıca defterde yer alan hükümlerin transkript edilerek konularının ve özetlerinin çıkarılması süreci oldukça zaman alan bir çalışma gerektirmektedir. Bu durumun tez çalışmasının süresini olumsuz etkileyeceği öngörülmüştür. Bu yüzden, transkripsiyonu tamamen yapılmış bir defter tez konusu olarak tercih edilmiştir.

Dört farklı tez çalışmasından elde edilen ve çevirileri yapılmış olan hükümler, hükümlerin tarihleri, sayfa numaraları, konu başlıkları ve hükümlerin özetleri Excel'e aktarılmıştır. Toplamda 1216 adet hükümden oluşan bir veri kümesi meydana gelmiştir. Bu veriler içerisinde 4 adet derkenâr ve 3 adet iptal edilmiş hüküm bulunmaktadır. Bu veriler çalışmanın sonuçlarına katkı sağlamayacağı gerekçesiyle kullanılmamıştır. Çalışmada yapılan iş ve işlemler 1209 adet hüküm veri üzerinde uygulanmıştır.

3.3. Verilerin Hazırlanması ve Veriyi Anlamlandırma

Ahkâm Defterleri içerikleri bakımından yazıldığı dönem ve toplum hakkında sosyolojik, psikolojik, ekonomik, siyasal, hukuksal ve kültürel konulara kaynaklık

edebilecek bilgiler içerebilmektedir. Fakat söz konusu bilginin bir kısmı hükümler içerisinde gömülü ya da gizli bir şekilde bulunmaktadır. Ayrıca hükümlerde verilen bilginin açık olması, her şeyi ile anlaşılabilmesi, hatta farklı zaman, mekân, kişi ve olaylar arasındaki muhtemel ilişkilerin kolaylıkla tespit edilebileceği anlamına gelmemektedir. Hükümler üzerinde çeşitli araç ve yöntemlerle bir dizi işlem yapılarak bu bilgilere ulaşılabilmesi ve muhtemel ilişkilerin/bağlantıların tespit edilmesi mümkündür. Yapılabilecek işlemlerden bir tanesi hüküm verilerinin sınıflandırılmasıdır. Hüküm verilerinin sınıflandırılması işlemi sonucunda sayısal birçok sonuç elde edilebilmektedir. Bu sonuçlar çok sayıda istatistiksel farklı şekil ve tablonun elde edilmesine olanak sağlamaktadır. Bu şekil ve tablolar vasıtasıyla veriler daha anlamlı hale getirilerek çeşitli çıkarımların yapılması ve yukarıda bahsedilen konulardaki bilgilere ulaşılması mümkün olabilmektedir. Fakat verilerin sınıflandırılmasının manuel olarak yapılması durumu, özellikle çok sayıdaki veri kümeleri için oldukça zaman alıcı ve hata payının yüksek olduğu bir yöntemdir. Dolayısıyla bu durum hata payının yüksek olabileceği anlamına gelmektedir. Bu çalışmada, metin madenciliği ile bir sınıflandırma modeli oluşturulması ve hüküm verilerinin sınıflandırılarak bahsedilen dezavantajlı durumların çözülmesi hedeflenmektedir. Verilerin manuel olarak tasnif edilmeden, otomatik olarak metnin analizinin yapılmasını amaçlaması (Bayer, 2011: 44) ve büyük miktarda veri kümeleri ile çalışılmasına imkân sağlaması, sınıflandırma işleminde metin madenciliğinin uygulanması nedenleri arasındadır. Bir diğer neden ise verinin yapısıdır. Ele alınan veri kümesi yarı yapılandırılmış veri türündedir. Bu tür yapıya sahip olan veri türleri üzerinde anlamlı sonuçlar çıkarılabilmesi metin madenciliği araç ve yöntemleri ile mümkündür.

Veri kümesi incelendiğinde hükümlerin farklı birçok konudan oluştuğu gözlemlenmektedir. Dört ayrı tezden alınan hükümlerin konuları incelendiğinde konu başlıkları üzerinde farklılıklar olduğu belirlenmiştir. Bu durum şu şekilde örneklendirilebilir; Hükümlerin konularını belirlerken Kantaroğlu'nun (2019: 21) kullandığı başlıklardan birisi daha genel bir kavram olan güvenlik kelimesidir. Susam (2019: 19) ise çalışmasında Kantaroğlu'nun güvenlik başlığı altında topladığı hükümler ile benzerlik gösteren hükümlerin konularını belirlerken daha özel bir

kavram olan şekavet başlığını kullanmıştır. Bu durum ile transkript çalışmalarının belirli bir üstveri standardına bağlı kalınmadan yapıldığı sonucuna varılmaktadır.

Verilerin alındığı çalışmalarda yer alan hükümlere verilen başlıklar yeniden düzenlenmiştir. Bu başlıkların neler olduğu ve her bir konu ile ilgili kaç adet hüküm bulunduğu yüzdelik oranları ile birlikte aşağıdaki Tablo 3.1.'de gösterilmiştir. Ayrıca sınıflandırma işleminin öneminin vurgulanması amacıyla, elde edilen verilere göre sınıflandırma işlemi ile birlikte elde edilebilecek bazı tablo ve grafikler verilmiştir.

Tablo 3. 1: Hükümlerin Düzenlenmiş Konularına Göre Dağılımı.

Hüküm Sınıfları	Frekans	Yüzde
	<i>n</i>	%
Asayiş	268	22,1
Sosyal, Ekonomik ve İdarî Konular	432	35,7
Arazi Meselesi ve Vergi	509	42,1

Hükümlerin %22,1'i asayiş, %35,7'si sosyal, ekonomik ve idarî konular, %42,1'i ise arazi meselesi ve vergi sınıfı altında yer almaktadır.

Hükümlerin sınıflandırılması işleminde daha başarılı bir performans elde edilebilmesi için aşağıdaki Şekil 3.1'de bulunan konu başlıkları, içeriklerine göre 3 ana konu başlığı altında tekrar sınıflandırılmıştır. Bu işlemin temel sebebi şu şekilde açıklanabilir; hükümler otomatik olarak sınıflandırılırken söz konusu veriler eğitim ve test kümesi olmak üzere ikiye ayrılır. Sınıflandırma işleminde kullanılacak algoritmaların eğitim kümesindeki veriler ile eğitilmesinin ardından test kümesinde bulunan verileri doğru bir şekilde sınıflandırması beklenir. Bu işlemin başarılı bir şekilde gerçekleşmesi için ise algoritmanın eğitilmesi işleminde kullanılan verilerin sayısı önemlidir. Bir konu başlığı altında bir araya getirilmiş çok sayıdaki hüküm verisiyle eğitilen algoritmanın, konu başlığını temsil eden hükümleri öğrenmesi daha başarılı olur ve bu durum gerçekleştirilecek olan test verilerinin sınıflandırılması işleminin başarısını doğrudan etkileyecektir. Bu amaç doğrultusunda sayıca az olan

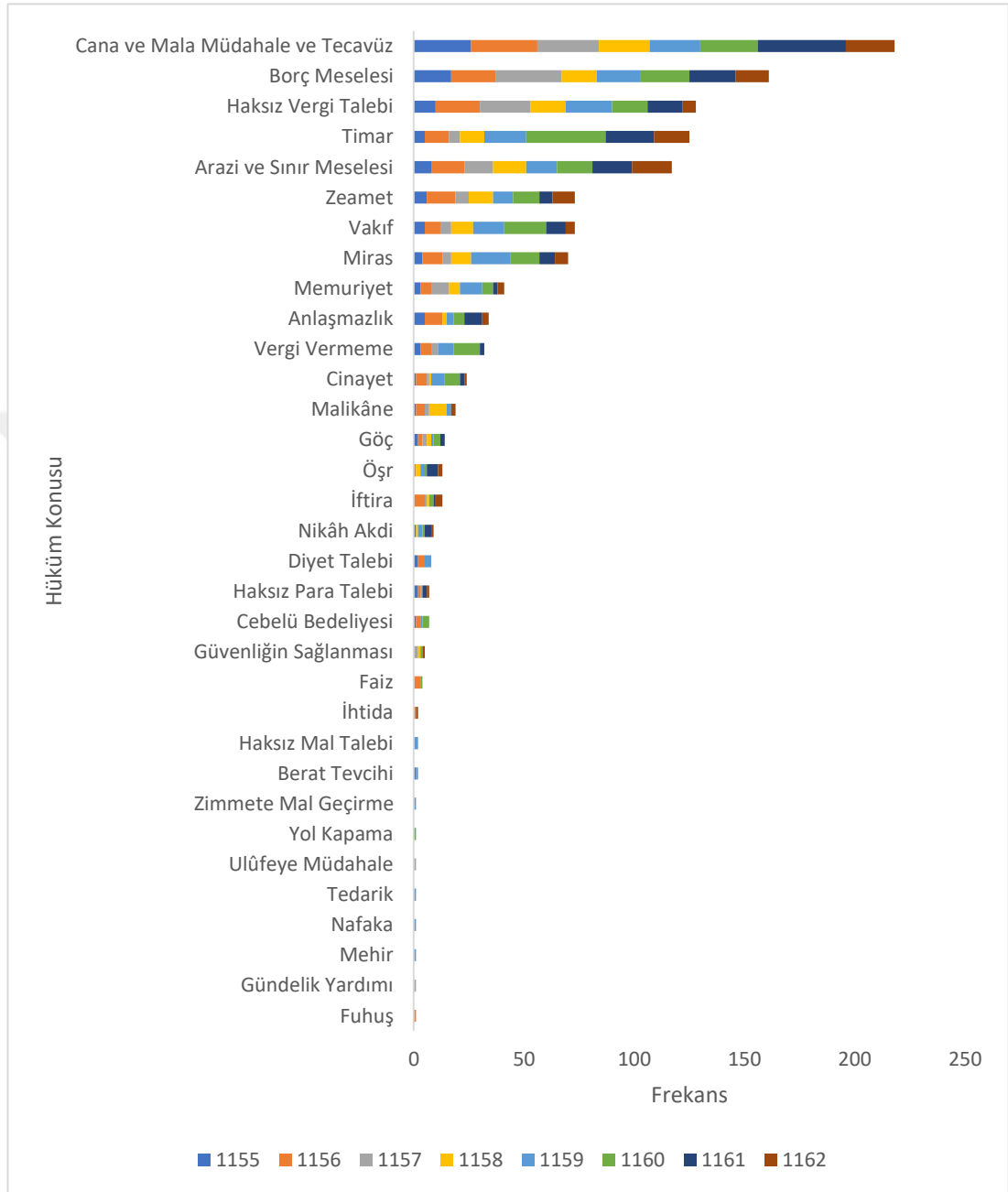
hükümler, daha kapsayıcı olan genel konu başlıkları altında yeniden düzenlenmiştir ve böylece gerçekleştirilen yeniden sınıflandırma işleminde hükümler belirli bir standarda göre bir araya getirilmiştir. Bu işlemle oluşturulan yeni konu başlıkları, alt konu başlıkları ve kapsadıkları hükümler şu şekildedir:

- **Asayiş:** Cana ve mala müdahale ve tecavüz (eşkîyalık, mal gaspı, para gaspı, tecavüz, darp ve hırsızlık ile ilgili hükümlerden oluşmaktadır), fuhuş, iftira, cinayet, güvenliğin sağlanması (can güvenliğinin sağlanması, resmi evrak ve değerli malların korunması ve şahıs güvenliğinin sağlanması ile ilgili hükümlerden oluşmaktadır), diyet talebi ve zimmete mal geçirme ile ilgili hükümler bu başlık altında bir araya getirilmiştir.
- **Sosyal, ekonomik ve idarî konular:** Vakıf (vakıf tevcihi, vakıf yönetimi tevcihi ve vakfa müdahale ile ilgili hükümlerden oluşmaktadır), borç meselesi, anlaşmazlık (mülk anlaşmazlığı, satış anlaşmazlığı ve su anlaşmazlığı ile ilgili hükümlerden oluşmaktadır), gündelik yardımı, yol kapama, göç, faiz, haksız para talebi, haksız mal talebi, ihtida, mehir, miras (haksız hisse talebi, miras ihtilafı, tereke taksimi ve mirasa müdahale ile ilgili hükümlerden oluşmaktadır), nikâh akdi, nafaka, berat tevcihi (sâ'ilik beratı ve müderrislik beratı ile ilgili hükümlerden oluşmaktadır), memuriyet (sorumluluklarını yapmama, göreve müdahale, vazifeyi kötüye kullanma, soruşturma, vazife tevcihi, ve tayin talebi ile ilgili hükümlerden oluşmaktadır) ve tedarik ile ilgili hükümler bu başlık altında bir araya getirilmiştir.
- **Arazi meselesi ve vergi:** Arazi ve sınır meselesi (mülke müdahale, sınır ihtilafı, meraya müdahale, yurtluğa ve ocaklığa müdahale, koruya müdahale ve yaylağa müdahale ile ilgili hükümlerden oluşmaktadır), zeamet (zeamete müdahale, zeamet tahsili ve zeamet mutasarrıflığı tevcihi ile ilgili hükümlerden oluşmaktadır), malikâne (malikâneye müdahale ve malikâne tahsili ile ilgili hükümlerden oluşmaktadır), timar (timar tevcihi, timara müdahale, timar mutasarrıflığı, timar tahsili

ve timar anlaşmazlığı ile ilgili hükümlerden oluşmaktadır), öşr (öşr tahsili ve öşr'e müdahale ile ilgili hükümlerden oluşmaktadır), haksız vergi talebi, vergi vermeme, cebelü bedeliyesi ve ulûfeye müdahale ile ilgili hükümler bu başlık altında bir araya getirilmiştir.

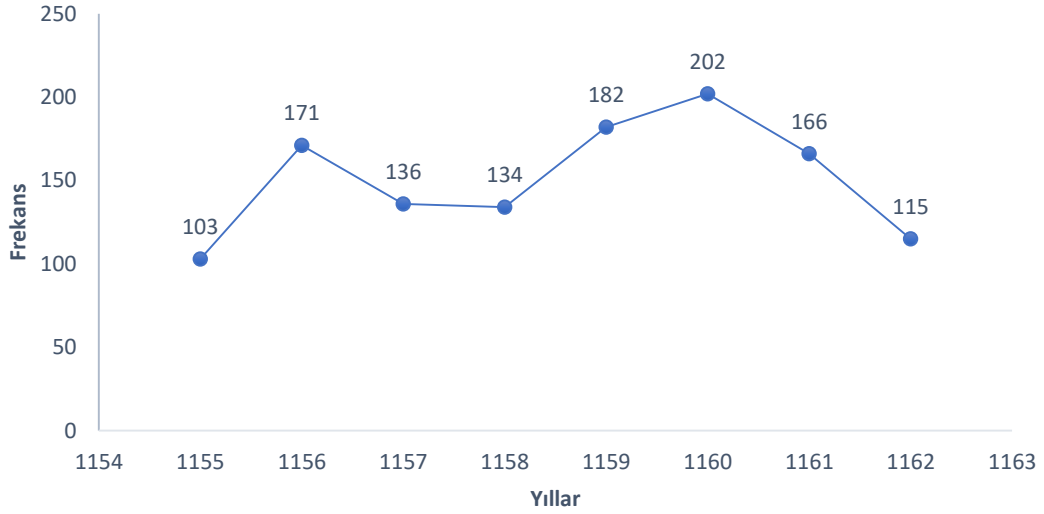


Şekil 3. 1: Yıllara Göre Hüküm Konularının Frekans Dağılımı.



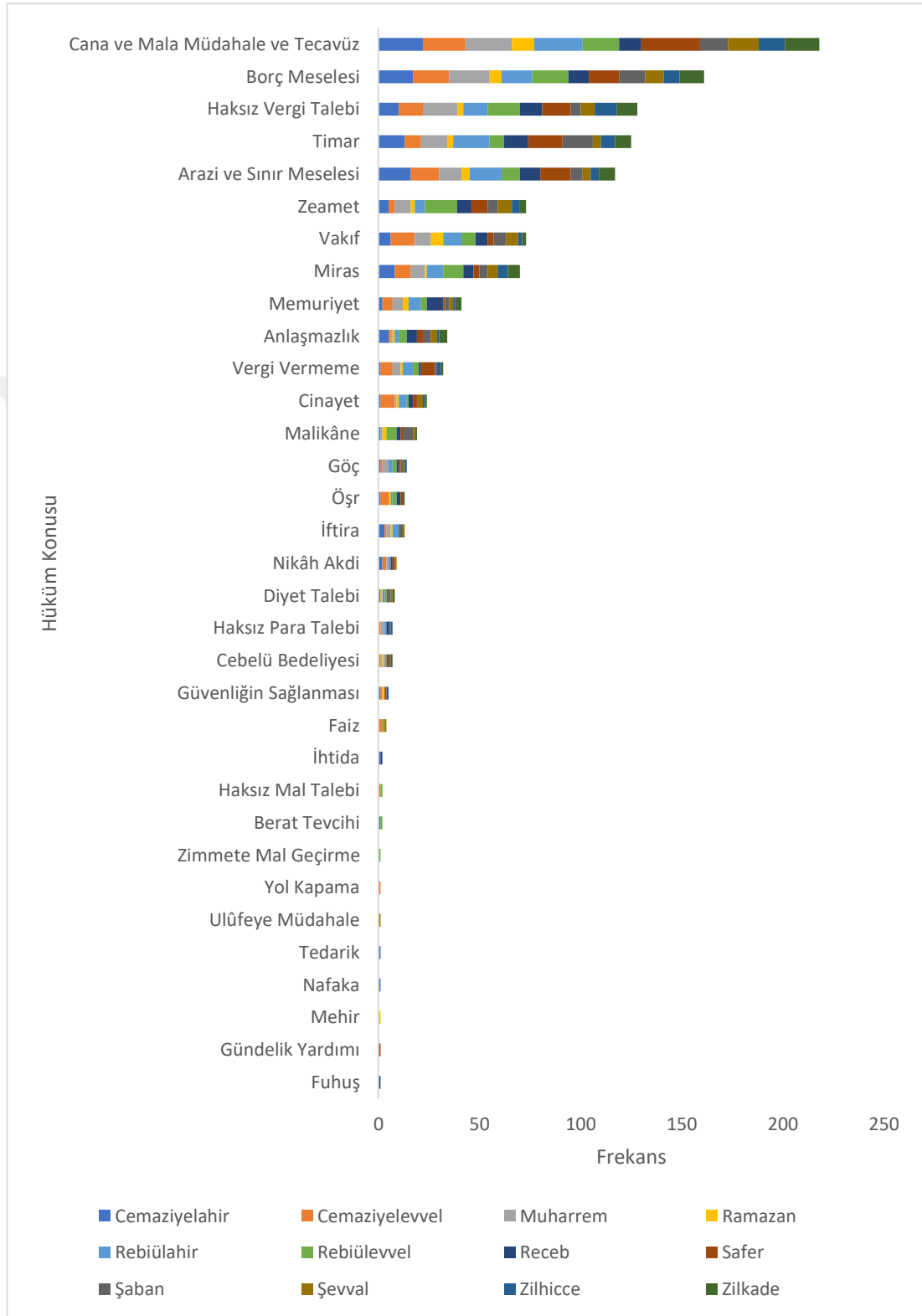
Yukarıdaki şekil incelendiğinde hüküm konularının frekans değerleri yıllara göre verilmiştir. Her bir yıl farklı renk ile temsil edilmiştir. En çok hükmün cana ve mala müdahale ve tecavüz konusu ile ilgili olduğu ve bu hükümlerin 40 adet hükümle en sık görüldüğü tarihin H.1161/M.1748 olduğu gözlemlenmektedir. Ayrıca H.1157/M.1744 tarihi, borç meselesi ile ilgili hükümlerin en sık görüldüğü tarihtir. Diğer hüküm konularının dağılımı Şekil 3.1.'de detaylı bir şekilde gösterilmiştir.

Şekil 3. 2: Yıllara Göre Hüküm Sayılarının Değişimine Ait Çizgi Grafiği.

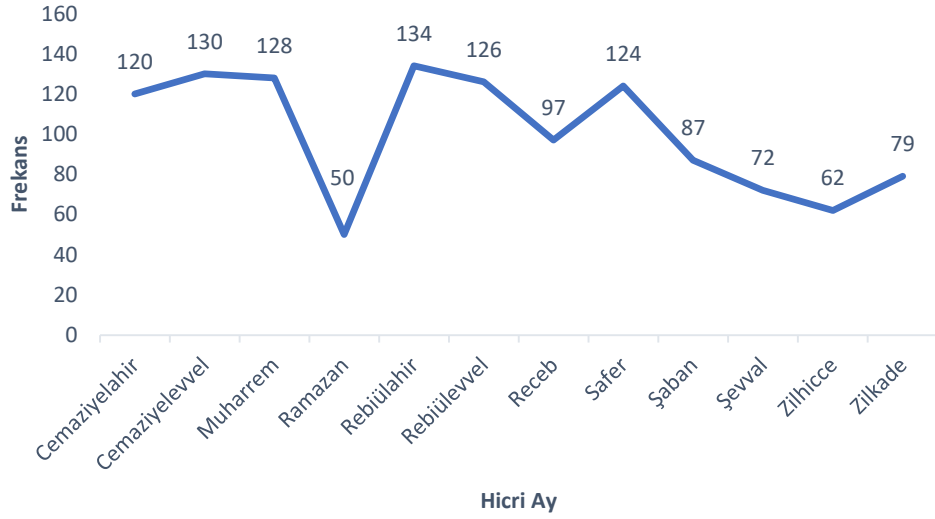


Şekil 3.2’de verilen grafikte H.1155/M.1742 ile H.1162/M.1749 tarihleri arasındaki hükümlerin frekans değerleri verilmiştir. H.1160/M.1747 yılında toplamda 202 hüküm deftere yazılmıştır. Bu tarih, Osmanlı İmparatorluğu padişahı I. Mahmud’un (M.1730-1754) dönemine denk gelmektedir (Özcan, 2003: 348). Grafikte bulunan bu veri I. Mahmud dönemi incelemelerinde yaşanan olaylara kaynak olma değeri taşımaktadır. Ayrıca Osmanlı İmparatorluğu’nun bahsi geçen dönemdeki mevcut durumu, iç ve dış siyasî ve askerî politikaları, bu grafiğin anlamlandırılmasını sağlayabilecektir. Bu durum verilerin sınıflandırılmasını ve dolayısıyla elde edilen grafiklerin önemini vurgulamaktadır.

Şekil 3. 3: Aylara Göre Hüküm Konularının Sayıları.



Şekil 3. 4: Aylara Göre Hüküm Sayılarının Değişimi.



Hüküm verileri kümesinden elde edilebilecek bir diğer tablo ise konuların aylara göre dağılımıdır. Her bir hüküm konusunun hangi ayda ne kadar görüldüğü ile toplamda bir ayda ne kadar hüküm görüldüğüne dair grafikleri elde etmek mümkündür. Böylece olay ve konu incelemelerinde bahsedilen grafikler hem başvurulabilecek birer kaynak niteliği taşımaktadır hem de çalışma yapılacak yeni konulara ilham olabilecektir.

3.4. Veri Ön İşleme

Excel’de oluşturulmuş olan veri kümesi yarı yapılandırılmış metinsel veri türündedir. Bu tür veriler üzerinde analizlerin yapılabilmesi için verilerin yapılandırılmış veri haline dönüştürülmesi gerekmektedir. 3 ana konu başlığı altında toplanmış olan verilere, KNIME programı vasıtasıyla bazı veri ön işleme adımları uygulanacaktır. Verilere uygulanan ön işlem aşamaları, verilerin analiz edilmeden önce işlenmeye uygun bir yapıya getirilmesini sağlamasının yanı sıra analizlerin doğruluğu ve kaliteli sonuçların elde edilmesi için önem arz etmektedir (Emre ve Selçukcan Erol, 2017: 165).

Veri ön işleme kısmında KNIME programında sınıflandırma modeli oluşturma işlemi için kullanılmış olan düğümlerden ve bu düğümlerin işlevlerinin neler olduğundan bahsedilecektir. Aşağıda KNIME programı içerisinde bulunan

düğümlerden, hüküm verilerini sınıflandırma modeli oluşturulurken kullanılan düğümler açıklanmıştır. Söz konusu düğümler vasıtasıyla hüküm verileri üzerinde metin madenciliği uygulaması gerçekleştirilmiştir.

3.4.1. Excel Reader (Excel Okuyucu)

KNIME programı farklı dosya türleri ile çalışılabilmeyi mümkün kılmaktadır. Bu çalışmanın verileri Excel’de bir araya getirilmiştir. Bu düğüm ile oluşturulan Excel veri dosyası KNIME programına yüklenmiştir. Söz konusu düğüm ve program ara yüzünde verilerin KNIME programına aktarıldıktan sonraki görünümü Resim 3.1. ile Resim 3.2.’de verilmiştir.

Resim 3. 1: Excel Reader Düğümü.



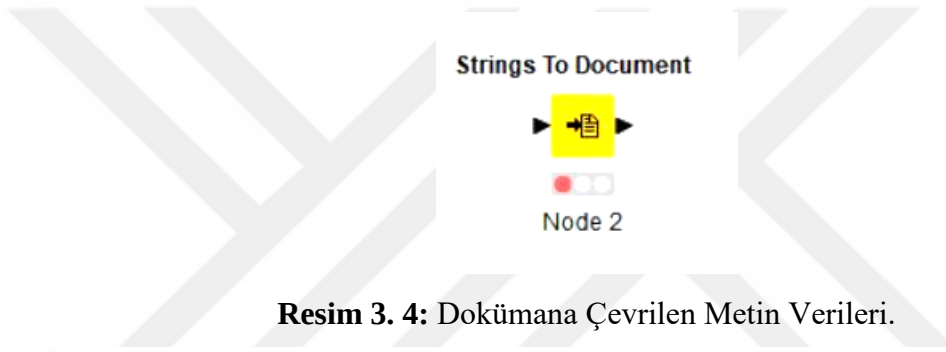
Resim 3. 2: Excel Reader Düğümüne Aktarılan Hüküm Verileri.

Row ID	Sıra	Sahife No	Hüküm no	Tarhi Hicri	Tarhi miladi	Hüküm	Hüküm...	Hükümün özet
Row12	13	3	13	[20] Cemaziyelahir 1...	[22 Ağustos 1742]	Erzurum vâlisine ve Ulubey ve () kâdılarına hüküm Abdülbaki nâm kimesne gelüp Ulub...	Toprak Mes...	Abdülbaki adı kişinin tasarrufunda olan
Row13	14	3	14	[20] Cemaziyelahir 1...	[22 Ağustos 1742]	Karahisar-ı Şariki sancağı mutasarrıfına, Bayramlı ve () kâdısına hüküm ki Bayramlı kazısına...	Alacak Verecek	Kadzâde Hasan'ın Sucuoğlu Hüseyin ac
Row14	15	3	15	[20] Cemaziyelahir 1...	[22 Ağustos 1742]	Erzurum vâlisine, Ulubey ve () kâdılarına hüküm Abdülbaki nâm kimesne gelüp Ulubey kaz...	Güvenlik	Abdülbaki adı kişinin kendisine intikal ec
Row15	16	3	16	[29] Cemaziyelahir 1...	[31 Ağustos 1742]	Erzurum vâlisine ve Pazarsuyu nâibine hüküm ki Pazarsuyu kazası mutâfâsından Kırık nâhi...	Güvenlik	Pazarsuyu kazasında Fark(?) nahiyesi
Row16	17	4	17	[29] Cemaziyelahir 1...	[31 Ağustos 1742]	Erzurum vâlisine ve Bayburd kâdısına hüküm Bayburd kazasına tâbi Feriuhadbey Evkâfı kar yerlerinden Pulur nâm kar ye...	Alacak Verecek	Bayburd'da Pulur adı kar ye ahalisinin b
Row17	18	4	18	[29] Cemaziyelahir 1...	[31 Ağustos 1742]	Erzurum vâlisine, Bayburd ve () kâdılarına hüküm ki E'immeden Ahmed Halife gelüp Baybur...	Vergi	Bayburd kazasına bağı Vaorek(?) kar ye
Row18	19	4	19	[29] Cemaziyelahir 1...	[31 Ağustos 1742]	Erzurum vâlisine ve Erzurum kâdısına hüküm ki Van sâkinlerinden Elhan ve Manok ve Ebercan ve Bedros ve Toma ve Parsak...	Alacak Verecek	Van'da ikamet eden bazı zimmilerin Erzu

3.4.2. Strings To Document (Doküman Dizeleri)

KNIME içerisinde yer alan diğer bir düğüm olan Strings To Document ile girişi yapılmış olan veri kümesi tablosunun içerisindeki satır ya da sütunlar yeni bir sütun oluşturularak dokümana dönüştürülür. Bu düğüm ile gerçekleştirilen işlemde Excel programında bir araya getirilen hüküm verileri satır ve sütun olarak ayrı birer doküman şeklinde tanımlanır. Bu durum ile amaçlanan, her bir verinin üzerinde işlem yapılabilir hale getirilmesi için öncelikle ayrı birer doküman olarak tanımlanmasıdır.

Resim 3. 3: String To Document Düğümü.



Resim 3. 4: Dokümana Çevrilen Metin Verileri.

String and document output table - 02 - Strings To Document (Metinleri dokümana)

File Edit Hilitte Navigation View

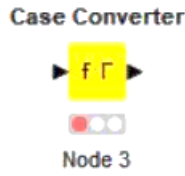
Table "default" - Rows: 1216 | Spec - Columns: 11 | Properties | Flow Variables

Row ID	Hüküm	Document
Row0	Bayburd kâdsına hüküm: Abdullah zide kadruhü gelup Erzurum sancağında Bayburd kazısına tâbi' Pulur nâm karyede vâkı' Ferruhâdbey Câmî' ve İmâretî evkâfının mütevellis olan Yakub nâm kimesne vakf-ı mezbûrû ferâğ ve sal müteakken vazifelerini vermişçüp kull gadr-ı valî' gerîf harâb eyleduğın ahâlden on beş nefer kimesneler haber verdiklerin sadr-ı Rumelî sâbık a'lemü'ülemâ'tî mütebahhırin Pîr-zâde Mevlânâ Muhammed	"Bayburd kâdsına hüküm:Abdullah zide kadruhü gelup Erzurum sancağında
Row1	Karahisar-ı Şarkî sancağı mutasarrıfına () kâdsına ve Yakubbey kazası nâ'bine hüküm ki. Sen ki Yakubbey kazası nâ'ibi Mevlânâ Mustafa zide ilmuhsun. Mektûb gön...	"Karahisar-ı Şarkî sancağı mutasarrıfına () kâdsına ve Yakubbey kazası r
Row2	Ber-vech-i arpalık Karahisar-ı Şarkî sancağı mutasarrıfına ve () kâdsına ve Yakubbey kazası nâ'bine hüküm Sen ki Yakubbey kazası nâ'ibi Mevlânâ Mustafa zide ilmuhsun.	"Ber-vech-i arpalık Karahisar-ı Şarkî sancağı mutasarrıfına ve () kâdsına
Row3	Erzurum vâlisine ve Erzurum monlâsına ve Kuruçay kâdsına hüküm ki Kuruçay kazısına tâbi' Kuçukarmudân nâm karye sâkinlerinden Yanı Kola ve Dimitri ve Yana...	"Erzurum vâlisine ve Erzurum monlâsına ve Kuruçay kâdsına hüküm
Row4	Erzurum vâlisine ve Erzurum monlâsına ve () ve () kâdlarına hüküm ki Uzun Ali nâm kimesne gelup bunun Erzurum'da sâkin Tebrizî Hacı Ali nâm kimesne zimmetinde bi...	"Erzurum vâlisine ve Erzurum monlâsına ve () ve () kâdlarına hüküm ki U
Row5	Erzurum vâlisine ve Erzincan kâdsına hüküm Ulemâdan Şeyh Ahmed zide ilmuhsun gelup bu Erzincan'da vâkı' Molla Ahmed Mükrimî vakfından olmak üzere vazife-i mu'ay... kâmlen kendüye alverilüp muğayir-ı fetvâ-yı şerife bir vechile ta'allül ve mümâna'at ve âhardan bi-vech dâhil ve mümâna'at etdirilmemek bâbında hüküm-i hümayûnı recâ eyleduğı eclden gerle görilmek için yazılmışdır.	"Erzurum vâlisine ve Erzincan kâdsına hüküm Ulemâdan Şeyh Ahmed zide
Row6	Erzurum vâlisine ve Hıns ve () kâdlarına hüküm ki Erbâb-ı tımdan Abdurrahman Sipahi gelup Hıns sancağında ve nâhiyesinde Sarumustafa nâm karye ve gayriden ...	"Erzurum vâlisine ve Hıns ve () kâdlarına hüküm ki Erbâb-ı tımdan Abd

3.4.3. Case Converter (Büyük-Küçük Harf Dönüştürücü)

Alfa-nümerik karakterler de dâhil olmak üzere veri içerisinde bulunan harflerin tamamının büyük ya da küçük harfe dönüştürülmesini sağlayan düğümdür. Bu düğümün kullanılma nedeni şu örnekle açıklanabilir; metin içerisinde baş harf büyük harfle yazılmış Kalem kelimesi ile baş harf küçük harfle yazılmış olan kalem kelimesi, KNIME programı tarafından farklı iki karakter olarak algılanacaktır. Bu durum aslında aynı olan Kalem kelimesi için ayrı, kalem kelimesi için ayrı frekans değerinin hesaplanmasına neden olacaktır. Bahsedilen durumun yaşanmaması için hüküm verilerini oluşturan kelimelerin tamamı küçük harfe dönüştürülmüştür. Bu işlem Case Converter düğümü ile gerçekleştirilmiştir. İlgili düğüm ve küçük harfe dönüştürülen hüküm verileri Resim 3.5. ve Resim 3.6.'da gösterilmiştir.

Resim 3. 5: Case Converter Düğümü.



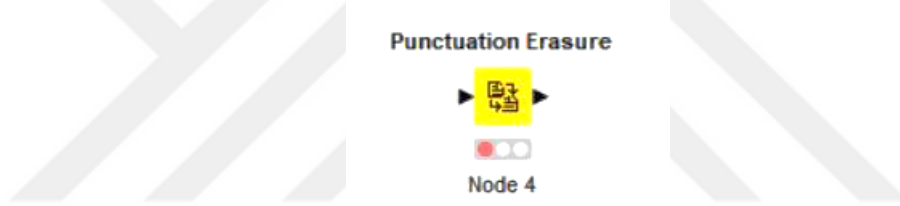
Resim 3. 6: Küçük Harfe Dönüştürülen Hüküm Verileri.

Preprocessed documents. - 0.3 - Case Converter (Büyük/küçük harf)		
File Edit Hilitte Navigation View		
Table "default" - Rows: 1216 Spec - Columns: 12 Properties Flow Variables		
Row ID	Document	Preprocessed Document
Row0	"Bayburd kâdsına hüküm:Abdullah zide kadruhü gelup Erzurum sancağında Bayburd kazâsına tâbi" Pulnâm karyede vâki' Ferruhşâdbey ...	"bayburd kâdsına hüküm:abdullah zide kadruhü gelup erzurum sancağında bayburd kazâsına tâbi"
Row1	"Karahisar-ı Şarkî sancağı mutasarrıfına () kâdsına ve Yakubbey kazâsı nâ'ibine hüküm ki. Sen ki Yakubbey kazâsı nâ'ibi Mevlânâ Mustafa ...	"karahisar-ı şarkî sancağı mutasarrıfına () kâdsına ve yakubbey kazâsı nâ'ibine hüküm ki. sen ki yak"
Row2	"Ber-vech-i arpaik Karahisar-ı Şarkî sancağı mutasarrıfına ve () kâdsına ve Yakubbey kazâsı nâ'ibine hüküm Sen ki Yakubbey kazâsı nâ'ibi ...	"ber-vech-i arpaik karahisar-ı şarkî sancağı mutasarrıfına ve () kâdsına ve yakubbey kazâsı nâ'ibin"
Row3	"Erzurum vâlisine ve Erzurum montâsına ve Kuruçay kâdsına hüküm ki Kuruçay kazâsına tâbi" Küçükarmudân nâm karye sâkinlerinden ...	"erzurum vâlisine ve erzurum montâsına ve kuruçay kâdsına hüküm ki kuruçay kazâsına tâbi" ki.
Row4	"Erzurum vâlisine ve Erzurum montâsına ve () ve () kâdlarına hüküm ki Uzun Ali nâm kimesne gelup bunun Erzurum'da sâkin Tebridi Hao Al...	"erzurum vâlisine ve erzurum montâsına ve () ve () kâdlarına hüküm ki uzun ali nâm kimesne gelup i"
Row5	"Erzurum vâlisine ve Erzincan kâdsına hüküm Ulemâdan Şeyh Ahmed zide ilimüh gelup bu Erzincan'da vâki' Molla Ahmed Mükrimî vakfında...	"erzurum vâlisine ve erzincan kâdsına hüküm ulemâdan şeyh ahmed zide ilimüh gelup bu erzincan"
Row6	"Erzurum vâlisine ve Hınıs ve () kâdlarına hüküm ki Fihrah-ı İmâretden Ahîd ırrahman Snahi nâlin Hınıs sancağında ve nâ'ibinde Sanım k...	"erzurum vâlisine ve hınıs ve () kâdlarına hüküm ki erhâh-ı imâretden ahîd ırrahman snahi nâlin hınıs"

3.4.4. Punctuation Erasure (Noktalama İşaretlerini Silme)

Bu düğüm, verilerin içerisinde bulunan noktalama işaretlerinin kaldırılmasını sağlamak için kullanılan düğümdür. Veri kümesini oluşturan hükümler transkript edilirken okunamayan kelimeler parantez içerisinde üç nokta ya da içi boş parantez ile belirtilmiştir. Case Converter düğümünün kullanılma nedeninde de bahsedildiği üzere, program, bu karakterleri ayrı birer karakter olarak algılayacaktır. Böylece yapılacak işlemlerde programın işlem hacminin artmasına neden olacaktır. Bu sebepler doğrultusunda hüküm verilerinde yukarıda belirtilmiş olan durumların yaşanmaması adına noktalama işaretlerinin kaldırılması için Punctuation Erasure düğümü kullanılmıştır.

Resim 3. 7: Punctuation Erasure Düğümü.



Resim 3. 8: Noktalama İşaretlerinin Kaldırıldığı Hüküm Verileri.

Preprocessed documents. - 0:5 - Punctuation Erasure (Noktalama işaretlerini)

File Edit Hilitte Navigation View

Table "default" - Rows: 1216 Spec - Columns: 12 Properties Flow Variables

Row ID	Document	Preprocessed Document
Row0	"Bayburd kâdsına hüküm:Abdullah zide kadruhü gelup Erzurum sancağında Bayburd kazısına tâbi' Pulunâm karyede v..."	"Bayburd kâdsına hükümAbdullah zide kadruhü gelup erzurum sancağında bayburd kazısına tâbi' pulunâm karyede v..."
Row1	"Karahisar-ı Şarîî sancağı mutasarrıfına () kâdsına ve Yakubbey kazısı nâ'bine hüküm ki. Sen ki Yakubbey kazısı nâ'bi m..."	"Karahisar-ı şarîî sancağı mutasarrıfına kâdsına ve yakubbey kazısı nâ'bine hüküm ki sen ki yakubbey kazısı nâ'bi m..."
Row2	"Ber-ved-i arpalık Karahisar-ı Şarîî sancağı mutasarrıfına ve () kâdsına ve Yakubbey kazısı nâ'bine hüküm Sen ki Yaku..."	"ber-ved-i arpalık karahisar-ı şarîî sancağı mutasarrıfına ve kâdsına ve yakubbey kazısı nâ'bine hüküm sen ki yakub..."
Row3	"Erzurum vâlsine ve Erzurum monlâsına ve Kuruçay kâdsına hüküm ki Kuruçay kazısına tâbi' Kuçukarmudân nâm k..."	"erzurum vâlsine ve erzurum monlâsına ve kuruçay kâdsına hüküm ki kuruçay kazısına tâbi' kuçukarmudân nâm k..."
Row4	"Erzurum vâlsine ve Erzurum monlâsına ve () kâdsına hüküm ki Uzun Ali nâm kimesne gelup bunun Erzurum'da s..."	"erzurum vâlsine ve erzurum monlâsına ve ve kâdsına hüküm ki uzun alı nâm kimesne gelup bunun erzurumda sâkin..."
Row5	"Erzurum vâlsine ve Erzurum kâdsına hüküm Ulemâdan Şeyh Ahmed zide ilmuhu gelup bu Erzurum'da vâki' Molla Ahme..."	"erzurum vâlsine ve erzurum kâdsına hüküm ulemâdan şeyh ahmed zide ilmuhu gelup bu erzurumda vâki' molla ahme..."

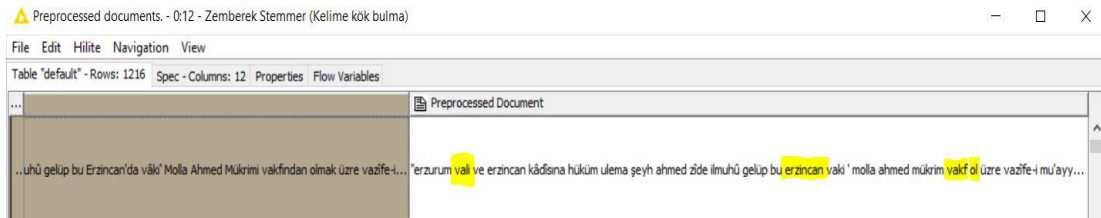
3.4.5. Zemberek Stemmer (Zemberek Kök Bulucu)

Zemberek, Türkçe için geliştirilmiş açık kaynak kodlu bir doğal dil işleme kütüphanesidir (Akın ve Akın, 2007: 1). Diğer bir ifade ile Türkçe metinler için kelimelerin kökünü elde etmeye yarayan algoritmadır. Bu işleme gövdeleme (kök bulma) denilmektedir ve bu konuya çalışmanın birinci bölümünde aynı isimle oluşturulmuş başlık altında değinilmiştir. Oluşturulan veri kümesi Osmanlı Türkçesi ile yazılmış hüküm verilerini içermektedir. Hükümler içerisinde oldukça fazla sayıda Arapça ve Farsa kelime ile tamlamalar mevcuttur. Farklı dil kaynaklı bu kelimelerin köklerini bulmak için geliştirilmiş bir kütüphane ya da algoritma henüz mevcut değildir. Dahası bu kelimelerin köklerini Zemberek Kütüphanesi kullanılarak elde etmek mümkün değildir. Zemberek Kütüphanesi, oluşturulan modelde hükümlerin içerisinde var olan Türkçe kelimelerin köklerine ayrılmasını sağlamak amacıyla kullanılmıştır. Zemberek Kütüphanesi tarafından köklerine ayıramayan kelimeler için manuel olarak işlem yapılmıştır. Resim 3.10.'da görüldüğü üzere hüküm verisinde bulunan ve Zemberek Kütüphanesi tarafından algılanabilen kelimeler eklerinden arındırılmıştır. Söz konusu kelimeler Resim 3.10.'da gösterilmiştir.

Resim 3. 9: Zemberek Stemmer Düzümü.



Resim 3. 10: Zemberek Stemmer ile Köklerine Ayrılan Kelimeler.



3.4.6. Number Filter (Sayı Filtresi)

Veri kümesi içerisinde bulunan sayısal ifadeleri kaldırmaya yarayan düğümdür. Hükümlerin içerisinde geçen sayısal ifadeler tek başına bir anlam ifade etmeyeceğinden veri boyutunda azaltma sağlanması hedeflenmiştir. Sayısal ifadelerin kaldırılması işleminin uygulanması sürecine karar verilirken veri ön işleme aşamasından önce veri kümesinden elde edilmek istenilen sonuçların belirlenmesi gerekmektedir şeklinde yorum yapılabilir. Bu yorumun sebebi ise veri kümesi içerisinde bulunan sayısal ifadelerin, verilerin anlamlandırılması için önemli olabilmesi durumudur. Örneğin bir kişinin öldükten sonra bıraktığı mallarının kayıtlarını içeren bir veride sayısal ifadeleri atmak doğru bir yaklaşım olmayacaktır. Bu sebeple üzerinde çalışma yapılacak olan veri kümesinin veri ön işleme adımlarından önce detaylı bir şekilde incelenmesi ve elde edilmek istenilen sonucun çalışma öncesinde detaylıca belirlenmesi gerekmektedir. Bu çalışmada hükümlerin konularına göre otomatik olarak sınıflandırılması işlemi hedeflenmiştir. Bu sebeple hükümlerin sınıflandırılması işleminde hükümlerin içerisinde bulunan sayısal ifadelerin çalışmanın sonucuna etki etmeyeceği öngörülmüştür ve bahsedilen düğüm kullanılarak sayısal ifadeler kaldırılması hedeflenmiştir. Fakat hüküm verileri incelendiğinde sayısal ifadelerin rakam kullanılarak yazılmadığı belirlenmiştir. Number Filter düğümünün oluşturulan sınıflandırma modelinde kullanılmasının sebebi ise şu şekilde ifade edilebilir: Hüküm verileri Excel programında bir araya getirilirken yanlışlıkla sayısal ifadelerin hüküm metinleri içerisine yazılmış olma ihtimali göz önünde bulundurulmuştur. Dolayısıyla hataların minimize edilmesi ve yukarıda bahsedildiği üzere veri boyutunun azaltılmasının sağlanması amacıyla söz konusu düğüm kullanılmıştır.

Resim 3. 11: Number Filter Düğümü.



3.4.7. Stop Word Filter (Durak Kelime Filtresi)

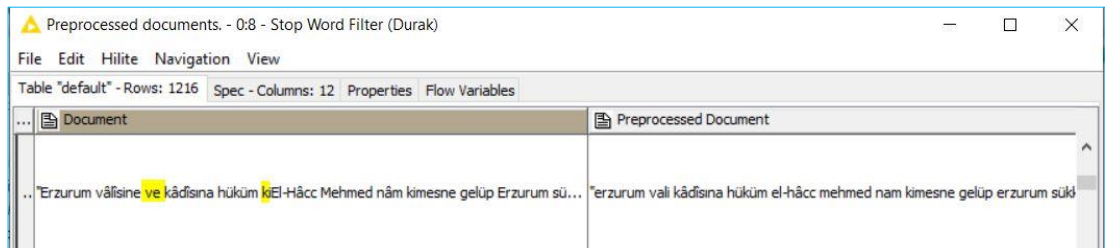
Bu düğüm ile veri kümesi içerisinde bulunan bağlaç ve edatlar çıkarılmaktadır. Bu kelimelerin ortak özelliği, yapılacak olan analizlerde ve hedeflenen sonuçların elde edilmesinde anlamlı bir katkılarının olmayışdır. Bu konu durak kelimelerin çıkarılması başlığı altında detaylı bir şekilde incelenmiştir. Veri setinde bulunan Arapça ve Farsça kelime ve tamlamalar içerisindeki durak kelimelerin çıkarılması için bir kütüphane bulunmamaktadır. Bu sebeple yapılacak olan yeni çalışmalarda kullanılmak üzere Osmanlı Türkçesi durak kelimelerini içeren bir sözlük oluşturulabilir. Böylece olası bir Osmanlı Türkçesi durak kelimeleri sözlüğü, gerçekleştirilecek olan çalışmaların başarı oranına büyük katkı sağlayacaktır. Söz konusu düğümün bu çalışmada kullanılmasının sebebi, hüküm verileri içerisinde bulunan ve algoritmanın algılayabildiği Türkçe bağlaç ve edatların çıkarılarak işlem hacminin düşürülmesini sağlamaktır.

Resim 3. 12: Stop Word Filter Düğümü.



Resim 3.13. incelendiğinde hüküm verisinin başında bulunan “ve” bağlacı ile “ki” bağlacı, Stop Word Filter düğümü ile veri kümesinden ayıklamıştır.

Resim 3. 13: Durak Kelimelerin Hüküm Verilerinden Çıkarılması.



3.4.8. Dictionary Replacer (Sözlük Değiştirici)

Veri kümesi içerisinde bulunan terimlerin istenilen terimlerle değiştirilmesine olanak sağlayan düğümdür. Verileri oluşturan hükümler dört farklı kişi tarafından çevrilmiştir. Bu durum aynı kelimenin farklı yazılmasına neden olmuştur. Örneğin veri kümesi incelenirken İsmail isminin İsma'il ve İsmâ'il olarak toplamda üç farklı şekilde yazıldığı gözlemlenmiştir. Frekans değeri belirlenirken doğru sonuçların elde edilebilmesi için bu gibi durumların giderilerek belirli bir standardın oluşturulması gerekmektedir. Hüküm verileri Osmanlı Türkçesi ile yazılmıştır. Bu sebeple günümüzde kullanılan bazı kelimelerin farklı şekillerde kullanılmış olduğu görülmektedir. Bu duruma örnek olarak “yazılmışdır” kelimesi verilebilir. Söz konusu kelime günümüzde “yazılmıştır” şeklinde kullanılmaktadır. Efendioğlu ve İşcan (2010: 138), çalışmalarında Türkçe dil bilgisine göre sert ünsüzlerden sonra sert ünsüzlerin geldiğine ve yumuşak ünsüzlerden sonra ise yumuşak ünsüzlerin geldiğine değinmişlerdir. Bu durumun ünsüz ile başlayan ekler için önemli olduğunu ve bu sebeple birçok ekin hem yumuşak ünsüz hem de sert ünsüz hallerinin bulunduğunu belirtmişlerdir. Bu ses olayını ise “ünsüz uyumu” başlığı altında incelemişlerdir. “Yazılmışdır” kelimesinin günümüzde “yazılmıştır” şeklinde kullanılmasının, yukarıda bahsedilen bilgi doğrultusunda ünsüz uyumu ile ifade edilmesi mümkündür. Sınıflandırma modelinde kullanılan düğümlerden, Türkçe kelimelerin kökünü bulma işleminde kullanılan Zemberek Stemmer ve Türkçe durak kelimelerin çıkarılması için kullanılan Stop Word Filter düğümleri, Osmanlı Türkçesi ile oluşturulmuş kelimeler üzerinde (yazılmışdır örneğinde olduğu gibi) başarılı olamamıştır. Bu sebeple oluşturulan sınıflandırma modelinde Dictionary Replacer düğümü kullanılmıştır. Verilen örneklerdeki duruma benzer kelimeler, bir standarda bağlı kalınmak koşuluyla kök haline getirilerek düzenlenmiş ve ortaya çıkan yeni kelime listesi farklı bir düğümlerle Dictionary Replacer düğümüne tanıtılmıştır. Böylece ilgili kelimelerin yer değiştirme işlemi gerçekleştirilmiştir ve ilgili düğümler tarafından algılanabilir hale getirilen kelimeler üzerinde ön işlem aşamaları uygulanmıştır.

Resim 3. 14: Dictionary Replacer Düğümü.



Resim 3. 15: Dictionary Replacer Düğümü İle Değiştirilmiş Kelime Verileri.

Preprocessed documents. - 0:48 - Dictionary Replacer (Kelimeleri)		Preprocessed documents. - 0:48 - Dictionary Replacer (Kelimeleri)	
File Edit Hilite Navigation View		File Edit Hilite Navigation View	
Table "default" - Rows: 1216 Spec - Columns: 12 Properties Flow Variables		Table "default" - Rows: 1216 Spec - Columns: 12 Properties Flow Variables	
Row ID		Row ID	
Row25	olmağın berât mücebine zabt etdirilmek için yazılmışdır."	Row25	ol berat müceb zabt et için yaz"

3.4.9. Table Creator (Tablo Oluşturucu)

Manuel olarak veri girişinin yapılmasına olanak sağlayan düğümdür. Yukarıda bahsedilen yeni kelime listesi, bu düğüm vasıtasıyla Dictionary Replacer düğümüne tanıtılmıştır.

Resim 3. 16: Table Creator Düğümü.



Resim 3. 17: Table Creator Düğümüne Aktarılan Veriler.

Manually created table - 0:14 - Table Creator (Kelime listesi)

File Edit Hilitte Navigation View

Table "default" - Rows: 1559 Spec - Columns: 2 Properties Flow Variables

Row ID	column1	column2
Row0	.alakas	alakas
Row1	.ali	ali
Row2	.amel	amel
Row3	.amremde	amire
Row4	.ammisi	ammisi
Row5	.an	an
Row6	.arz	arz
Row7	.avanz	avanz
Row8	.ilet	ilet
Row9	.orf	orf
Row10	.orfe	orfe
Row11	.ogr	ogr
Row12	.oglerin	ogr
Row13	abbas	abbas
Row14	abdullah	abdullah
Row15	abdurrahman	abdurrahman
Row16	aded	aded
Row17	adem	adem
Row18	adem-i	adem
Row19	adem-i	adem
Row20	adet-i	adet
Row21	adet-i	adet
Row22	agob	agob
Row23	agrâzdan	agrâz
Row24	ağa	ağa
Row25	ağasa	ağa
Row26	agrâz	agrâz
Row27	ahali	ahali
Row28	ahâf-i	ahali
Row29	ahâf-i	ahali
Row30	ahar	ahar
Row31	âher	âher
Row32	âherde	âher
Row33	âherden	âher
Row34	âhere	âher
Row35	âheri	âher
Row36	âherin	âher
Row37	ahkam	ahkam

3.4.10. N Chars Filter (N Karakter Filtresi)

Karakter filtreleme olarak da bilinen bu düğüm vasıtası ile n sayıdaki karakterlerin veri kümesinden çıkarılması mümkün olmaktadır. Daha açık bir şekilde ifade edilecek olunursa n değişkenine verilen değer kadar harfle oluşturulmuş kelimelerin veri kümesinden çıkarılmasını sağlamaktadır. Örnek vermek gerekirse; bir ya da iki harfli kelimelerin yapılacak olan analizlerde anlam ifade etmeyeceği düşünüldüğünden n değeri iki seçilirse iki harfli olan kelimeler veri kümesinin dışında tutulacaktır. Çalışmada n değeri iki olarak belirlenmiştir.

Resim 3. 18: N Chars Filter Düğümü.



3.4.11. Dictionary Tagger (Sözlük Etiketleyici)

Yukarıda bahsedilen ve elle yeni tablo oluşturulmasını mümkün kılan Table Creator düğümü ile verilen varlıkları tanıyarak etiketleme yapmak için kullanılan düğümdür. Hüküm verilerinden elde edilmiş olan 1559 adet anahtar kelime içerisinde kişi adları ve yer adları belirlenmiştir. Bu adlar Table Creator düğümüne yeni bir tablo olarak işlenmiştir. Tablo Creator düğümü ile Dictionary Tagger düğümü birbirine bağlanarak ilgili adların etiketlenmesi gerçekleştirilmiş olur.

Resim 3. 19: Dictionary Tagger Düğümü.



3.4.12. Bag of Words Creator (Kelime Çantası Oluşturma)

Kelime çantası oluşturma olarak da adlandırılan bu düğüm ile her bir veri kaynağı kelimelere ayrılır. Bu işlem frekans hesaplamaları için kolaylık sağlamaktadır. Söz konusu düğüm vasıtasıyla hüküm verilerini oluşturan metinlerin kelimelere ayrılması işlemi gerçekleştirilmiştir. Metinleri oluşturan her bir kelimenin ayrı birer veri olarak ele alınmasını ve böylelikle kelimelerin frekans değerlerinin hesaplanabilir hale gelmesini sağlamak söz konusu düğüm ile mümkündür.

Resim 3. 20: Bag of Words Creator Düğümü.



Resim 3.21.'de gösterilen tabloda kelime çantası oluşturma işlemi gerçekleştirilmiştir. Söz konusu işlemde her bir hüküm, içerisinde bulunan her bir kelime için ayrı birer satır olarak ele alınmaktadır. Bir hüküm içerisinde bulunan kelimeler

bittikten sonra diğer hükümler için de söz konusu işlem tekrarlanmaktadır. Yukarıda bahsedilen Dictionary Tagger düğümü ile kişi ve yer bilgileri etiketlenmiş olan kelimeler de görülebilmektedir.

Resim 3. 21: Kelime Çantası Oluşturma İşlemi.

Documents output table - 09 - Bag Of Words Creator (Kelime çantası)

File Edit Hilite Navigation View

Table 'default' - Rows: 108720 Spec - Columns: 3 Properties Flow Variables

Row ID		Term
Row0	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	baybur[LOCATION(N...]
Row1	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	kad[]
Row2	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	hukm[]
Row3	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	abduhah[PERSON(NE)]
Row4	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	zâde[]
Row5	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	kadr[hu[]]
Row6	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	ge[]
Row7	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	erzum[LOCATION(N...]
Row8	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	sancâ[]
Row9	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	kaza[]
Row10	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	tab[]
Row11	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	bulur[]
Row12	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	nam[]
Row13	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	karne[]
Row14	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	vaki[]
Row15	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	feruhşâdbey[]
Row16	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	dâm[]
Row17	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	inaret[]
Row18	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	evkaf[]
Row19	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	mütevell[]
Row20	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	yakub[PERSON(NE)]
Row21	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	kimesne[]
Row22	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	vak[]
Row23	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	mezbûr[]
Row24	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	ferâg[]
Row25	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	asim[]
Row26	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	mürtezkanm[]
Row27	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	vazife[]
Row28	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	ver[]
Row29	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	külü[]
Row30	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	gadr[]
Row31	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	gerif[]
Row32	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	harâb[]
Row33	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	eyle[]
Row34	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	ahâl[]
Row35	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	nefer[]
Row36	ezbür evla megruta merium abduhah sâhib evvel bil tevilyet mezbûr merium yakub fevt oğî hüseyin ref kendu tevch bab istida inayet mahal hasm mezbûr terâfu ger tevilyet mezbûr kangas ehak elyak sabit arz y...	haber[]

3.4.13. DF (Document Frequency, Doküman Frekansı)

Document Frequency düğümü ile Bag of Words Creator düğümü birleştirilerek söz konusu tablo halinde olan veriler içerisindeki terimlerin geçtiği belge sayıları hesaplanmaktadır. Bu düğümün kullanılması ile hüküm verilerindeki kelimelerin her biri için, tüm veri koleksiyonu içerisindeki hükümlerde hangi sıklıkla kullanıldıklarının belirlenmesi mümkün olmuştur. Resim 3.23.'de söz konusu sıklık değerleri görülmektedir.

Resim 3. 22: DF Düğümü.



Resim 3. 23: Hüküm Verileri İçerisindeki Kelimelerin DF Ağırlıklandırma Yöntemine Göre Hesaplanan Frekans Değerleri.

Terms and documents output table - 0:10 - DF (Ağırlıklandırma)

File Edit Hilitte Navigation View

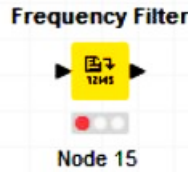
Table "default" - Rows: 108720 Spec - Columns: 4 Properties Flow Variables

Row ID	Term	DF
Row0	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	102
Row1	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1060
Row2	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1191
Row3	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	91
Row4	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	415
Row5	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	66
Row6	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	930
Row7	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	817
Row8	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	344
Row9	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	866
Row10	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	658
Row11	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	9
Row12	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1069
Row13	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	777
Row14	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	589
Row15	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1
Row16	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	5
Row17	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	10
Row18	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	47
Row19	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	91
Row20	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	33
Row21	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	849
Row22	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	147
Row23	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1033
Row24	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	22
Row25	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	3
Row26	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	20
Row27	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	32
Row28	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	543
Row29	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	170
Row30	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	617
Row31	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1040
Row32	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	13
Row33	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	1024
Row34	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	243
Row35	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	136
Row36	uta merkur abdullah sâhib evvel bil tevîyet mezbûr merkur yakub fevt oğî hüseyin ref kendü tevch bab istida inayet mahal hasm mezbûr terâfı ger tevîyet mezbûr kangis ehak elyak sabit arz y...	32

3.4.14. Frequency Filter (Frekans Filtresi)

Frekans değerleri hesaplanmış olan veriler arasında filtreleme yapılmasına olanak sağlayan düğümdür. Bu düğüm ile maksimum ve minimum frekans değerlerini aşan durumlarda terim filtrelenir. Hüküm verilerinin yarısından fazlasında sıklıkla geçmiş olan kelimelerin hükümlerin sınıflandırılması işleminde belirleyici olmayacağı düşünüldükçe söz konusu düğüm vasıtasıyla frekans maksimum değeri 450 olarak belirlenmiştir. Böylece 450 ve üzeri hükümde sıklıkla geçmiş olan kelimeler filtrelenerek algoritmaların gerçekleştirdiği sınıflandırma işleminden daha başarılı sonuçların elde edilmesi hedeflenmiştir.

Resim 3. 24: Frequency Filter Düğümü.

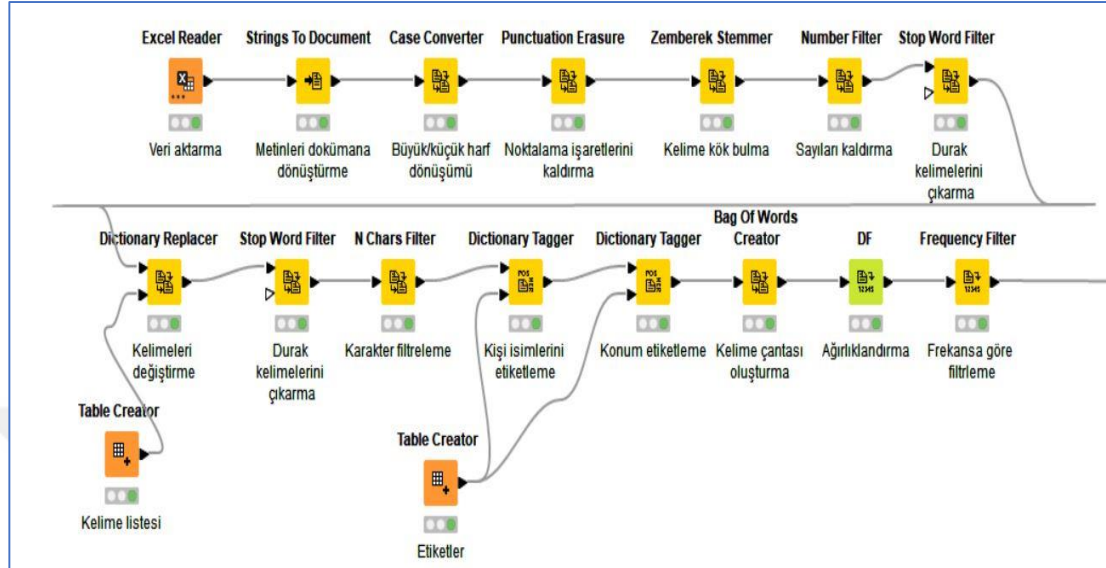


Resim 3. 25: Frekanslarına Göre Filtreleme İşleminin Çıktıları.

Terms and documents output table - 023 - Frequency Filter (Frekans göre)			
File Edit Hilitte Navigation View			
Table "default" - Rows: 80832 Spec - Columns: 4 Properties Flow Variables			
Row ID	Term	DF	
Row0	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	102	
Row3	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	91	
Row4	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	415	
Row5	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	66	
Row7	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	817	
Row8	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	344	
Row11	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	9	
Row15	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	1	
Row16	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	5	
Row17	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	10	
Row18	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	47	
Row19	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	91	
Row20	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	33	
Row22	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	147	
Row24	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	22	
Row25	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	20	
Row26	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	3	
Row27	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	32	
Row29	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	170	
Row32	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	13	
Row34	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	243	
Row35	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	136	
Row36	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	32	
Row37	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	99	
Row38	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	6	
Row39	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	110	
Row40	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	2	
Row41	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	3	
Row42	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	3	
Row43	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	162	
Row44	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	8	
Row45	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	38	
Row46	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	66	
Row47	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	5	
Row48	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	95	
Row50	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	3	
Row51	olumunup deffer hakan teviyet evla megruta merum abdullah sâhib envel teviyet merum yakub fevt oğî hüseyin refi kendü tevch istda inayet hasm terâfû teviyet kangas ehak elyak sabit a...	440	

Yukarıda verilen, hüküm verilerinin sınıflandırılması için oluşturulan modelin ön işlem aşamaları genel olarak aşağıdaki Resim 3.26’da verilmiştir. Sınıflandırma modeli oluşturmak için uygulanan ön işlem aşamaları sayesinde yapılandırılmış veriye dönüştürülen hüküm verileri içerisindeki kişi isimleri ve konum/yer etiketlemesi de otomatik bir şekilde kolaylıkla gerçekleştirilebilecektir.

Resim 3. 26: KNIME Sınıflandırma Modeline Ait Veri Ön İşlem Aşamasının Genel Görünümü.



Veri kümesine uygulanan ön işlem aşamalarından sonra hükümlerde en çok kullanılan 50 kelime ve bu kelimelerin geçiş sıklıkları Tablo 3.2.'de sunulmuştur. Toplamda 5 ve üzeri sıklıkla kullanılan toplam 1860 kelime bulunmuştur. Sıklık değişkeninin 5 olarak seçilmesinin nedeni daha az sıklıkla geçen kelimelerin metin sınıflarının belirlenmesinde ayırt edici bir özelliğe sahip olmayacağı düşüncesidir. Genel olarak hükümlerde hükm, bab, yaz, nam, kadı, bil, şerif gibi kelimeler sıklıkla geçmektedir.

Tablo 3. 2: En Çok Kullanılan 50 Kelime ve Sıklıkları.

Kelime	Frekans	Kelime	Frekans	Kelime	Frekans
hükm	1188	gör	670	zîde	415
bab	1144	vali	660	üzeri	408
yaz	1119	tabi	657	ahz	406
nam	1059	mûceb	630	men	400
kadı	1056	gadr	617	mehmed	396
bil	1053	hal	597	defter	395
şerif	1039	vaki	589	zabt	391
mezbûr	1032	taaddî	583	dahl	388
eyle	1024	amel	575	icra	364

gel	928	ecil	575	akçe	361
kaza	864	ver	543		
kimesne	848	mahal	536		
erzurum	816	sakin	535		
reca	811	kanun	463		
hilaf	795	mutasarrıf	438		
karye	775	şarki	432		
şer	768	icab	431		
sene	753	karahisâr	428		
emr	719	arz	424		
hümayun	686	mezkur	418		

Verilerin görselleştirilmesi elde edilen sonuçların daha kolay anlaşılandırılmasına olanak sağlamaktadır. KNIME programı birçok görselleştirme düğümü sunmaktadır. Verilerin görselleştirilmesinde sıklıkla kullanılan yöntemlerden birisi de kelime bulutudur. Bu şekil ile hükümlerde en sık geçen 1.000 kelime gösterilmiştir. Kelimelerin geçiş sıklığına göre görsel üzerindeki yazı boyutu büyümektedir. Bahsedilen şekil aşağıdaki gibidir;

Resim 3. 27: En Çok Kullanılan 1.000 Kelimenin Kelime Bulutu İle Gösterimi.



Hükümlerde geçen kişi isimleri etiketlenerek algoritmaya tanıtılabilir ve frekans değerleri hesaplanabilir. Bu duruma Dictionary Tagger düğümü anlatılırken değinilmiştir. Aşağıdaki isim listesi incelendiğinde Kirkor, Bogos ve Agop adlı yabancı kişilerin isimleri dikkat çekmektedir. Bu kişilerin isimlerinin az da olsa hükümlerde geçtiği belirlenmiştir. Bu isimlerden yola çıkılarak ilgili hüküm metinleri incelenebilir ve varsa bu kişiler ile alakalı diğer hükümler arasında ilişkiler kurulabilir. Birbirleri ile tutarlı bağlantıları olan (yer ve zaman gibi) hükümler vasıtasıyla ismi geçen kişiler hakkında hükümlerin içeriğine bağlı olarak farklı birçok konuda yeni bilgiler elde edilebilir. Ayrıca hükümlerde geçen isimlere bakıldığında erkek isimlerinin fazla olduğu gözlemlenmektedir. Tablo 3.3. incelendiğinde hükümlerde en çok geçen kişi isimlerinin Mehmed, Ali, Mustafa, Ahmed, Hasan, Murad ve İbrahim olduğu görülürken Malik, Agop, Bayram, Murtaza, Feyzullah isimlerinin ise daha az sıklıkla geçtiği sonucuna varılmaktadır.

Tablo 3. 3: Hükümlerde En Çok Geçen Kişi İsimleri.

İsim	Frekans	İsim	Frekans
Mehmed	396	Musa	35
Ali	321	Yakub	33
Mustafa	314	Salih	30
Ahmed	288	Abdurrahman	28
Hasan	244	Kasım	25
Murad	231	Yahya	22
İbrahim	207	Abbas	16
Osman	206	Fatıma	16
Hüseyin	199	Kirkor	16
Ömer	146	Bogos	14
İsmail	145	Aişe	13
Halil	129	İshak	13
Süleyman	118	Cafer	12
Abdullah	90	Feyzullah	12

Yusuf	81	Murtaza	12
Mahmud	80	Bayram	11
Ebubekir	66	Agop	10
Alişan	40	Malik	10
Bekir	35		

Hükümlerde geçen şehir isimleri incelendiğinde Erzurum, Karahisar, Erzincan, Bayburt, Kemah, Kars ve Gümüşhane şehirlerinin sıklıkla hükümlerde geçtiği elde edilmişken Hamur, Patnos, Kuzucan, Kiğı gibi şehirlerin az sıklıkla hükümlerde geçtiği sonucuna varılmıştır.

Tablo 3. 4: Hükümlerde Geçen Şehir İsimleri ve Sıklıkları.

Şehir	Frekans	Şehir	Frekans
Erzurum	816	Bolaman	20
Karahisar	428	İran	17
Erzincan	155	Trabzon	15
Bayburt	102	Gercanis	11
Kemah	59	Bayezid	10
Kars	57	Haleb	10
Gümüşhane	43	Van	10
Medine	42	Kiğı	4
İskefser	39	Kuzucan	2
İstanbul	39	Patnos	1
Pasin	39	Hamur	1
Kuruçay	37		
Tercan	33		
Habsmana	29		
İspir	28		
Sivas	27		
Hınıs	21		

3.5. Sınıflandırma İşlemi

3.5.1. Hükümlerin Sayısal Olarak İfade Edilmesi

Veri kümesine uygun adımlar doğrultusunda yapılan veri ön işleme çalışmalarının ardından veriler, üzerlerinde analiz yapılabilecek bir yapıya getirilmiş bulunmaktadır. Bu yapı şu şekilde açıklanabilir; algoritmalar kelimeler üzerinde işlem yapmamaktadır. Yapılan işlemler kelimelerin vektör olarak sayısal ifadelere dönüştürülmesiyle gerçekleştirilmektedir. Söz konusu dönüşüm işleminin yapılabilmesi için de iki ana öğeye ihtiyaç vardır. Bunlar;

- **Sözlük:** Hüküm içerisinde geçen kelimelerin tekil olarak gösterildiği tablodur. Ayrıca sözlük içerisinde bulunan her bir kelimeye de benzersiz bir numara verilir. Kelimeler, verilen numaralar ile temsil edilir ve algoritmalar bahsedilen numaralar üzerinden işlemler gerçekleştirir.
- **Hüküm-kelime eşleşmesi (Hüküm-kelime tablosu):** Satırları hüküm verilerinden, sütunları ise sözlükteki bütün kelimelerden oluşan bir tablodur. Bu tabloda kelimelerin hükümler içerisinde geçip geçmediği ve toplamda kaç defa geçtiği gösterilir. “0” değeri kelimenin hükümde geçmediğini, “0” değerinden farklı değerler ise o kelimenin toplamda kaç defa ve hangi hükümde geçtiğini gösterir. Örneğin Tablo 3.6’daki Hüküm 1’de “Erzincan” kelimesi geçmektedir fakat “Meblağ” kelimesi ise geçmemektedir. Ayrıca Hüküm 1’de “Kadı” kelimesi geçmiştir ve başka bir hükümde daha geçtiği için “2” ile gösterilmiştir.

Çalışmada kullanılan algoritmaların hükümler üzerinde yaptığı işlemlerin daha anlaşılabilir olması için, örnek olarak kelime sayısı az olan üç hüküm seçilmiş ve yukarıda bahsedilen sözlük ile hüküm-kelime eşleşmesi tablosu aşağıda gösterilmiştir.

Bu hükümler içerisinde tekil 84 adet kelime bulunmaktadır ve bu kelimeler sözlüğü oluşturmaktadır. Her bir kelimeye ise 1’den başlanarak 84’e kadar numara verilmiştir.

Tablo 3. 5: Sözlük

KELİME	NUMARA
Erzincan	1
Kadı	2
Kaza	3
...	...
Meblağ	83
Tahsil	84

Tablo 3. 6: Hüküm-Kelime Eşleşmesi.

	Erzincan	Kadı	Kaza	...	Meblağ	Tahsil	Konu
Hüküm 1	1	2	2	...	0	0	Sosyal, Ekonomik ve İdarî Konular
Hüküm 2	0	2	2	...	0	0	Sosyal, Ekonomik ve İdarî Konular
Hüküm 3	0	0	0	...	1	1	Arazi Meselesi ve Vergi

Tablo 3.6.'da gösterilen her bir satır, ilgili hükmün vektörünü oluşturmaktadır. Örneğin Hüküm 1 için 1x84'lük vektörün gösterimi aşağıdaki gibidir;

$$\text{Hüküm 1} = [1 \ 2 \ 2 \ \dots \ 0 \ 0]$$

Sınıflandırma algoritmaları her bir hükmü sayısal olarak yukarıdaki gibi gösterilen vektör şeklinde ele almakta ve hüküm vektörleri üzerinden işlem yapmaktadır.

Çalışmada ele alınan 1 numaralı Erzurum Ahkâm Defterinde 1209 adet hüküm ve 11766 kelimeden oluşan bir sözlük ile 1.209 satır ve 11766 sütundan oluşan hüküm-kelime eşleşmesi tablosu bulunmaktadır. Sözlük ve hüküm-kelime eşleşmesi tablosu KNIME programında Document Vector düğümü ile elde edilmiştir.

Resim 3. 28: Document Vector Dügümü Çıktısının Örnek Gösterimi.

Documents output table - 0:15 - Document Vector (Vektör oluşturma)

File Edit Hilite Navigation View

Table "default" - Rows: 1216 Spec - Columns: 11766 Properties Flow Variables

Row ID	Document	D akçeşehr	D naib	D sehven	D kayd	D olmışdı...	D şahin	D tımar
Row6	"aybasdı boş karahisâr şarki hass voyvoda aybasdı bağçecik ahalı taha b...	0	0	0	0	0	0	0
Row7	"aybasdı boş karahisâr şarki hass voyvoda aybasdı bağçecik fatı hatun t...	0	0	0	0	0	0	0
Row8	"aybasdı boş karahisâr şarki hass voyvoda aybasdı bağçecik sükkân tah...	0	0	0	0	0	0	0
Row9	"aybasdı boş karahisâr şarki hass voyvoda aybasdı surşin sükkân alı kız k...	0	0	0	0	0	0	0
Row10	"aybasdı boş karahisâr şarki mukataa voyvoda ömer şer'an virecek deyn...	0	0	0	0	0	0	0
Row11	"aybasdı boş mîlas naib karahisâr şarki voyvoda mîlas naib mevlana taha ...	0	107	0	0	0	0	241
Row12	"aybasdı hasnucak sükkân immeden mevlana ahmed zide ilmuhi cami şer'i...	0	0	0	0	0	0	0
Row13	"aybasdı karahisâr şarki hass voyvoda aybasdı abahuk ahmed şer'an ne...	0	0	0	0	0	0	0
Row14	"aybasdı karahisâr şarki voyvoda hasan mehmed südde sa'âdetime arz a...	0	0	0	0	0	0	0
Row15	"bayburt abdullah zide kadruhi erzurum sancağ bayburt pulur feruhsâd...	0	0	0	157	0	0	0
Row16	"bayburt bayburt bul müteveffa ferahşad beğ cami imaret evkaf evlad m...	0	0	0	0	0	0	0
Row17	"bayburt bayburt uzun sükkân kahraman südde saadet arz eza lazm av...	0	0	0	0	0	0	0
Row18	"bayburt baybürd'da şaban ağa mescid bil'fîl berat müteveli imam abbas...	0	0	0	0	0	0	0
Row19	"bayburt baybürd'da ahi emir ahmed kuddise sirruhul-aziz evkaf defter m...	0	0	0	0	0	0	0

Veri kümesinde bulunan hükümlerin konularına göre sınıflandırılabilmesi için çeşitli sınıflandırma yöntemleri kullanılmıştır. Farklı kelime ağırlıklandırma yöntemlerine göre sınıflandırma algoritmalarının başarı oranları değerlendirilmiştir. Sınıflandırma işlemine 1.209 adet hüküm dâhil edilmiştir. Bu hükümlerin %70'lik kısmı eğitim verisi ve geri kalan %30'luk kısmı ise test verisi olarak kullanılmıştır. 847 hüküm eğitim verisi olarak kullanılırken 362 hüküm ise test verisi olarak kullanılmıştır. Söz konusu bölme işlemi KNIME programının Partitioning (Bölme) düğümü ile gerçekleştirilmiştir. KNIME düğümleri kullanılarak veri kümesi üzerinde uygulanan sınıflandırma algoritmaları aşağıdaki gibidir;

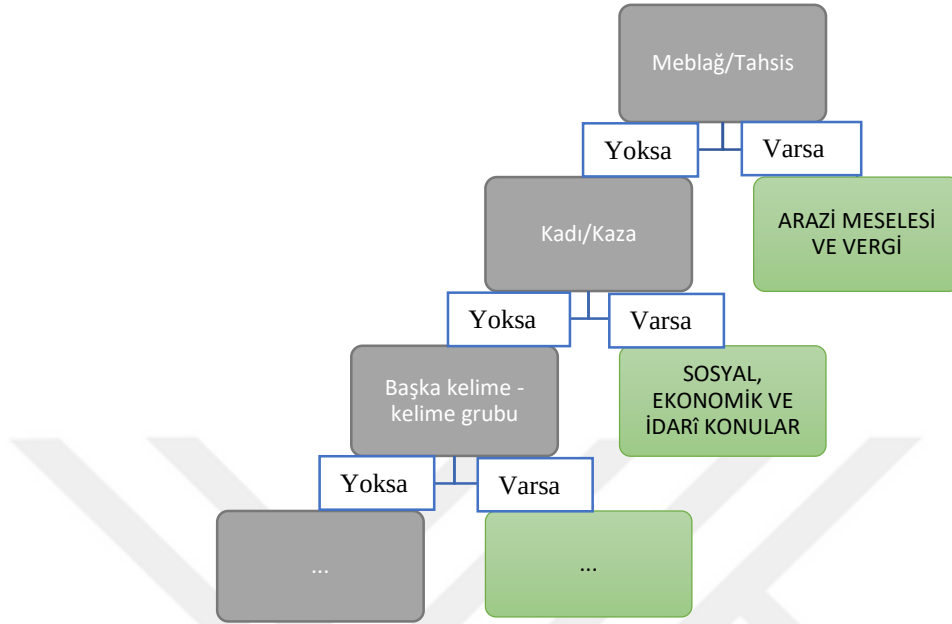
3.5.2. Decision Tree Learner (Karar Ağacı Öğrenicisi)

Karar ağacı, kararların ve muhtemel sonuçların gösterilmesi için ağaç benzeri bir grafik ya da model kullanan (Uysal ve Güyer, 2014: 34) kolay yorumlanabilir ve anlaşılabilir olan, ayrıca sınıflandırma ile tahmin için kullanılan bir veri madenciliği yaklaşımıdır (Chien ve Chen, 2008: 282). Karar ağaçlarında sınıflandırma işlemi öğrenme ve tahmin etme (sınıflandırma) olmak üzere iki aşamada gerçekleştirilmektedir (Özekeş ve Çamurcu, 2002; Çalış ve diğerleri, 2014; aktaran Çetinkaya ve Horasan, 2021: 141).

Bir karar ağacı, kök düğüm adı verilen tek bir değişkenle başlayan, ağaç benzeri bir yapıda organize edilmiş hiyerarşik bir ilişkiler grubudur. Bu kök düğüm, kök düğümün kategorik ise ayrı sınıflarını, sürekliyse düğüm ölçeği boyunca belirli aralıkları temsil eden iki ya da ikiden fazla dala ayrılır. Her bölmede, bölünen değişkenin sınıfları veya aralığı açısından yanıtı olan bir soru sorulur. “Bu vaka bir erkek mi yoksa bir kadın mı?” sorusu buna bir örnek olabilir. Buna benzer sorular ikili karar ağacı oluşturmak için kullanılabilir. Karar ağaçları aynı zamanda ikiden fazla çoklu bölmelerle de oluşturulabilir. Her bölmede sorulan sorular, bölünmelerde ortaya çıkan durumların ne kadar tekil olması gerektiğini yansıtan bazı belirsizlik ölçüleri açısından tanımlanır. Her dal, diğer değişkenlerin sınıfları veya aralıkları kullanılarak daha da alt dallara bölünür. Her bölmede, bölünen düğüme ana (üst) düğüm ve bölünmüş düğümlere de çocuk (alt) düğümler denir. Bu süreç, belirli bir durdurma kuralı sağlanana kadar devam eder. Bu işleme özyinelemeli bölümlenme denir (Nisbet ve diğerleri, 2009: 241).

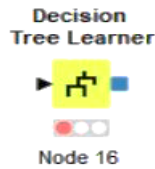
Bu çalışmada karar ağacı algoritması şu şekilde işlem yapmaktadır; konuları bilinen hüküm verilerinin (eğitim verilerinin) içerisinde geçen kelimeler göz önünde bulundurularak konuları en iyi şekilde ifade edebilecek kelimeler belirlenir. Ardından bir kök düğüm (kelime/kelime grubu) seçildikten sonra bu kelime/kelime grubuna hükümde geçip geçmediğine dair soru sorulur. Sorulan sorunun neticesinde ya hükmün konusu belirlenir ya da hükmün konusu belirleninceye kadar başka kelime/kelime gruplarına aynı soru sorulmaya devam eder.

Şekil 3. 5: Karar Ağacı Örneği.



Karar ağacının oluşturulması KNIME programında Decision Tree Learner düğümü ile gerçekleştirilir. Bu aşamaya öğrenme aşaması denir ve öğrenme işlemi için eğitim veri kümesi kullanılır. Bu düğüm KNIME programında aşağıdaki gibidir;

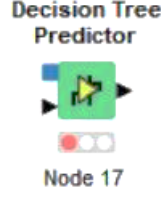
Resim 3. 29: Decision Tree Learner Düğümü.



3.5.3. Decision Tree Predictor (Karar Ağacı Tahmincisi)

Sınıflandırma tahmini aşamasında test veri kümesinde, yukarıda öğrenilen karar ağacı kullanılarak hükmün konu tahmini gerçekleştirilir. Konusu belirlenmek istenen hüküm verisinin içerisinde geçen kelimeler, yukarıda eğitilmiş karar ağacının kök düğümünden başlayarak kontrol edilir. Hüküm içerisinde geçen kelimeler varsa söz konusu dal üzerinden konusu belirlenir, bahsi geçen kelimeler yoksa hükmün konusu belirlenene kadar karar ağacı üzerinde ilerleme işlemine devam edilir. KNIME programında ilgili düğümün gösterimi aşağıdaki gibidir;

Resim 3. 30: Decision Tree Predictor D ğ m .



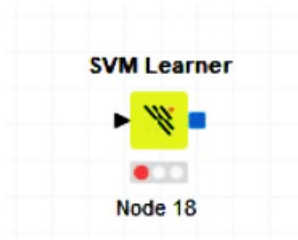
3.5.4. Support Vector Machine (SVM) Learner (Destek Vekt r Makinesi  ğrenicisi)

SVM algoritması literat rde Destek Vekt r Makinesi olarak da bilinen, b y k miktardaki verilerin sınıflandırılması sorununun c z lmesine yardımcı olan klasik bir makine  ğrenmesi y ntemidir (Suthaharan, 2016: 207) ve bu y ntem, eğitim ařamasında verilerin hangi sınıfa ait olduklarının  nceden belirli olduėu denetimli  ğrenme y ntemi kapsamına girmektedir (K c ksille ve Ateř, 2013, s. 2).

Bu algoritma iki ya da ikiden fazla (c oklu) sınıflı sınıflandırma problemlerinde kullanılmaktadır. Buradaki amaç, eğitim veri k mesindeki verileri iki ya da ikiden fazla sınıfa ayıracak en iyi genelleřtirici sınıflandırıcıyı bulmaktır. Verileri ayıracabilecek birc ok olası sınıflandırıcı olabilir ancak her sınıfın birbirine en yakın veri noktası arasındaki mesafeyi en uzak mesafeye c ıkaran yalnızca bir sınıflandırıcı bulunmaktadır. Bu sınıflandırıcıya ise optimal ayırma d zlemi denir (Gunn, 1998: 5).

Bu c alıřmada SVM algoritmasıyla h k m konularını en iyi ve genel řekilde temsil edecek kelimeleri birbirinden ayıracak sınıflandırıcı/optimal ayırma d zlemi  ğrenilmiřtir. S z konusu algoritmanın KNIME'da bulunan d ğ m n n g sterimi ařaėıdaki gibidir;

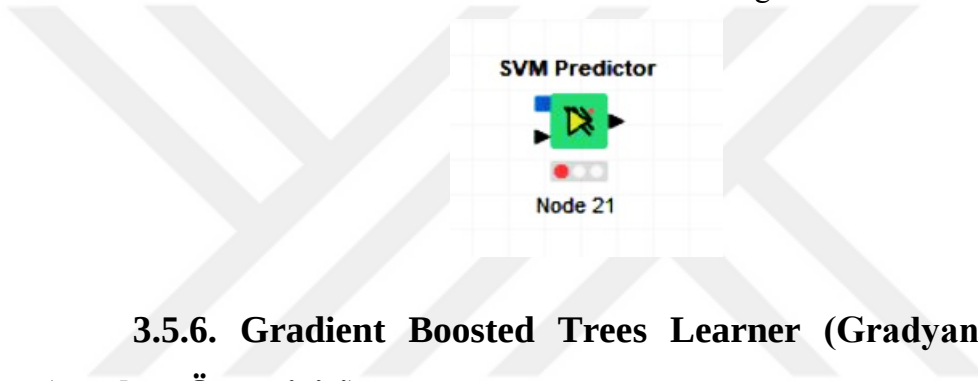
Resim 3. 31: SVM Learner D ğ m .



3.5.5. Support Vector Machine (SVM) Predictor (Destek Vektör Makinesi Tahmincisi)

SVM Learner düğümü ile elde edilmiş sınıflandırıcı/optimal ayırma düzlemi tarafından, sınıflandırılması istenen hükmün kelimelerinin düzlem üzerindeki yeri belirlenir. Ardından düzlem üzerindeki yeri belirlenen kelimelerin etrafında bulunan kelimelere olan uzaklığına bakılır. En yakın kelimenin konusu yeni hükmün konusu olarak belirlenir. KNIME programında ilgili düğümün gösterimi şu şekildedir;

Resim 3. 32: SVM Predictor Düğümü.

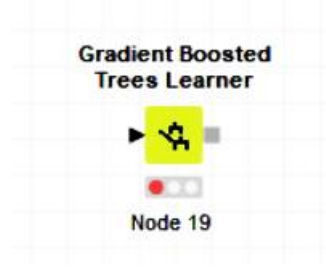


3.5.6. Gradient Boosted Trees Learner (Gradyan Artırma Ağaçları Öğrenicisi)

Gradyan Artırma Algoritması olarak da bilinen bu düğüm ile Decision Tree benzeri birden fazla ağaç oluşturulmaktadır. Daha sonra tahminlerde yapılan hatalar göz önüne alınarak ağaç oluşturma işlemi tekrar etmektedir. Bahsedilen süreç modelin daha fazla geliştirilemeyeceği düşünülene kadar devam etmektedir (Keleş, Keleş ve Keleş, 2020: 516).

Bu çalışmada Gradient Boosted algoritması, hükmün konusunun belirlenmesinde yardımcı olacak kelimeleri seçer. Ardından seçilen kelimelerin hükmün konusunu belirlemesindeki başarısı hesaplanır. Burada amaç, başarı oranını arttırmak dolayısıyla hata oranını azaltmaktır. Bu işlem, hükmün konusunu en iyi temsil edecek kelimeleri içeren ağaç yapısını bulana kadar devam eder. Bu düğüm KNIME programında aşağıdaki gibidir;

Resim 3. 33: Gradient Boosted Trees Learner Düzümü.

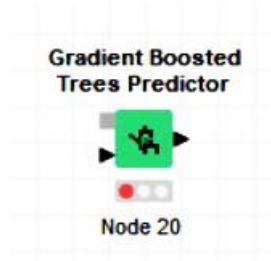


3.5.7. Gradient Boosted Trees Predictor (Gradyan Artırma Ağaçları Tahmincisi)

Gradient Boosted Learner düğümüne verilen test veri kümesinin algoritma ile tahmin edilerek çıktıların sunulduğu düğümdür.

Bu çalışmada konusu belirlenecek olan hükmün içerisinde geçen kelimeler kullanılarak Learner düğümü tarafından eğitilen ağaç üzerinde ilerlenir. İlerlemenin sonucunda bulunan konu, hükmün konusu olarak tahmin edilir. KNIME programında düğüm aşağıdaki gibidir;

Resim 3. 34: Gradient Boosted Trees Predictor Düzümü.



3.5.8. Random Forest Learner (Rastgele Orman Öğrenicisi)

Rastgele orman algoritması, aynı yapıya sahip rastgele oluşturulmuş birden çok karar ağacı kullanarak sınıflandırma yapan bir algoritmadır. Sınıflandırılmak istenen yeni bir veri geldiğinde bu veri her bir ağaca gönderilir. Her ağaç kendi içerisinde sınıflandırma yapar ve verinin sınıfını çıktı olarak verir. Ardından, karar ağaçlarının çoğunluğu tarafından belirlenen sınıf, sınıflandırılmak istenen verinin nihai sınıfı olarak atanır (Breiman, 2001: 6).

Bu çalışmada Random Forest Learner düğümü ile hükümlerin içerisinde geçen kelimelerinin var olup olmama durumlarına göre oluşturulan karar ağaçlarının bir araya getirilerek çıktıların algoritma tarafından öğrenilmesi işlemi gerçekleştirilmiştir. Bu işlem KNIME programında Random Forest Learner düğümü ile gerçekleştirilir.

Resim 3. 35: Random Forest Learner Düğümü.



3.5.9. Random Forest Predictor (Rastgele Orman Tahmincisi)

Random Forest Learner düğümünde oluşturulan birden çok karar ağacının çıktıların öğrenilmesinden sonra, test verisi içerisindeki sınıfları tahmin edilecek verilerin girdi olarak işleme alınmasını ve sınıflarının tahmin edilmesini sağlayan düğümdür. Yukarıda bahsedildiği üzere oluşturulan karar ağaçlarından elde edilen sonuçlara göre, test veri kümesi içerisindeki verilerin sınıfları tahmin edilir.

Bu çalışmada, Random Forest Learner düğümü ile, hükümlerin içerisinde geçen kelimelerinin var olup olmama durumlarına göre oluşturulan rastgele karar ağaçlarından elde edilen konular içerisinde çoğunluğun vermiş olduğu konu yeni hükmün konusu olarak tahmin edilir.

Resim 3. 36: Random Forest Predictor Düğümü.



3.5.10. K Nearest Neighbor (K En Yakın Komşu)

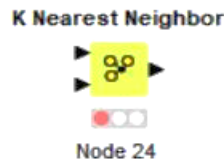
K en yakın komşu algoritması (KNN – K Nearest Neighbor), sınıflandırma algoritmaları içinde sıklıkla kullanılmaktadır. K en yakın komşu algoritması, sınıfı tahmin edilecek verinin veri kümesindeki diğer verilere en yakın olduğu noktayı arar. Bu algoritmanın iki ana parametresi bulunmaktadır. Bunlar;

- **K değeri:** Kaç tane en yakın komşunun dikkate alınacağını belirtmek için kullanılır. Genellikle tek sayılar seçilmektedir. Bunun sebebi, yakın olan komşu sayılarının eşit olma durumunun önlenmesidir.
- **Uzaklık ölçüsü:** Tahmin edilecek verinin, veri kümesindeki diğer verilere olan uzaklığını hesaplamak için kullanılan ölçüdür. Bu uzaklık genellikle Öklid uzaklığı kullanılarak hesaplanır. Alternatif olarak Manhattan uzaklığı da kullanılabilir (Nisbet ve diğerleri, 2009: 239).

Bu çalışmada K en yakın komşu algoritması şu şekilde işlem yapmaktadır; konusu belirlenmek istenen bir hüküm verisinin, eğitim için kullanılan verilere olan uzaklığı hüküm vektörleri kullanılarak hesaplanır. Uzaklıklar küçükten büyüğe doğru sıralanır ve belirlenen k değeri kadar yakın komşu hüküm verilerin konularına bakılır. K adet komşu içerisinde sayıca fazla olan verinin konusu, yeni hükme konu olarak atanır. Çalışmada k değişkeni, sıklıkla tercih edilen 3, 5 ve 7 olarak belirlenmiş ve her bir k değerine göre sınıflandırma işlemi yapılmıştır.

Yukarıda bahsedilen işlemler KNIME programında K Nearest Neighbor düğümü kullanılarak gerçekleştirilmiştir.

Resim 3. 37: K Nearest Neighbor Düğümü.



3.5.11. Scorer (Performans Hesaplayıcı)

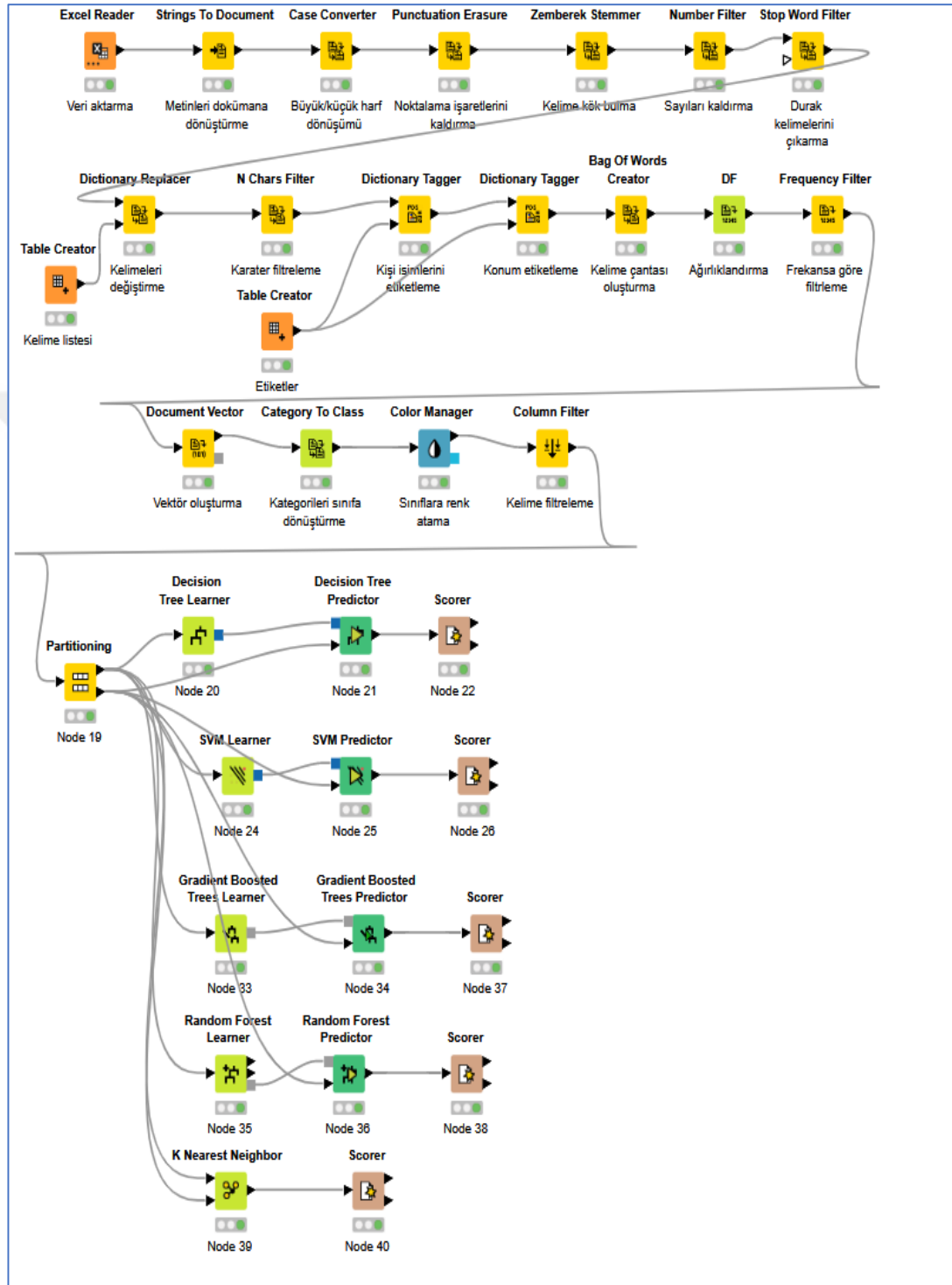
Bu düğüm, veri kümesine uygulanan algoritmalarından elde edilen sonuçların gösterilmesine olanak sağlamaktadır. Algoritmaların çeşitli ağırlıklandırma yöntemlerine göre hüküm verilerinin konularını tahmin etmesi ve sınıflandırması işlemlerinin başarı oranları bu düğüm vasıtasıyla öğrenilir. KNIME programında düğümün gösterimi şu şekildedir;

Resim 3. 38: Scorer Düğümü.



Aşağıdaki şekilde KNIME programı ile hüküm verilerinin sınıflandırılması için oluşturulmuş olan sınıflandırma modelinin tüm işlem düğümleri bulunmaktadır. Söz konusu model ile gerçekleştirilecek yeni çalışmalara kaynaklık edilebilmesi ve muazzam bir veri kaynağı teşkil eden Osmanlı dönemine ait arşiv belgeleri üzerinde yapılabilecek olan çalışmalara dikkat çekilmesi hedeflenmektedir.

Resim 3. 39: KNIME İle Oluşturulan Sınıflandırma Modeli.



3.6. Sınıflandırma Sonuçları

Yukarıda açıklamaları yapılan sınıflandırma algoritmaları, KNIME programıyla hüküm metinlerinden elde edilen veri kümesi üzerinde uygulanmıştır. Bu uygulamalar sonucunda farklı ağırlıklandırma yöntemlerine göre başarı oranları elde edilmiştir. Sınıflandırma işleminde başarı, verilerin doğru başlıklar altında bir araya getirilmesi şeklinde açıklanabilir. Hükümlerin sınıflandırılmasında elde edilen başarı oranları doğruluk (accuracy) oranı ve f-ölçüsü (f-measure) değerlendirme ölçütlerine göre belirlenmiştir. Daha sonra en başarılı sonuçlara ulaşılan algoritma ile gerçekleştirilen sınıflandırma işleminde doğru ya da yanlış sınıflandırılan hüküm sayılarının dağılımı karışıklık matrisi (confusion matrix) ile gösterilmiştir. Karışıklık matrisi ile verilerin gösterimi şu şekildedir;

Tablo 3. 7: Karışıklık Matrisi.

		Tahmin Edilen Sınıf	
		Pozitif	Negatif
Gerçek Sınıf	Pozitif	TP (Doğru Pozitif)	FP (Yanlış Pozitif)
	Negatif	FN (Yanlış Negatif)	TN (Doğru Negatif)

- **TP (True Positive);** Pozitif sınıfında bulunan ve doğru tahmin edilen verilerin adedini,
- **FP (False Positive);** Pozitif sınıfında bulunan fakat yanlış tahmin edilen verilerin adedini,
- **TN (True Negative);** Negatif sınıfında bulunan ve doğru tahmin edilen verilerin adedini,
- **FN (False Negative);** Negatif sınıfında bulunan fakat yanlış sınıflandırılan verilerin adedini ifade etmektedir.

Verilerinin sınıflandırılmasında elde edilen başarı oranları hesaplanırken kullanılan değerlendirme ölçütlerinden doğruluk oranı, doğru tahmin edilerek sınıflandırılan verilerin sayısının toplam veri sayısına oranı ile elde edilmektedir (Amanet, 2017: 50).

Denklem 3. 1: Doğruluk Oranı.

$$\text{Doğruluk Oranı} = \frac{TP + TN}{TP + FP + TN + FN}$$

Kaynak: Amanet, 2017: 50.

F-ölçütü ise kesinlik ve duyarlılık ölçütlerinin harmonik ortalaması ile elde edilmektedir ve her iki ölçütün bir arada kullanılmasını sağlamaktadır (Kartal, 2015: 34).

Denklem 3. 2: F-ölçütü.

$$F - \text{Ölçütü} = \frac{2 * \text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}}$$

Kaynak: Kartal, 2015: 34.

Yukarıdaki denklemde geçen kesinlik (precision), pozitif sınıfta bir araya getirilmiş olan verilerden kaç tanesinin gerçekten pozitif sınıfa ait olduğuna dair hesaplama yapılmasına olanak sağlayan ölçüttür (Gündüz, 2013: 23).

Denklem 3. 3: Kesinlik.

$$\text{Kesinlik} = \frac{TP}{TP + FP}$$

Kaynak: Ekrem, Salman, Aksoy ve İnan, 2020: 246.

Duyarlılık ölçütü ise doğru olarak sınıflandırılmış pozitif veri sayılarının tahmin edilen toplam pozitif veri sayılarına oranı ile elde edilmektedir (Kartal, 2015: 33).

Denklem 3. 4: Duyarlılık.

$$Duyarlılık = \frac{TP}{TP + FN}$$

Kaynak: GümüŖsu, 2019: 54.

Hüküm verileri üzerinde farklı ağırlıklandırma yöntemlerine göre uygulanan algoritmalar sonucunda sınıflandırma işlemleri gerçekleştirilmiştir. Bu işlemler sonucunda oluşturulan sınıflandırma modellerinin başarı oranları aşağıdaki tablolarda Ŗu Ŗekilde verilmiştir;

Tablo 3. 8: DF Ağırlıklandırma Yöntemiyle Sınıflandırma Algoritmalarının Başarısının İncelenmesi.

Sınıflandırma algoritmaları	F-ölçüsü	Doğruluk oranı
Karar ağaçları	0,715	0,724
SVM (Destek Vektör Makineleri)	0,724	0,735
Rastgele Orman (Random Forest)	0,805	0,812
K-En yakın komŖu (k=3)	0,523	0,552
K-En yakın komŖu (k=5)	0,512	0,541
K-En yakın komŖu (k=7)	0,521	0,555
Gradient Boosted ağaçları	0,799	0,812

Tablo 3.8.'de kelime ağırlıklandırma yöntemi olarak DF ağırlıklandırma yönteminin kullanıldığı modelde en yüksek doğruluk oranı ve F-ölçüsü değeri Rastgele Orman (Random Forest) yönteminde elde edilmiştir. En düşük başarı oranı ise 0,541 ile K-En yakın komŖu (k=5) algoritmasında elde edilmiştir.

Tablo 3. 9: Binary Ağırlıklandırma Yöntemiyle Sınıflandırma Algoritmalarının Başarısının İncelenmesi.

Sınıflandırma algoritmaları	F-ölçüsü	Doğruluk oranı
Karar ağaçları	0,711	0,718
SVM (Destek Vektör Makineleri)	0,748	0,757
Rastgele Orman (Random Forest)	0,760	0,765
K-En yakın komşu (k=3)	0,579	0,605
K-En yakın komşu (k=5)	0,575	0,608
K-En yakın komşu (k=7)	0,614	0,644
Gradient Boosted ağaçları	0,765	0,773

Tablo 3.9.'da kelime ağırlıklandırma yöntemi olarak binary ağırlıklandırma yönteminin kullanıldığı modelde en yüksek doğruluk oranı Gradient Boosted ağaçları algoritmasından elde edilmiştir. Bu oran 0,773'tür. Ayrıca en yüksek F-ölçüsü değeri de Gradient Boosted ağaçları algoritmasından elde edilmiştir. En düşük başarı oranı ise 0,605 ile K-En yakın komşu (k=3) algoritmasında elde edilmiştir.

Tablo 3. 10: TF Ağırlıklandırma Yönteminde Sınıflandırma Algoritmalarının Başarısının İncelenmesi.

Sınıflandırma algoritmaları	F-ölçüsü	Doğruluk oranı
Karar ağaçları	0,736	0,749
SVM (Destek Vektör Makineleri)	0,726	0,751
Rastgele Orman (Random Forest)	0,793	0,804
K-En yakın komşu (k=3)	0,627	0,660
K-En yakın komşu (k=5)	0,646	0,680
K-En yakın komşu (k=7)	0,662	0,696
Gradient Boosted ağaçları	0,767	0,782

Tablo 3.10.'da kelime ağırlıklandırma yöntemi olarak TF ağırlıklandırma yönteminin kullanıldığı modelde en yüksek doğruluk oranı Rastgele Orman (Random

Forest) algoritmasından elde edilmiştir ve bu oran 0,804'tür. En yüksek F-ölçüsü değeri de 0,793 oran ile yine Rastgele Orman (Random Forest) algoritmasından elde edilmiştir. En düşük başarı oranı ise 0,660 ile K-en yakın komşu (k=3) algoritmasında elde edilmiştir.

Tablo 3. 11: IDF Ağırlıklandırma Yönteminde Sınıflandırma Algoritmalarının Başarısının İncelenmesi.

Sınıflandırma algoritmaları	F-ölçüsü	Doğruluk oranı
Karar ağaçları	0,710	0,718
SVM (Destek Vektör Makineleri)	0,744	0,754
Rastgele Orman (Random Forest)	0,789	0,793
K-En yakın komşu (k=3)	0,575	0,627
K-En yakın komşu (k=5)	0,541	0,624
K-En yakın komşu (k=7)	0,542	0,624
Gradient Boosted ağaçları	0,809	0,812

Tablo 3.11.'de kelime ağırlıklandırma yöntemi olarak IDF ağırlıklandırma yönteminin kullanıldığı modelde en yüksek doğruluk oranı ve F-ölçüsü değeri Gradient Boosted ağaçları algoritmasında elde edilmiştir. Bu değerler doğruluk oranı için 0,812 ve F-ölçüsü için 0,809'dur. En düşük başarı oranı ise 0,624 ile K-en yakın komşu (k=5 ve k=7) algoritmalarından elde edilmiştir.

Tablo 3. 12: ICF Ağırlıklandırma Yönteminde Sınıflandırma Algoritmalarının Başarısının İncelenmesi.

Sınıflandırma algoritmaları	F-ölçüsü	Doğruluk oranı
Karar ağaçları	0,753	0,751
SVM (Destek Vektör Makineleri)	0,795	0,801
Rastgele Orman (Random Forest)	0,760	0,765
K-En yakın komşu (k=3)	0,537	0,577
K-En yakın komşu (k=5)	0,513	0,572

K-En yakın komşu (k=7)	0,505	0,572
Gradient Boosted ağaçları	0,783	0,787

Tablo 3.12.'de kelime ağırlıklandırma yöntemi olarak ICF ağırlıklandırma yönteminin kullanıldığı modelden en yüksek doğruluk oranı ve F-ölçüsü değeri SVM algoritmasında elde edilmiştir. Bu başarı değerleri doğruluk oranı için 0,801 ve F-ölçüsü değeri için de 0,795'tir. En düşük başarı oranı ise 0,572 ile K-En yakın komşu (k=5 ve k=7) algoritmalarından elde edilmiştir.

Genel olarak yapılan sınıflandırma çalışmasında elde edilen en yüksek başarılı sonuç, IDF ağırlıklandırma yöntemine göre veri kümesine uygulanan Gradient Boosted ağaçları algoritmasında elde edilmiştir. Bu sonuca dayanarak hükümlerin konularına göre sınıflandırılmasında Gradient Boosted ağaçları algoritması ve IDF ağırlıklandırma yöntemi kullanılabilir.

Tablo 3. 13: Gradient Boosted Ağaçları Algoritması Sonucunda Elde Edilen Karışıklık Matrisi (n=362).

		Tahmin Edilen Sınıf			
		Asayiş	Sosyal, Ekonomik ve İdarî Konular	Arazi Meselesi ve Vergi	Başarı Oranı %
Doğru Sınıf	Asayiş	60	14	8	73,2
	Sosyal, Ekonomik ve İdarî Konular	8	108	14	83,1
	Arazi Meselesi ve Vergi	4	17	129	86

Tablo 3.13.'te test verisinde kullanılan 362 hükmün kaç tanesinin hangi sınıflara atandığını gösteren karışıklık matrisi verilmiştir. Gradient Boosted ağaçları algoritmasının veri kümesine uygulanması ile elde edilen sonuçların sunulduğu karışıklık matrisi incelendiğinde asayiş ile ilgili toplam 82 adet hüküm verisinin

60'ının doğru bir şekilde sınıflandırıldığı gözlemlenmektedir. 22 adet hüküm ise farklı konu başlıkları altında yanlış sınıflandırılmıştır. Bu durum, yaklaşık olarak %73,2 oranında başarılı bir sınıflandırma yapıldığını göstermektedir. Ayrıca sosyal, ekonomik ve idarî konular ile ilgili hükümlerin sınıflandırma sonuçları incelendiğinde toplamda 130 adet olan hükmün 108'inin doğru konu başlığı altında sınıflandırıldığı gözlemlenmiştir. 22 adet hüküm ise yanlış sınıflandırılmıştır. Böylece doğru sınıflandırılan hükümlerin oranı yaklaşık olarak %83,1'dir. Ayrıca arazi meselesi ve vergi ile ilgili hükümlerin sınıflandırma sonuçları incelendiğinde toplamda 150 adet olan hükmün 129'unun doğru konu başlığı altında sınıflandırıldığı gözlemlenmiştir. 21 adet hüküm ise yanlış sınıflandırılmıştır. Böylece doğru sınıflandırılan hükümlerin oranı yaklaşık olarak %86'dır. Elde edilen bu oranlar ile geliştirilen sınıflandırma modelinin başarılı olduğu sonucuna ulaşılabilir.

SONUÇ VE ÖNERİLER

Gün içerisinde yapılan eylemler sonucunda farklı alan ve konularda birçok veri üretilmektedir. Oluşturulan bir belge, gönderilen elektronik bir posta ya da bir marketten belirli bir ay içerisinde satın alınan ürünlerin bilgileri ve adedi birer veri olarak kabul edilebilmektedir. Kolaylıkla ve herkes tarafından her alanda üretilebilen veriler, yapıları gereği farklılık gösterebilmektedir. Yapılandırılmış, yarı yapılandırılmış ya da yapılandırılmamış şekilde bulunan verilerden, oluşmalarına etki eden durumlar başta olmak üzere birçok konu hakkında bilgi elde edilebilmektedir. Fakat bu bilgiler her zaman basit bir şekilde anlaşılabilir olamamaktadır. Elde edilmek istenilen bilgiler veri içerisinde örtük bir şekilde bulunabilmektedir. Söz konusu bilgilere, veri türüne ve yapısına göre uygulanan deney ve gözlemler sonucunda verinin anlamlandırılması ile ulaşılabilir olmaktadır. Verinin anlamlandırılabilmesi için çeşitli araç ve yöntemler geliştirilmiştir. Veri madenciliği ve metin madenciliği bu anlamda en sık kullanılan yöntemlerdendir. Bu yöntemler ile veriler çeşitli ön işlemlerden geçirilerek uygun algoritmalar ile analiz edilmektedir. Ortaya çıkan sonuçların yorumlanması ile bilgi elde edilebilmektedir.

Geçmişin anlaşılabilmesi, mevcut durumun değerlendirilebilmesi, geleceğe yönelik tahminlerin yapılması ve kararların alınması, verilerden elde edilebilecek bilgiler ile mümkün olabilmektedir. Kurum ve kuruluşlar, şirketler ve kişiler elde edilen bu bilgiler doğrultusunda faaliyetlerini belirlemekte ve gelecek planlarını belirlemektedir. Sunulan hizmetlerin kalitesinin artırılması, iş ve işlemlerin düzenlenerek geliştirilmesi, yeni imkân ve fırsatların değerlendirilmesi durumları için verilerden elde edilen bilgiler belirleyici olabilmektedir. Ayrıca nedeni merak edilen durumların açıklığa kavuşturulması, yeni bilgilerin üretiminin kolaylaştırılması ve başvurulması halinde kaynak niteliği taşıması durumları bakımından verilerden bilginin çıkarılması oldukça önemli bir işlem olarak kabul edilmektedir.

Verilerin analiz edilerek amaca yönelik bilgilere ulaşılmasında verinin kaynağı ve hacmi oldukça belirleyicidir. Kaynağın güvenilir olması ve içerdiği verilerin doğruluğu ayrıca veri miktarının yapılan analizler sonucunda anlamlı sonuçlar

alınabilmesi için yeterli düzeyde olması, elde edilecek bilgiye doğrudan etki etmektedir. Güvenilirliğinden şüphe edilen ya da içerisinde tutarlı olmayan değerler barındıran veri kaynağından bilgi çıkarımının yapılması bir anlam ifade etmeyecektir. Ayrıca kısıtlı veriler üzerinde yapılan işlemler sonucunda çıkarılan bilgilerin geçerliliği sorgulanabilecektir. Bu sebeple seçilecek verilerin alındığı kaynak ve üzerinde işlem yapılacak verinin miktarı bilgi çıkarımının gerçekleştirilmesinde önemli birer unsur olarak kabul edilmektedir. Söz konusu unsurlar göz önünde bulundurularak yapılan verilerden bilgi çıkarma işlemi her bir disiplin için uygulanabilmektedir.

Veriler üzerinde yapılacak analiz işlemleri için gerekli olan veriler birçok ortamdan elde edilebilmektedir. Literatür incelendiğinde genellikle verilerin veri madenciliği ya da metin madenciliği ile analiz edilerek incelendiği çalışmalarda kaynak olarak sosyal medya platformlarından (Amanet, 2017: 84), e-ticaret ve haber sitelerinden (Erten, 2015: 34), basılı ya da elektronik formatta bulunan kitap ile dergilerden elde edilen verilerin (Toplu, 2019: 2) kullanıldığı görülmektedir. Kurum ve kuruluşların arşivleri içerisinde bulunan ve faaliyetler sonucu meydana gelen ya da bir araya getirilen materyaller de veri kaynağı olarak veri madenciliğinde ve metin madenciliğinde kullanılmaktadır. Kılıç, Demirtaş, Candan, Akbayrak ve Kutlu'nun yapmış olduğu çalışmada (2019: 311) yapay zekâ teknolojilerinden biri olan metin madenciliğinin arşivlerde yaşanan sorunlara çözümler getirdiğini belirtmeleri söz konusu durumu kanıtlar niteliktedir. Büyük miktarda verilerin elde edilebileceği ve kanıt niteliği taşıyan, farklı birçok disiplin tarafından incelenebilecek sonuçların elde edilebileceği önemli bir diğer veri kaynağının ise Osmanlı dönemi arşiv evrakı olduğu söylenebilir. Osmanlı İmparatorluğu'ndan günümüze kadar ulaşan büyük miktardaki arşiv malzemesi, İmparatorluğun köklü bir arşivleme geleneğine sahip olduğuna işaret ettiği sonucuna varılabilmektedir. Osmanlı dönemi arşiv evrakı sahip olduğu veriler bakımından ele alındığında, söz konusu veriler üzerinde yapılacak analiz çalışmaları sonucunda ortaya çıkarılacak olan bilgiler Coğrafya, Etnoloji, Kronoloji, Filoloji, Felsefe, Hukuk, Tarih, İktisat, Edebiyat, Sosyoloji ve Tıp bilimlerine kaynak olabilmektedir.

Osmanlı arşivinde bulunan veri kaynakları genellikle metin formatındadır ve bu veriler üzerinde gerçekleştirilecek analizlerin manuel olarak yapılması çeşitli zorluklar ile karşılaşılmasına sebep olmaktadır. Söz konusu zorlukların başında ise zaman problemi gelmektedir. Büyük miktardaki verinin elle düzenlenmesi ve üzerinde bilgi çıkarımı yapılabilecek hale dönüştürülmesi oldukça zaman alıcı ve zorlu bir işlemdir. Karşılaşılabilecek bir diğer sorun ise, arşivde bulunan kaynakların Osmanlı Türkçesi ile kaleme alınmış olmasıdır. Bu durum söz konusu verilerin anlaşılabilmesi için ayrı bir uzmanlık gerektirmektedir. Ayrıca konu ile ilgili yapılacak olan analiz çalışmalarında ele alınacak veri kaynaklarının neler olduğunu doğru bir şekilde saptayabilmek adına İmparatorluğun ve arşivin yapısı hakkında genel bir bilgi sahibi olunması gerekmektedir. Bu durum analiz çıktılarının sonuçlarını doğrudan etkileyebilmektedir.

Çalışmanın amacı, oluşturulan veri kümesi içerisinde bulunan hükümlerin metin madenciliği yöntemiyle otomatik olarak sınıflandırılmasıdır. Böylece elle yapıldığında oldukça fazla iş yüküne ve zaman kaybına neden olan sınıflandırma işlemi uygun algoritmalar vasıtasıyla gerçekleştirilecektir. Ayrıca araştırmacıların yeni çalışmalarında kullanabilecekleri bilgi çıkarımının yapılması mümkün kılınacaktır. Oluşturulan sınıflandırma işlemi ile çeşitli istatistiksel şekil ve tablo elde edilebilecektir. Bu şekil ve tablolar, metin verisi içerisinde bulunan gizli örüntülerin ve olaylar, tarihler, kurumlar, makamlar, unvanlar, şahıslar, mekânlar ve nesneler gibi belgelerin düzenlenmesine yardımcı olan ayrıntılar arasındaki ilişkilerin ortaya çıkarılmasına olanak sağlayabilecektir.

Çalışmada 1 numaralı Erzurum Ahkâm defteri veri kaynağı olarak seçilmiştir. Söz konusu defter M.1742 ve M.1749 (H.1155-1162) tarihlerini kapsamaktadır. Defter içerisinde dönemin yerel mahkemelerinde çözüme kavuşturulamayan ve Dîvân-ı Hümayun'a intikal eden davalar bulunmaktadır. Defterde 4 adet derkenar ve 3 adet iptal edilmiş hüküm ile birlikte toplamda 1216 adet kayıt bulunmaktadır. Bahsedilen derkenar ve iptal edilen hükümler çalışmanın kapsamı dışında bırakılmıştır. Toplamda 1209 adet hüküm verisi çalışmanın örneklemini oluşturmaktadır. 1 numaralı Erzurum Ahkâm Defteri içerisinde bulunan hükümler Arap harfleri ile yazılmıştır. Osmanlı Türkçesi ile kaleme alınan hükümlerde Arapça ve Farsça kelimeler ile tamlamalar

bulunmaktadır. Bahsedilen hükümlerin transkript edilmesi ve konularının belirlenmesi oldukça kapsamlı ve zaman alıcı bir işlemdir. Bu işlemin yapılması ile tez çalışmasının süresinin fazlaca etkileneceği öngörülmüştür. 1 numaralı Erzurum Ahkâm Defterinin belirli sayfa aralıklarına göre transkript edildiği dört farklı tez çalışması mevcuttur. Defter üzerinde yapılacak olan çalışmanın verileri, bahsedilen bu dört tez çalışmasından elde edilmiştir. Söz konusu dört çalışmada Latin harflerine çevrilen hükümler özetlenmiştir ve ayrıca manuel olarak hükümlerin konuları da belirlenmiştir. Bu çalışmalardan elde edilen 1209 adet yarı yapılandırılmış hüküm verisi Excel programı vasıtasıyla bir araya getirilmiştir. Oluşturulan veri kümesi incelendiğinde ilk olarak, aynı içeriğe sahip hükümleri kapsayabilecek 140 farklı konu başlığının belirlendiği saptanmıştır. Söz konusu konu başlıkları sayısının bu derece fazla olmasının sebebi şu örnekle açıklanabilir; mal ve para gasbı ile ilgili hükümlerin asayiş başlığı altında toplanabilmesi mümkün iken bahsi geçen çalışmalarda mal gasbı, para gasbı ve mal ve para gasbı şeklinde üç ayrı konu başlığına ait hükümlerin olduğu görülmektedir. Bu durum konu başlıklarının belirli bir standarda dayandırılarak yapılmaması ile açıklanabilir. Söz konusu problem, özel konu başlıklarına sahip hükümlerin daha kapsayıcı genel başlıklar altında bir araya getirilmesi ile çözülmüştür. Böylece hükümler 3 ana başlık altında toplanmıştır ve başlıklar altında bulunan hüküm sayıları artmıştır. Bu durum sınıflandırma işleminde algoritmalarından daha başarılı sonuçlar elde edilmesinde doğrudan etkili olmuştur. Dolayısıyla incelenecek yarı yapılandırılmış veri miktarı artırıldığında, ana başlıklardan başka alt başlıklar üzerinden analizlerin yapılması ve böylelikle başarılı sonuçların elde edilmesi mümkün olabilecektir.

Yarı yapılandırılmış metin verilerinden oluşan hüküm veri kümesi üzerinde analizler yapılarak bilgi çıkarımının gerçekleştirilebilmesi için metin madenciliği yöntemleri uygulanmıştır. Açık kaynak kodlu olan KNIME programı vasıtasıyla verilere bir dizi veri ön işleme adımları uygulanmıştır. İlk olarak veri kümesi içerisinde bulunan metinler dokümanlara dönüştürülmüştür. Ardından tüm büyük harfler küçük harflere dönüştürülmüştür ve metin içerisinde bulunan noktalama işaretleri kaldırılmıştır. Kelimeler gövdeleme işlemi sonucunda köklerine ayrılmıştır. Veri kümesi içerisinde bulunan sayısal ifadeler ile bir anlam ifade etmeyen fakat sıklıkla

tekrar eden durak kelimeler kaldırılmıştır. Yazım yanlış ve farklılıklarının giderilerek kelimelerde standartlaştırma oluşturulması amacıyla kelime dönüştürme işlemi gerçekleştirilmiştir. Karakter, kişi ve konum etiketlemesi yapılarak ağırlıklandırma yapılmıştır. Ardından frekanslarına göre 5 ve daha az sıklıkla geçen kelimeler analiz dışında bırakılarak veri indirgeme yöntemi uygulanmıştır. Veri indirgeme ile veri boyutu küçültülerek daha hızlı ve başarılı sonuçların elde edilmesi amaçlanmıştır. Ardından doküman vektörü oluşturularak veriler sayısallaştırılmıştır. Böylece metin verilerinden oluşan veri kümesi algoritmalar tarafından anlaşılabilir (var ise 1 ve yok ise 0) bir yapıya dönüştürülmüştür. Gerçekleştirilen ön işlem aşamasının ardından veriler %70 eğitim verisi ve %30 test verisi olarak bölünmüştür. %70 oranının altındaki veri miktarının algoritmaların eğitilmesi işleminde yeterli olmayacağı, %70'in üzerindeki veri miktarının ise test işlemi için az sayıda verinin kalmasına sebep olacağı öngörülmüştür. Bu durum söz konusu oranın belirlenmesinde etkili olmuştur. Sınıflandırma işlemleri bitsel ağırlıklandırma (Binary), terim frekansı (TF), doküman frekansı (DF), ters doküman frekansı (IDF) ve ters sınıf frekansı (ICF) ağırlıklandırma yöntemlerine göre karar ağacı, destek vektör makineleri, rassal orman, k en yakın komşu ve Gradient Boosted ağaçları algoritmaları ile gerçekleştirilmiştir. İşlem sonuçlarının çıktıları karışıklık matrisi ile sunulmuştur. Yapılan sınıflandırma işlemlerinde en başarılı sonuç ters doküman frekansı ağırlıklandırma yöntemine göre Gradient Boosted ağaçları algoritması ile elde edilmiştir. Söz konusu algoritmanın F-ölçüsü değeri 0,809 olarak ve doğruluk oranı ise 0,812 olarak elde edilmiştir. Arazi meselesi ve vergi konu başlığı altında bir araya getirilen hükümler %86'lık bir doğruluk oranı ile sınıflandırılmıştır. Sosyal, ekonomik ve idarî konular ile ilgili hükümler ise %83,1'lük bir doğruluk oranı elde edilerek sınıflandırılmıştır. Doğruluk oranları göz önünde bulundurulduğunda oluşturulan sınıflandırma modelinin başarılı olduğu sonucuna varılmaktadır.

Sınıflandırma işlemlerinde asayiş konu başlığının doğruluk oranının diğer konu başlıklarına göre daha düşük olduğu gözlemlenmiştir. Bu duruma, konu başlıkları altında bulunan hüküm verilerinin sayıları arasındaki dengesizlik sebep olmaktadır. Bazı konu başlıkları altında asgari seviyede hüküm verisinin olmayışı bu durumun sebebi olarak gösterilebilir. Algoritmaların eğitilmesi için ayrılan veri sayısı

ile doğruluk oranı arasında doğru orantı vardır denilebilir. Veri sayısının ideal asgari düzeyi, her bir konu başlığına ait hüküm verilerinin birbirine yakın sayıda olması olarak belirlenebilir. 1 numaralı Erzurum Ahkâm Defteri özelinde meydana getirilen sınıflandırma modeli diğer Ahkâm Defterleri için de kullanılabilecektir. Bahsedilen veri sayısındaki asgari düzeyin belirlenmesi diğer veri setleri için de geçerlidir. Veri setinde bulunan hükümlerin birbirine yakın sayıda olmaması durumunda ise şu yöntemlerden biri tercih edilebilir: eğer çalışmada bir konu başlığı altında az sayıda (1 ile 5 adet arasındaki hükümler kabul edilebilir) hüküm var ise söz konusu hükümler çalışmanın dışında bırakılabilir ya da daha kapsayıcı konu başlıkları belirlenerek bu hükümler belirlenen konu başlıkları altında bir araya getirilebilir. Bu çalışmada ikinci yöntem tercih edilmiştir ve hüküm konu başlıkları daha genel başlıklar altında birleştirilmiştir.

Gerçekleştirilen literatür taraması sonucunda çalışmaya konu olan Ahkâm Defterleri dışında Osmanlı arşiv belgeleri üzerinde benzer fonları oluşturan verilerin metin madenciliği yöntemiyle ele alındığı bir çalışma bulunamamıştır. Çalışma; Ahkâm Defterleri üzerinde ilk kez metin madenciliği yöntemiyle sınıflandırma yapılması ve ayrıca yapılacak olan diğer çalışmalara katkı sağlayabilecek olması durumu sebebiyle önemlidir. El ile yapılan sınıflandırma işlemlerinin neden olduğu zaman kaybının minimize edilmesi, oluşturulan sınıflandırma modeli ile veri kümesine yeni eklenen verilerin çalışmaya kolaylıkla ve yüksek doğruluk oranıyla dahil edilmesinin mümkün kılınması ve elde edilecek istatistiksel veriler ile bilgi çıkarımlarına olanak sağlanması çalışmanın önemine işaret eden diğer unsurlardır. Bu çalışmanın ardından, oluşturulmuş olan sınıflandırma modelinin geliştirilerek başarı oranının artırılması ve bu hükümler içerisinde geçen kişiler, mekanlar, nesneler, tarihler, unvanlar, kurumlar ve olaylar arasında bulunan organik bağların incelenmesi hedeflenmektedir. Osmanlı arşiv belgeleri üzerinde gerçekleştirilecek metin madenciliği uygulamaları ile ilgili çalışmalara ve söz konusu çalışmanın geliştirilmesine yönelik ayrıca şu öneriler sunulabilir:

- Gerçekleştirilecek çalışmalarda verilerin çoğunluğu Osmanlı Türkçesi ve Arap harfleri ile oluşturulmuş kaynaklardan oluşacaktır. Bu kaynaklar üzerinde metin madenciliği işlemlerinin gerçekleştirilebilmesi için eksiksiz

ve doğru bir şekilde transkript edilmeleri gerekmektedir. Ele alınan veri kümesi içerisinde okunamayan kelimeler ya da bölümler veri kaybına ve dolayısıyla gerçekleştirilecek çalışmanın başarı oranının etkilenmesine neden olabilecektir. Transkript işlemlerinin başarıyla yapılabilmesi için, Osmanlı Türkçesi dil bilgisi hakkında yetkin olunması gerekmektedir. Bahsedilen alfabenin öğrenilmesi ve bunun yanı sıra büyük miktarlardaki verilerin transkripsiyonu oldukça zaman alıcı bir süreç olabilecektir. Bu zaman alıcı ve özel beceri gerektiren durumun giderilmesi için literatürde Arap harflerini Latin harflerine çevirmek için geliştirilen çalışmaların varlığı bilinmektedir. Örnek olarak Uzun, Özer ve Türkmen'in (2021: 581) yapmış olduğu çalışma gösterilebilir. Söz konusu çalışmada meydana getirilen sistem ile belgelerin %97 oranında doğru şekilde satırlara ayrıldığı ve satırlar üzerinde bulunan kelimelerin harflere ayrılması işleminin de %96 oranında başarılı bir şekilde gerçekleştirildiği belirtilmektedir. Örnek olarak gösterilebilecek bir diğer çalışma ise Dölek ve Kurt (2022: 1) tarafından gerçekleştirilmiştir. Söz konusu çalışmada geliştirdikleri Evrişimli Tekrarlayan Sinir Ağı (Convolutional Recurrent Neural Network-CRNN) tabanlı Optik Karakter Tanımlama (Optical Character Recognition-OCR) modeli ile 21 sayfalık veri seti içerisindeki karakterlerin %96 oranında bir başarı ile tanındığı belirtilmiştir. Ayrıca geliştirmiş oldukları alfabe çeviri sistemi ile 7500 kelimeyi içeren veri setinden %98'lik bir oran ile kelimelerin doğru bir şekilde çevrildiği de ifade edilmiştir. İlgili çalışmalar incelendiğinde söz konusu başarı oranlarının üzerinde yazı türlerinin belirleyici bir etkisi olduğu görülmektedir ve bu çalışmaların, yazılışı diğer türlere göre daha anlaşılır olan matbu metinlerde daha başarılı sonuçlar elde edebileceğini söylemek mümkündür. Gerçekleştirilecek metin madenciliği çalışmalarında transkripsiyonu yapılmış olan verilerin seçilmesi ya da tanıdıkları imkânlar doğrultusunda ilgili program ve algoritmalar kullanılarak transkripsiyon işleminin gerçekleştirilmesi gerekmektedir.

- Gerçekleştirilen transkript işlemlerinde aynı kelimelerin farklı şekillerde yazılmasının önüne geçilmesi için kelimelerin yazılışları hakkında bir standart oluşturulmalıdır.
- Veriler transkript edilirken söz konusu veriyi, veri kümesini, veri kaynağını, konusunu ve metin içerisinde bulunan özel isimleri tanımlayabilecek kapsamlı standart bir üstveri sistemi oluşturulmalıdır.
- Konu başlıklarının belirlenmesi için de bir standart geliştirilmelidir. Ele alınan veri kümesinin hacmine bağlı olarak bir araya getirilecek verilerin konu başlıkları değiştirilmelidir. Bu değişim ise önceden belirlenmiş konu başlıklarına göre olmalıdır. Her bir konu başlığı altında bulunan verilerin miktarı çalışmayı anlamlandırabilecek seviyede ise seçilen konu başlıkları daha dar anlamlı ve özel olmalıdır. Veri miktarı az ise daha kapsayıcı ve genel başlıklar seçilmelidir. Bu genel başlıkların hangi konuları içereceği önceden belirlenmiş olan standart konu başlıklarına göre şekillendirilmelidir.
- Metin madenciliği uygulamalarında kelimelerin aldığı eklerin veri kümesi içerisinde çıkarılması ve böylece kelimelerin köklerinin bulunması işlemlerinde Türkçe metinler için kullanılan Zemberek Kütüphanesi gibi Osmanlı Türkçesi için de kullanılabilecek bir kütüphane oluşturulması gerekmektedir. Böyle bir kütüphanenin olmayışı bu işlemlerin gerçekleştirilmesini zorlaştırmakta ve oldukça zaman almaktadır.
- Gerçekleştirilen işlemler sonucunda hedefe uygun algoritmalar tercih edilmeli ve elde edilen çıktılar disiplinler arası bir bakış açısıyla irdelenmelidir.
- Çıktılar farklı görselleştirme yöntemleri ile daha anlaşılır hale getirilmeli böylece gerçekleştirilebilecek yeni çalışmalara zemin hazırlanmalıdır.
- Günümüzde kurum ve kuruluşların faaliyetleri sonucu elektronik ortamda üretilen ya da elektronik ortama aktarılarak bir araya getirilen metinsel

verileri düzenlemek, yeni oluşturulan metinsel verileri otomatik olarak kategorilerine ayırmak ve metinsel veriler hakkında ilk bakışta anlaşılamayan çıkarımlarda bulunmak söz konusu model ile mümkündür. Osmanlı Türkçesi ve dönemin alfabesi ile yazılmış olan metinlere göre ise bu işlemler daha kolay ve daha başarılı bir şekilde gerçekleştirilebilecektir. Bu duruma sebep olarak metinsel verilerin transkript edilmesini gerektiren zorlu ve zaman alıcı ek bir sürecin olmaması, ilgili metinsel verilerin elektronik ortamda olması ya da elektronik ortama kolaylıkla aktarılabilmesi sayesinde hızlı ve basit bir şekilde erişilebilmesi, Latin harfleri ile oluşturulmuş Türkçe metinsel verilerin kelime köklerine ayrılmasını mümkün kılan Zemberek Kütüphanesinin var olması gösterilebilir. Yukarıda bahsedilen durum ve sebepler dolayısıyla günümüz kurumlarında üretilen Latin harfli metinlerin söz konusu model ile yapılandırılmış veri haline getirilerek üzerinde analiz ve çıkarım yapılabilmesi mümkündür. Bu doğrultuda günümüzde kurum ve kuruluşların faaliyetleri sonucu meydana gelen metinsel veriler ele alınmalıdır ve üzerlerinde gerçekleştirilen analiz ve çıkarımlar ile geleceğe yönelik yapılacak olan planlamalara ve çalışmalara kaynak oluşturulmalıdır.

KAYNAKÇA

- Adalı, Eşref: “Doğal Dil İşleme”, **Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi**, C.5, No:2, 2016, s. 1-19, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/207209>, 5 Ekim 2021.
- Akadal, Emre,
Özdemir, Şebnem,
Ayvaz Reis, Zerrin: “GNU Özgür Belgeleme Lisansı (GFDL) Kapsamındaki Dokümanlar İçin Bir Çevrimiçi Arşiv Geliştirilmesi”, **XV. Akademik Bilişim Konferansı Bildirileri 23-25 Ocak 2013**, Antalya, Türkiye, Ed. Mustafa Akgül vd., s. 855-859.
- Akgöbek, Ömer,
Kaya, Serkan: “Veri Madenciliği Teknikleri ile Veri Kümelerinden Bilgi Keşfi: Medikal Veri Madenciliği Uygulaması” **e-Journal of New World Sciences Academy**, C.6, No:1, 2011, s. 237-245, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/186263>, 1 Temmuz 2022.
- Akın, Ahmet Afşın,
Akın, Mehmet Dünder: “Zemberek, an Open Source NLP Framework for Turkic Languages”, **Structure**, 2007, C.10, (Çevrimiçi) <https://www.academia.edu/8066909/>, 3 Şubat 2022.
- Akyüz, Fatma,
Gülten, Selçuk: **Metin Madenciliği Kelime Muhasebesi ve Denetimi**. Ankara, Gazi, 2022.
- Amanet, Hasan: “Türkçe Sosyal Medya Metinlerinde Duygu Analizi”, Yayınlanmamış Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2017.
- Argüden, Yılmaz,
Erşahin, Burak: **Veri Madenciliği: Veriden Bilgiye, Masraftan Değere**. İstanbul, Arge Danışmanlık, 2008.

- Arslan, Alphan A.: **“Türkçe Metinlerden Anlamsal Bilgi Çıkarımı İçin Bir Veri Madenciliği Uygulaması”**, Yayınlanmamış Yüksek Lisans Tezi, Başkent Üniversitesi, Fen Bilimleri Enstitüsü, 2011.
- Atan, Suat: **“Metin Madenciliği: İmkânlar, Yöntemler ve Kısıtlar”**, **Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, C.31, 2020, s. 220-239, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/1117243>, 7 Ocak 2021
- Aydın, Bilgin: **“Osmanlı Bürokrasisinde Dîvân-ı Hümâyûn Defter Formlarının Ortaya Çıkışı ve Gelişimi (XV-XVI. Yüzyıl)”**, Yayınlanmamış Doktora Tezi, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, 2003.
- Aytaç, Muhammed B.,
Bilge, Hasan Şakir: **“Tele Pazarlama Verilerinin Birlikte Kurallarıyla ve CRISP-DM Yöntemiyle Analiz Edilmesi”**, **Aksaray Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2013, C.5, No:2, s. 25-40, (Çevrimiçi) <http://aksarayiibd.aksaray.edu.tr/pub/issue/22549/240981>, 19 Ocak 2022
- Babadağ, Melis: **“İstihbaratta Büyük Veri ve Veri / Metin Madenciliği: Bir Uygulama”**, Yayınlanmamış Yüksek Lisans Tezi, Polis Akademisi, Güvenlik Bilimleri Enstitüsü, 2021.
- Bağcı, Ömer: **İleri Düzey Excel: Excel 2016, 2019, 2021 ve Microsoft 365**. Ankara, Seçkin, 2022.
- Bardakoğlu, Ali: **“Diyet”**, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.9, İstanbul, 1994, s. 473-479.
- Başar, Fahmeddin: **“Osmanlı Devleti’nin Kuruluş ve Yükseliş Dönemi Siyasî Tarihi Üzerine Cumhuriyet Döneminde Yapılan Çalışmalar”**, **Türkiye Araştırmaları Literatür Dergisi**, C.1, No:2, 2003, s. 25-56, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/651512>, 12 Mart 2022.

Başbakanlık Osmanlı Arşiv Rehberi. Yayına Hazırlayan: İskender Türe ve Salim Kaynar, İstanbul, Seçil Ofset, 2017.

Bayer, Harun:

“Veri Madenciliğinde Bir Metin Madenciliği Uygulaması”, Yayımlanmamış Yüksek Lisans Tezi, Beykent Üniversitesi, Fen Bilimleri Enstitüsü, 2011.

Baykal, Abdullah,
Coşkun, Cengiz:

“Performance Analysis of Classification Algorithms of Several Data Mining Softwares”, **Middle East Journal of Science**, 2018, C.4, No:2, s.104-112, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/604010>, 1 Şubat 2022.

Berthold, R. Michael,
Cebron, Nicolas, Dill, Fabian,
Gabriel, Thomas R., Kötter,
Tobias, Meinl, Thorsten, Ohl,
Peter, Sieb, Christoph, Thiel,
Killian, Wiswedel, Bernd;

“KNIME: The Konstanz Information Miner”, içinde **Data Analysis, Machine Learning and Applications**, Ed. Christine Preisach v.d., s. 319-326, Springer, 2008.

Bilgin, T. Tugay:

“Veri Akış Diyagramları Tabanlı Veri Madenciliği Araçları ve Yazılım Geliştirme Ortamları”, **Akademik Bilişim’09-XI. Akademik Bilişim Konferansı Bildirileri 11-13 Şubat 2009**, Şanlıurfa, s. 807-814.

Binici, Kasım

“Kütüphane ve Bilgi Bilimi Çalışmalarında Dönemsel Konu Analizi”, Yayımlanmamış Doktora Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, 2016.

Bolasco, Segio, Canzonetti,
Alessio, Capo, Federico M.,
Ratta-Rinaldi, Francesca
della, Singh, Bhupesh K.:

“Understanding Text Mining: A Pragmatic Approach”, içinde **Knowledge Mining**, Ed. Spiros Sirmakessis, y.y., 2005, s. 31–50. C. 185. (Çevrimiçi) doi:10.1007/3-540-32394-5_4

Breiman, Leo:

“Random Forests”, **Machine Learning**, C. 45, 2001, s. 5-32, (Çevrimiçi) <https://link.springer.com/content/pdf/10.1023/A:1010933404324.pdf>, 4 Eylül 2021.

- Budak, Hüseyin: “Özellik Seçim Yöntemleri ve Yeni Bir Yaklaşım”, **Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi**, C.22, 2018, s. 21-31 (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/552933>, 30 Mart 2021.
- Canhası, Ercan: “**Edebiyat Eserlerinin Web Verilerine Dayanarak Sınıflandırılması**”, Yayımlanmamış Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, 2007.
- Cevger, Gürkan: “**1 Numaralı ve 1155-1158/1742-1745 tarihli Karaman Ahkâm Defterinin Transkripsiyonu ve Değerlendirilmesi (ss.1-60)**”, Yayımlanmamış Yüksek Lisans Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, 2012.
- Chien, Chen-Fu,
Chen, Li-Fei: “Data Mining to Improve Personnel Selection and Enhance Human Capital: A Case Study in High-Technology Industry”, **Expert Systems with Applications**, C.34, No:1, 2008, s. 280-290, (Çevrimiçi) <https://www.sciencedirect.com/science/article/pii/S0957417406002776>, 11 Şubat 2022.
- Covington, Michael A.: **Natural Language Processing for Prolog Programmers**. New Jersey, Prentice-Hall, 1994.
- Çalış, Kübra,
Gazdağı, Oya,
Yıldız, Oktay: “Reklam İçerikli Epostaların Metin Madenciliği Yöntemleri ile Otomatik Tespiti”, **Bilişim Teknolojileri Dergisi**, C.6, No:1, 2013, s. 1-7 (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/75317>, 30 Mart 2021.
- Çapkın, Çağdaş: “Üstverinin Tam-Metin Bilgi Erişim Performansı Üzerindeki Etkisi: Küçük Ölçekli Külliyyat Üzerinde Deneysel Bir Araştırma”, **Türk Kütüphaneciliği**, C.30, No:4, 2016, s. 678-701, (Çevrimiçi) <http://www.tk.org.tr/index.php/TK/article/view/2731/2691>, 2 Temmuz 2022.

- Çelik, Buket: **“Ahkâm Defterlerine Göre 18. Yüzyılda Sivas Eyaleti (1742-1774)”**, Yayınlanmamış Yüksek Lisans Tezi, Sivas Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, 2020.
- Çetinkaya, Zeynep,
Horasan, Fahrettin: **“Decision Trees in Large Data Sets”**, **Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi**, C.13, No:1, 2021, s.140-151, (Çevrimiçi) <https://dergipark.org.tr/tr/download/articlefile/1186009>, 16 Şubat 2022.
- Çiçek, Hikmet: **“Erzurum Vilayeti 1 Numaralı ve 1155-1162/1742-1749 Tarihli Ahkâm Defteri’nin Transkripsiyonu ve Değerlendirilmesi (s. 65-130)”**, Yayınlanmamış Yüksek Lisans Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, 2011.
- Çiftçi, Yunus: **“7 Numaralı Anadolu Ahkâm Defteri’nin Değerlendirilmesi ve Transkripsiyonu (H.1158/M.1745)”**, Yayınlanmamış Yüksek Lisans Tezi, Kafkas Üniversitesi, Sosyal Bilimler Enstitüsü, 2020.
- Demirtel, Hakan,
Sağıroğlu, Şeref: **“Özgür Dağıtıma Konu Olan Yazılımlar: Nedir, Ne Kadar Güvenlidir?”**, **TÜBAV Bilim Dergisi**, 2011, C.4, No:1, s. 52-63, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/200948>, 13 Ocak 2022.
- Devellioğlu, Ferit: **Osmanlıca-Türkçe Ansiklopedik Lügat**, Ankara, Aydın Kitapevi, 2008.
- Doğan, Korcan,
Arslantekin, Sacit: **“Elektronik Belge Yönetimi, Dijital Arşivleme Sistemleri ve Büyük Veri”**, içinde **Bilgi Sistemleri ve Bilişim Yönetimi: Beklentiler ve Yeni Yaklaşımlar**, Ed. Fahrettin Özdemirci, Zeynep Akdoğan, 2017, s. 65-80.

- Dolgun, Özgür M.,
Özdemir, Güzel Tülin,
Oğuz, Doruk:
- “Veri Madenciliği’nde Yapısal Olmayan Verinin Analizi: Metin ve Web Madenciliği”, **İstatistikçiler Dergisi**, 2009, C.2, s. 48-55, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/105653>, 11 Mart 2021.
- Dominich, Sandor:
- The Modern Algebra Of Information Retrival.** Veszprém, Springer, 2008.
- Dölek, İshak,
Kurt, Atakan:
- “Osmanlıcadan Türkçeye Uçtan Uca Aktarım”, **Journal of Smart Systems Research**, C.3, No:1, 2022, s. 1-10, (Çevrimiçi) <https://dergipark.org.tr/en/download/article-file/2171923>, 4 Temmuz 2022.
- Efendioğlu, Süleyman,
İşcan, Adem:
- “Türkçe Ses Bilgisi Öğretiminde Ses Olaylarının Sınıflandırılması”, **A. Ü. Türkiyat Araştırmaları Enstitüsü Dergisi (TAED)**, C.43, 2010, s. 121-143, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/33519>, 5 Temmuz 2020.
- Ekrem, Özge,
Salman, K. M. Osamah,
Aksoy, Bekir,
İnan, Seyit Ahmet:
- “Yapay Zekâ Yöntemleri Kullanılarak Kalp Hastalığının Tespiti”, **Mühendislik Bilimleri ve Tasarım Dergisi**, C.8, No:5, 2020, s. 241-254, (Çevrimiçi) <https://dergipark.org.tr/en/download/article-file/1392668>, 21 Şubat 2022.
- Emecen, Feridun M.:
- Osmanlı Klasik Çağında Hanedan, Devlet ve Toplum.** İstanbul, Timaş Yayınları, 2011.
- Emre, İlkin Ecem,
Selçukcan Erol, Çiğdem:
- “Veri Analizinde İstatistik mi Veri Madenciliği mi?”, **Bilişim Teknolojileri Dergisi**, C.10, No:2, 2017, s. 161-167, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/297864>, 3 Şubat 2022.

- Ergün, Mustafa: “Using The Techniques Of Data Mining And Text Mining İn Educational Research”, **Electornic Jourbal Of Education Sciences**, C.6, No:12, 2017, s. 180-189, (Çevrimiçi) [https:// atif. sobiad. com/ index. jsp?modul= makale-goruntule&id= AWCuoHwQoDuH9Br_d8tr](https://atif.sobiad.com/index.jsp?modul=makale-goruntule&id=AWCuoHwQoDuH9Br_d8tr), 3Temuz 2022.
- Erten, Filiz: “**Metin Madenciliği Tabanlı Bir Web Sitesi Sınıflandırma Aracı Tasarımı**”, Yayınlanmamış Yüksek Lisans Tezi, Maltepe Üniversitesi, Fen Bilimleri Enstitüsü, 2015.
- Feldman, Ronen,
Sanger, James: **The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data**. Cambridge, New York, Cambridge University Press, 2007,
- Frawley, William J.,
Piatetsky-Shaprio ,Gregory,
Matheus, Christoper J.: Knowledge Discovery in Databases an Overview, **AI Magazine**, C.13, No:3, 1992, s. 57-60, [https:// ojs. aai. org/ index. php/ aimagazine/ article/ view/ 1011/ 929](https://ojs.aaai.org/index.php/aimagazine/article/view/1011/929), 30 Aralık 2021.
- Gahi, Youssef, Guennoun,
Mouhcine Mouftah, Hussein
T.: "Big Data Analytics: Security and privacy challenges". **2016 IEEE Symposium on Computers and Communication (ISCC)**, 27-30 Haziran 2016, Messina, İtalya, s.952-957.
- Gedikli, Fethi: “Osmanlı Hukuk Tarihi Kaynağı Olarak Şer’iyye Sicilleri”, **Türkiye Araştırmaları Literatür Dergisi**, C.3, No:5, 2005, s. 187-213, (Çevrimiçi) [https:// dergipark. org. tr/ tr/ download/ article-file/ 652331](https://dergipark.org.tr/tr/download/article-file/652331), 4 Temmuz 2022.
- Gökbilgin, M. Tayyip: **Osmanlı İmparatorluğu Medeniyet Tarihi Çerçevesinde: Osmanlı Paleografya ve Diplomatik İlmî**. İstanbul, Enderun, 1992.

- Göker, Hanife,
Tekedere, Hakan: “Fatih Projesine Yönelik Görüşlerin Metin Madenciliği Yöntemleri İle Otomatik Değerlendirilmesi”, **Bilişim Teknolojileri Dergisi**, C.10, No:3, 2017, s. 291-299, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/328852>, 7 Ekim 2021.
- Gunn, R, Steve: “**Support Vector Machines for Classification and Regression**”, Technical Report, Faculty of Engineering, School of Electronics and Computer Science, University of Southampton, 1998.
- Gupta, Vishal,
Lehal, Gurpreet S: “A Survey of Text Mining Techniques And Applications”, **Journal of Emerging Technologies in Web Intelligence**, C.1, No:1, 2009, s. 60-76.
- Gülüm, Türker,
Köksal, Mehmet Ali: “Genel Kamu Lisansı Çerçevesinde Yazılımlarda Yeterlilik Oranı”, **İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi**, 2021, C.3, No:3, s. 193-196, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/1509446>, 13 Ocak 2022.
- Günay, Ramazan: “Osmanlı Arşiv Kaynakları İçerisinde Ahkâm Defterleri: Gelişim Seyri, Muhtevası ve Önemi”, **Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, C.1, No:17, 2013 (a), s. 9-29, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/215765>, 11 Kasım 2021.
- Günay, Ramazan: “Diyarbakir Ahkâm Defterleri ve Muhtevaları”, **Elektronik Sosyal Bilimler Dergisi**, C.12, No:45, 2013 (b), s. 290-307, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/70472>, 5 Temmuz 2022.
- Gündüz, Hakan: “**Borsa İstanbul (BIST) 100 Endeksi Yönünün Ekonomi Haberleri ile Tahmin Edilmesi**”, Yayınlanmamış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2013.

Güneş, İsmail:

“Kamu Kurumlarında Açık Kaynak Kodlu Yazılımların Kullanılmasının Ekonomik Faydaları: Yerel Yönetimler İçin Pilot Uygulama Önerisi”, **Karamanoğlu Mehmetbey Üniversitesi Sosyal ve Ekonomik Araştırmalar Dergisi**, No:3, 2007, s. 151-162, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/107495>, 13 Ocak 2022.

Gümüşsu, Ertuğrul:

“Metin Madenciliği Teknikleri İle Sosyal Medya Gönderilerinin Analiz Edilmesi: Ankilozan Spondilit Hastalığı Örneği”, Yayınlanmamış Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, 2019.

Gümüşsu, Ertuğrul,
Murat, Naci:

“Tamam ve Devam Etiketleriyle Atılan Tweetler ile Cumhuriyet ve Millet İttifakı Grupları Arasındaki İlişkinin İncelenmesi”, **Bilişim Teknolojileri Dergisi**, C.12, No:4, 2019, s. 287-298, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/843031>, 1 Temmuz 2020.

Güven, Ahmet:

“Türkçe Belgelerin Anlam Tabanlı Yöntemlerle Madenciliği”, Doktora Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2007.

Haltaş, Ahmet,
Alkan, Ahmet,
Karabulut, Mustafa:

“Metin Sınıflandırmada Sezgisel Arama Algoritmalarının Performans Analizi”, **Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi**, C.30, No:3, 2015, s.417-427, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/76587>, 22.01.2022.

Hamde, Muhammet Abu:

“Kurumsal Belgelere (Metin Verilerine) Metin Madenciliği Tekniği İle Erişimin Değerlendirilmesi: Türk Özel Sektörüne Yönelik Bir İnceleme”, Doktora Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, 2018.

Hearst, Marti:

“What is Text Mining?” (Çevrimiçi) <http://www.sims.berkeley.edu/~hearst/text-mining.html>, 7 Ekim 2003.

- İlhan, Ufuk: **“Application of K-NN and FPTC Based Text Categorization Algorithms to Turkish News Reports”**, Yayınlanmamış Yüksek Lisans Tezi, Bilkent Üniversitesi, Mühendislik ve Fen Bilimleri Enstitüsü, 2001.
- İlhan, Sevinç, Duru, Nevcihan, Karagöz, Şenol, Sağır, Merve: **“Metin Madenciliği ile Soru Cevaplama Sistemi”**, **ELECO Elektrik- Elektronik ve Bilgisayar Mühendisliği Sempozyumu**, Bursa, 2008, s. 26-30.
- İnalcık, Halil: **“Şikâyet Hakkı: ‘Arz-ı Hal ve ‘Arz-ı Mazhar’lar”**, **Osmanlı Araştırmaları Dergisi**, C.07-08, No:7-8, 1988, s. 33-54, (Çevrimiçi) <https://dergipark.org.tr/en/download/articlefile/112703>, 6 Ekim 2021.
- İpşirli, Mehmet: **“Cuma Selâmlığı”**, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.8, İstanbul, 1993, s. 90-92.
- Kantaroglu, Nazmiye Gül: **“1 Numaralı ve 1155-1162/1742-1749 Tarihli Erzurum Ahkâm Defterinin Transkripsiyonu ve Değerlendirilmesi (s. 212-293)”**, Yayınlanmamış Yüksek Lisans Tezi, Amasya Üniversitesi, Sosyal Bilimler Enstitüsü, 2019.
- Karaca, Mehmet Fatih, Günel Mustafa, Taştan, Akif Alkan: **“Metin sınıflandırmada Benzerlik Hesaplama Tekniklerinin Değerlendirilmesi”**, **XVII. Akademik Bilişim Konferansı Bildirileri 4-6 Şubat 2015**, Eskişehir, Türkiye, Ed. Mustafa Akgül v.d., s. 808-812.
- Kartal, Elif: **“Sınıflandırmaya Dayalı Makine Öğrenmesi Teknikleri ve Kardiyolojik Risk Değerlendirmesine İlişkin Bir Uygulama”**, Yayınlanmamış Doktora Tezi, İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, 2015.

Kaşıkcı, Tuğba,
Gökçen, Hadi:

“Metin Madenciliği ile E-Ticaret Sistemlerinin Belirlenmesi”, **Bilişim Teknolojileri Dergisi**, C.7, No:1, 2014, s. 25-32, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/75338>, 5 Şubat 2022

Keleş, Ahmet:

“3 Numaralı Bosna Ahkâm Defteri Transkripsiyon ve Değerlendirme (s.1-100)”, Yayımlanmamış Yüksek Lisans Tezi, Kırşehir Ahi Evran Üniversitesi, Sosyal Bilimler Enstitüsü, 2020.

Keleş, Mustafa Berk,
Keleş, Aytürk,
Keleş, Ali:

“Yapay Zekâ Teknolojisi ile Uçuş Fiyatı Tahmin Modeli Geliştirme”, **Turkish Studies-Information Technologies and Applied Sciences**, C.15, No:4, 2020, s.511-520, (Çevrimiçi) <https://app.tr.dizin.gov.tr/makale/TkRFNE5qVTFOU09/yapay-zeka-teknolojisi-ile-ucus-fiyati-tahmin-modeli-gelistirme>, 16 Şubat 2020.

Kılıç, Özcan, Demirtaş, Pınar,
Candan, Hüseyin, Akbayrak,
Emre Hasan, Kutlu, Osman:

“Arşivlerde Yapay Zekâ Uygulamaları: GEODİ-Atılım Üniversitesi Kadriye Zaim Kütüphanesi Ankara Dijital Kent Arşivi Örneği”, içinde **Bilgi Yönetimi ve Bilgi Güvenliği: eBelge-eArşiv-eDevlet-Bulut Bilişim-Büyük Veri-Yapay Zekâ**, Ed. Bahattin Yalçinkaya v.d., s. 307-324, Ankara, Türkiye, Ankara Üniversitesi Yayınları, 2019.

Kobayashi, Vladimer B.,
Mol, Stefan T., Vrolijk,
Jarno, Kismihok, Gabor:

“Text Mining in Career Studies: Generating Insights from Unstructured Textual Data”, içinde **Handbook of Research Methods in Careers**, Ed. Wendy Murphy v.d., s. 139-163, Cheltenham, UK: Edward Elgar Publishing, 2021.

Kodaman, Bayram:

“Osmanlı Devleti’nin Yükseliş ve Çöküş Sebeplerine Genel Bakış”, **Süleyman Demirel Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi**, No:16, 2007, s. 1-24, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/118020>, 15 Mart 2022.

- Kotu, Vijay,
Deshpande, Bala:
Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner. Waltham, MA, Morgan Koufmann, 2015. (Çevrimiçi) <https://www.sciencedirect.com/book/9780128014608>, 25 Şubat 2021.
- Kul, Muhittin:
“1 Numaralı ve 1155-1162/1742-1749 Tarihli Erzurum Ahkâm Defterinin Transkripsiyonu ve Değerlendirilmesi (s. 1-65)”, Yayımlanmamış Yüksek Lisans Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, 2011.
- Kutluca, Tamer:
“Excel Yazılımı İle Geliştirilen Bilgisayar Destekli Bir Öğretim Materyalinin Tasarlanması” Elektronik Eğitim Bilimleri Dergisi, C.2, No:4, 2013, s.40-49, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/147687>, 9 Ocak 2022
- Küçüksille, Engin:
“Veri Madenciliği Süreci Kullanılarak Portföy Performansının Değerlendirilmesi ve İMKB Hisse Senetleri Piyasasında Bir Uygulama”, Yayımlanmamış Doktora Tezi, Süleyman Demirel Üniversitesi, Sosyal Bilimler Enstitüsü, 2009.
- Küçüksille, Ecir Uğur,
Ateş, Nurulah:
“Destek Vektör Makineleri ile Yaramaz Elektronik Postaların Filtrelenmesi”, **Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi**, C.6, No:1, 2013, s. 1-7, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/207219>, 16 Ocak 2022.
- Kütükoğlu, Mübahat S.:
Osmanlı Belgelerinin Dili (Diplomatik). Kubbealtı Akademisi Kültür ve Sanat Vakfı, İstanbul, 1998.
- Ladha, L.
Deepa, T.:
“Feature selection methods and algorithms”, **International Journal on Computer Science and Engineering**, C. 3, No:5, 2011, s. 1787-1797, (Çevrimiçi) <http://www.enggjournals.com/ijcse/doc/IJCSE11-03-05-051.pdf>, 7 Ocak 2022

- Lernattee, Verayuth,
TheeramunkonG, Thanaruk: “Improving Thai Academic Web Page Classification Using Inverse Class Frequency and Web Link Information”, **22nd International Conference on Advanced Information Networking and Applications-Workshops (aina workshops 2008)**, 25-28 Mart 2008, Okinawa, s. 1144-1149.
- Menne-Haritz, Angelika: “Eriřim-Bir Arřivsel Paradigmanın Yeniden Formülasyonu” içinde **Arřivcilik Metinleri**, Yay. Haz. İřhak Keskin, s. 71-103, İstanbul, Türkiye, Yeditepe Yayınevi, 2008.
- Miner, Garry, Delen, Dursun, Elder, John, Fast, Andrew, Hill, Thomas, Nisbet, Robert A.: **Practical Text Mining and Statistical Analysis for Nonstructured Text Data Applications**. Waltham, MA, Academic Press, 2012. (Çevrimiçi) [http:// www. sciencedirect. com/ science/ book/ 9780123869791](http://www.sciencedirect.com/science/book/9780123869791), 8 Eylül 2021.
- Mumcu, Ahmet: “Dîvân-ı Hümâyûn”, **Türkiye Diyanet Vakfı İřlam Arařtırmaları Ansiklopedisi**, C.9, İstanbul, 1994, s. 430-432.
- Nisbet, Robert, Elder, John, Miner, Gary: **Handbook of Statistical Analysis and Data Mining Applications, Burlington**. MA, Academic Press, 2009, (Çevrimiçi) [https:// www. sciencedirect. com/ book/ 9780123747655](https://www.sciencedirect.com/book/9780123747655), 9 Eylül 2021.
- North, Matthew: **Data Mining for the Masses**. 2012. (Çevrimiçi) <https://docs.rapidminer.com/downloads/DataMiningForTheMasses.pdf>, 11 Kasım 2019.
- Oğuz, Başak, Bilge, Uğur, Saka, Osman: “Tıpta Metin Madenciliğı”, **Tıp Biliřimi Dergisi**, 2007, C.7, s. 15-18, (Çevrimiçi) [https:// turkmiia. net/ kongre2007/ cd/ pdf/ 108-112. pdf](https://turkmiia.net/kongre2007/cd/pdf/108-112.pdf), 18 Ocak 2021.
- Oğuzlar, Ayře: **Temel Metin Madenciliğı**. Bursa, Dora, 2011.

- Özcan, Abdülkadir: “Mahmud I”, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.27, Ankara, 2003, s. 348-352.
- Özdemirci, Fahrettin,
Torunlar, Mehmet: **Bilgi Çağında Arşivsel Bilgi Analizi: Bilgi-İktidar-İdeoloji-Devlet**, Ankara, Ankara Üniversitesi Basımevi, 2015.
- Özel, Ahmet: “Ahkâm”, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.1, İstanbul, 1988, s. 550-551.
- Öztürk, Said: “Sosyo-Ekonomik Tarih Kaynağı Olarak Ahkam Defterleri”, içinde **Pax Ottomana Studies in Memoriam Prof. Dr. Nejat Göyünç**, Ed. Kemal Çiçek, s. 611-639, C.1, No:1, Haarlem-Ankara, Sota-Yeni Türkiye, 2001.
- Öztürk, Yunus: “Etkileşimli Tahta Kullanımına Yönelik Sosyal Medya Yorumlarının Metin Madenciliği Yöntemleri İle Analizi”, Yayınlanmamış Yüksek Lisans Tezi, Sivas Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, 2021.
- Pakalın, Mehmet Zeki: **Osmanlı Tarih Deyimleri ve Terimleri Sözlüğü**, C.1, İstanbul, Millî Eğitim Bakanlığı Yayınları, 2004.
- Pasin, Ezgi: “Investigation Of Text Mining Methods On Turkish Text”, Yayınlanmamış Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, 2018.
- Pilavcılar, İsmail Ferhat: “Metin Madenciliği ile Metin Sınıflandırma”, Yayınlanmamış Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2007.
- Sahillioğlu, Halil: “Ahkâm Defteri”, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.1, İstanbul, 1988, s. 551.

- Sami, Şemseddin: **Kâmûs-ı Türkî: Latin Harfleriyle.** İstanbul, İdeal Kültür, 2011.
- Sarıcaoğlu, Fikret: “Dîvân-ı Hümayûn’un Kronolojik Toplanma ve Merasim Günleri (1153-1210, 1740-1795)”, **Osmanlı Araştırmaları Dergisi**, C.30, No:30, 2007, s. 87-143, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/112321>, 9 Ocak 2021.
- Seçkin, Keziban: **“Metin Madenciliğinde Kullanılan Yöntemlerin Karşılaştırılması: Siyasi Parti Liderlerinin Grup Genel Toplantı Konuşmaları İle Bir Uygulama”**, Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, 2011.
- Shearer, Colin: “The CRISP-DM Model: The New Blueprint For Data Mining”, **Journal Of Data Warehousing**, C.5, No:4, 2000, s. 13-23, (Çevrimiçi) <https://mineracaodedados.files.wordpress.com/2012/04/the-crisp-dm-model-the-new-blueprint-for-data-mining-shearer-colin.pdf>, 4 Temmuz 2022.
- Stallman, Richard M.: **Free Software, Free Society: Selected Essays of Richard M. Stallman.** Boston, MA, GNU Press, 2002.
- Suthaharan, Shan: **Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning.** Greensboro, USA, Springer, 2016.
- Sukanya, M.,
Biruntha. S.: “Techniques on Text Mining”, **2012 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)**, 2012, s. 269-271.

- Susam, Feyzanur: **“1 Numaralı ve 1155-1162/1742-1749 Tarihli Erzurum Ahkâm Defterinin Transkripsiyonu ve Değerlendirilmesi (s. 131-212)”**, Yayınlanmamış Yüksek Lisans Tezi, Amasya Üniversitesi, Sosyal Bilimler Enstitüsü, 2019.
- Şahin, İlhan,
Emecen, Feridun M.: **“Osmanlılarda Divân-Bürokrasi-Ahkâm: II. Bâyezid Dönemine Ait 906/1501 Tarihli Ahkâm Defteri”**, İstanbul, Türk Dünyası Araştırmaları Vakfı, 1994.
- Şeker, Şadi Evren: **“CRISP-DM: Endüstriler Arası Standart İşleme-Veri Madenciliği İçin (Cross Industry Standard Processing-Data Mining)”**, **YBS Ansiklopedi**, C.5, No:2, 2018, s.10-16.
- Şekeroğlu, Songül: **“Hizmet Sektöründe Bir Veri Madenciliği Uygulaması”**, Yayınlanmamış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2010.
- Şık, Muhammed Şamil: **“Veri Madenciliği Ve Kanser Erken Teşhisinde Kullanımı”**, Yayınlanmamış Yüksek Lisans Tezi, İnönü Üniversitesi, Sosyal Bilimler Enstitüsü, 2014.
- Şimşir, Nahide: **“Ahkâm Defterlerinin Tarihi Kıymeti ve 107 No’lu Ahkâm Defterindeki İzmir ile İlgili Hükümler”**, **Tarih İncelemeleri Dergisi**, C.9, No:1, 1994, s. 357-390, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/58334>, 14 Aralık 2021.
- Şimşir, Nahide: **“103 No’lu Anadolu Ahkâm Defteri’ndeki İzmir İle İlgili Hükümler”**, **Tarih İncelemeleri Dergisi**, C.10, No:1, 1995, s. 329-341, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/58355>, 15 Aralık 2021.

- Taner, Sönmez: “Bilgi Tanımlama ve Bilgi Erişim Aracı Olarak Thesaurus”, **Bilgi ve Belge Araştırmaları**, C.1, 2008, s. 85-107, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/170866>, 2 Temmuz 2022.
- T.C. Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı (BOA), **Bab-ı Asaî Ahkâm Defterleri**, Erzurum Ahkâm Defterleri, No:1.
- Tekerek, Adem: “Veri Madenciliği Süreçleri ve Açık Kaynak Kodlu Veri Madenciliği Araçları”, **Akademik Bilişim’11-XIII. Akademik Bilişim Konferansı Bildirileri**, 2-4 Şubat 2011, Malatya, s. 161-169.
- Temelkuran, Tefik: “Divân-ı Hümayun: Mühimme Kalemî”, **İstanbul Üniversitesi Edebiyat Fakültesi Tarih Enstitüsü Dergisi**, No:6, 1975, s. 128-175, (Çevrimiçi) http://isamveri.org/pdfs/D00120/1975_6/1975_6_TE_MELKURANR.pdf, 23 Haziran 2021.
- Thiel, Kilian: “Introduction to the Knime Text Processing Extension”, içinde **Text Mining and Visualization; Case Studies Using Open-Source Tools**, Ed. Markus Hofmann, Andrew Chisholm, s. 63-79, Boca Raton, FL, CRC Press, 2016.
- Tonta, Yaşar: “İnternet, Elektronik Kütüphaneler ve Bilgi Erişim”, **Türk Kütüphaneciliği**, C.10, No:3, 1996, s. 215-23, (Çevrimiçi) <http://www.tk.org.tr/index.php/TK/article/view/1475/1467>, 3 Temmuz.
- Torii, Manabu, Waghlikar, Kavishwar, Liu, Hongfang: “Using Machine Learning For Concept Extraction On Clinical Documents From Multiple Data Sources”, **Journal of American Medical Informatics Association**, C.18, No:5, 2011, s. 580-587, (Çevrimiçi) <https://academic.oup.com/jamia/article/18/5/580/832942?login=false>, 3 Temmuz 2022.

- Toplu, Selçuk Göksel: **“Metin Madenciliği ve Sağlık Alanında Bir Uygulama”**, Yayınlanmamış Yüksek Lisans Tezi, Düzce Üniversitesi, Sosyal Bilimler Enstitüsü, 2019.
- Tunalı, Volkan,
Bilgin, Turgay Tugay: **“Türkçe Metinlerin Kümelenmesinde Farklı Kök Bulma Yöntemlerinin Etkisinin Araştırılması”**, **ELECO Elektrik- Elektronik ve Bilgisayar Mühendisliği Sempozyumu**, 29 Kasım- 01 Aralık 2012, Bursa, s. 598-602.
- Tunç, Muhammet Nuri: **“Ahkâm Defterleri’nin Vakıf Araştırmalarındaki Yeri (1 Numaralı Erzurum, Karaman ve Maraş Ahkâm Defterleri), Çankırı Karatekin Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, C.13, No:1, s. 1-34, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/1953715>, 2 Temmuz 2022.
- Uğur, Yunus: **“Şer’iyye Sicilleri”**, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.39, İstanbul, s. 8-11.
- Uysal, Mevlüt,
Güyer, Tolga: **“İstatistiksel Veri Analizine İlişken Genişleyebilir Bir Karar Ağacı Tasarımı”**, **Bilişim Teknolojileri Dergisi**, C.7, No:3, 2014, s. 33-43, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/75346>, 7 Şubat 2022.
- Uzun, Alp,
Özer, Alperen,
Türkmen, H. İrem: **“Evrişimsel Sinir Ağı Tabanlı Osmanlıca Belge Çözümleyici”**, **International Journal of Advances in Engineering and Pure Sciences**, C.33, No:4, 2021, s. 581-591, (Çevrimiçi) <https://dergipark.org.tr/en/download/article-file/1607142>, 4 Temmuz 2021.
- Uzunçarşılı, İsmail Hakkı: **Osmanlı Devleti’nin Saray Teşkilatı**. Ankara, Türk Tarih Kurumu, 1998.
- Üzüm, İlyas: **“Hüküm”**, **Türkiye Diyanet Vakfı İslam Araştırmaları Ansiklopedisi**, C.18, İstanbul, 1998, s. 464-466

- Vlahovic, Nikola: “Web 2.0 and its Impact on Information Extraction Practices”, içinde **Proceedings of the International Conference on Applied Computer Science**, Eylül 2010, s. 574-579.
- Wirth, Rüdiger,
Hipp, Jochen: “CRISP-DM: Towards a Standart Process Model For Data Mining”, içinde **Proceedings Of The 4th International Conference On The Practical Applications Of Knowledge Discovery and Data Mining**, Nisan 2000, s. 29-39.
- Yıldırım, Hacı Osman, Atik, Vahdettin, Cebecioğlu, Murat, Demirbaş, Uğurhan, Karazeybek, Mustafa, Safi, Muhammed, Serin, Mustafa, Uslu, Osman, Yekeler, Numan, Yıldırım, Zahit: **5 Numaralı Mühimme Defteri (973/1565-1566): Özet ve İndeks**, Ankara, T.C. Başbakanlık Devlet Arşivleri Genel Müdürlüğü, 1994.
- Yıldırım, Zafer,
Hayırlıoğlu, Nezaket İslam: “Mühimme Defterlerinin Tarih Araştırmaları Açısından Önemi, 82 Numaralı Mühimme Defteri Örneği”, **Uluslararası Sosyal Bilgilerde Yeni Yaklaşımlar Dergisi**, C.3, No:2, 2019, s. 205-214, (Çevrimiçi) <https://dergipark.org.tr/tr/download/article-file/911500>, 7 Kasım 2021.
- Zhang, Yuntao, Gong, Ling, Wang, Yongcheng, Yin, Zhonghang: “An Effective Concept Extraction Method For Improving Text Classification Performance”, **Geo-Spatial Information Science**, C.6, No:4, 2003, s. 66-72, (Çevrimiçi) <https://www.tandfonline.com/doi/pdf/10.1007/BF02826953>, 4 Temmuz 2022.

EKLER

Berat Tevcihi	236, 580
Borç Meselesi	4, 5, 11, 12, 14, 17, 19, 21, 33, 42, 46, 47, 53, 69, 78, 79, 96, 104, 107, 117, 123, 126, 129, 131, 138, 139, 148, 150, 175, 176, 210, 211, 213, 234, 266, 271, 274, 281, 287, 292, 293, 294, 295, 299, 300, 317, 322, 325, 326, 333, 338, 339, 348, 355, 358, 360, 362, 364, 367, 368, 371, 373, 377, 379, 385, 396, 409, 434, 438, 447, 463, 471, 478, 481, 485, 486, 504, 510, 513, 518, 537, 543, 545, 553, 557, 575, 577, 585, 588, 590, 594, 599, 605, 636, 637, 660, 677, 691, 695, 714, 715, 718, 719, 732, 741, 778, 780, 787, 792, 800, 802, 809, 810, 825, 828, 853, 859, 870, 882, 898, 901, 907, 911, 914, 930, 934, 938, 979, 989, 992, 1005, 1007, 1016, 1024, 1031, 1049, 1054, 1055, 1060, 1062, 1074, 1075, 1079, 1081, 1083, 1089, 1099, 1104, 1105, 1106, 1113, 1118, 1119, 1130, 1152, 1161, 1183, 1187, 1193, 1199, 1204, 1212
Can Güvenliğinin Sağlanması	405, 1149
Cebelü Bedeliyesi	87, 106, 202, 569, 799, 888, 894
Cinayet	23, 177, 198, 214, 227, 255, 392, 539, 574, 626, 639, 640, 694, 702, 743, 793, 811, 816, 833, 840, 843, 960, 1091, 1128
Darp	1051, 1198
Diyyet Talebi	63, 82, 157, 187, 251, 653, 684, 685
Eşkiyalık	16, 44, 48, 49, 52, 56, 64, 65, 66, 73, 90, 102, 103, 114, 120, 121, 122, 134, 147, 153, 156, 168, 173, 196, 197, 199, 200, 201, 243, 245, 260, 263, 269, 288, 309, 310, 314, 315, 318, 319, 330, 332, 340, 344, 345, 347, 353, 361, 365, 369, 370, 378, 384, 387, 397, 399, 411, 413, 417, 425, 433, 435, 439, 444, 445, 451, 453, 464, 468, 476, 477, 479, 488, 490, 497, 501, 533, 534, 540, 541, 551, 558, 571, 579, 582, 591, 592, 609, 621, 661, 665, 717, 725, 744, 758, 803, 831, 832, 846, 849, 862, 869, 884, 897, 917, 919, 927, 931, 936, 942, 943, 944, 947, 950, 952, 953, 954, 955, 962, 967, 974, 975, 1008, 1022, 1027, 1030, 1045, 1048, 1069, 1080, 1086, 1116, 1123, 1125, 1138, 1141, 1142, 1147, 1148, 1189, 1191, 1197
Faiz	205, 206, 254, 786
Fuhuş	222
Göreve Müdahale	36, 119, 248, 249, 286, 334, 354, 528, 548, 607, 655, 670, 679, 697, 776, 1173

Göç	38, 70, 128, 140, 349, 407, 418, 421, 687, 782, 795, 851, 961, 991
Gündelik Yardımı	302
Haksız Hisse Talebi	160, 190, 552, 646
Haksız Mal Talebi	616, 643
Haksız Para Talebi	28, 98, 108, 280, 1006, 1066
Haksız Vergi Talebi	18, 22, 31, 32, 37, 54, 80, 83, 85, 94, 112, 118, 124, 137, 144, 149, 154, 161, 165, 203, 217, 225, 226, 230, 235, 241, 256, 268, 272, 273, 278, 283, 284, 296, 301, 304, 307, 308, 324, 341, 343, 350, 351, 357, 359, 374, 380, 381, 383, 386, 390, 401, 408, 415, 420, 427, 430, 440, 441, 450, 452, 473, 480, 489, 494, 520, 530, 535, 538, 547, 562, 564, 572, 589, 593, 595, 596, 597, 604, 610, 617, 658, 659, 662, 664, 673, 701, 703, 711, 726, 731, 735, 755, 765, 770, 783, 814, 815, 824, 847, 864, 874, 904, 915, 916, 928, 935, 937, 941, 980, 984, 986, 995, 1001, 1012, 1013, 1017, 1029, 1047, 1088, 1092, 1097, 1101, 1139, 1168, 1182, 1188, 1194, 1209
Hırsızlık	92, 136, 194, 699, 1033, 1214
Koruya Müdahale	700, 1100
Mal Gaspı	25, 41, 57, 59, 68, 88, 89, 145, 151, 228, 258, 277, 279, 555, 556, 682, 796, 807, 836, 850, 876, 957, 1000, 1010, 1072, 1078, 1082, 1085, 1087, 1107, 1111, 1145, 1146, 1162, 1202
Malikâne Tahsili	158, 159
Malikâneye Müdahale	84, 221, 261, 312, 375, 422, 424, 446, 495, 515, 519, 523, 524, 613, 622, 1171
Mehir	692
Meraya Müdahale	298, 531, 763, 819, 1036, 1103, 1109, 1144
Miras İhtilafı	223, 270, 282, 502, 503, 583, 638, 644, 649, 817, 993, 1184, 1210
Mirasa Müdahale	15, 101, 171, 209, 366, 388, 428, 466, 484, 522, 529, 566, 576, 584, 675, 681, 683, 723, 728, 733, 737, 774, 820, 821, 823, 839, 856, 883, 906, 925, 966, 973, 988, 1090, 1093, 1136, 1154, 1155, 1160, 1167, 1203
Mülk Anlaşmazlığı	721, 722, 739, 1205
Mülke Müdahale	2, 3, 13, 24, 39, 58, 61, 100, 109, 111, 113, 135, 167, 170, 172, 174, 178, 182, 191, 193, 238, 239, 240, 285, 313, 320, 321, 331, 336, 346, 363, 382, 398, 404, 429, 436, 448, 454, 456, 467, 474, 487, 498, 499, 532, 536, 542, 546, 625, 645, 647, 654, 657, 663, 672, 676, 686, 693, 712, 713, 716, 740, 747, 748, 752, 829, 830, 852, 857, 893, 896, 913, 940, 949, 965, 976, 978, 981, 982, 990, 1002, 1014, 1046, 1052, 1053, 1065, 1108, 1110, 1122, 1153, 1156, 1164, 1174, 1175
Nafaka	667

Nikâh Akdi	67, 492, 669, 678, 801, 945, 1018, 1025, 1114
Para Gaspı	8, 9, 40, 91, 97, 164, 166, 169, 544, 560, 602, 635, 698, 706, 760, 804, 861, 867, 891, 899, 968, 1015, 1026, 1034, 1037, 1042, 1084, 1176, 1200
Resmî Evrak ve Değerli Malların Korunması	525
Satış Anlaşmazlığı	29, 43, 45, 267, 926, 948, 971, 1032, 1038, 1211
Sorumluluklarını Yapmama	6, 470
Soruşturma	224, 372
Su Anlaşmazlığı	34, 86, 184, 195, 218, 233, 244, 250, 259, 432, 496, 581, 754, 790, 791, 1020, 1035, 1039, 1059, 1169
Sınır İhtilafı	865
Tayin Talebi	95
Tecâvüz	74, 311, 400, 1207
Tedarik	652
Tereke Taksimi	20, 99, 105, 155, 262, 297, 483, 516, 586, 611, 612, 651, 929, 1098, 1179, 1206
Timar Anlaşmazlığı	767, 1120
Timar Mutasarrıflığı	680
Timar Tahsili	808, 998, 1040, 1127
Timar Tevcihi	133, 189, 220, 406, 512, 784, 855, 858, 872, 877, 985, 1011, 1063, 1064, 1112, 1121, 1131, 1165, 1185
Timara Müdahale	7, 26, 27, 60, 76, 130, 162, 183, 188, 242, 247, 257, 265, 291, 316, 403, 410, 412, 426, 442, 443, 449, 465, 472, 500, 514, 521, 550, 567, 568, 570, 603, 608, 619, 620, 623, 633, 634, 642, 656, 674, 688, 704, 729, 730, 745, 749, 750, 756, 759, 761, 766, 773, 797, 813, 822, 834, 837, 866, 871, 873, 879, 885, 889, 892, 900, 903, 908, 909, 912, 920, 922, 923, 924, 933, 939, 959, 970, 983, 996, 1019, 1028, 1043, 1044, 1057, 1058, 1067, 1070, 1094, 1096, 1102, 1124, 1140, 1143, 1150, 1157, 1172, 1178, 1192, 1201
Ulûfeye Müdahale	394
Vakfa Müdahale	30, 50, 71, 72, 125, 152, 192, 252, 335, 342, 352, 389, 391, 416, 423, 455, 469, 475, 491, 493, 505, 508, 517, 549, 559, 561, 565, 631, 632, 641, 650, 666, 689, 696, 707, 724, 727, 738, 746, 769, 779, 785, 789, 794, 798, 806, 818, 826, 845, 848, 860, 863, 875, 880, 881, 887, 956, 987, 997, 999, 1041, 1071, 1076, 1077, 1132, 1133, 1151, 1181, 1195, 1208, 1215
Vakıf Tevcihi	237, 264
Vakıf Yönetimi Tevcihi	1, 204, 977
Vazife Tevcihi	886, 1050, 1177
Vazifeyi Kötüye Kullanma	253, 327, 329, 337, 395, 437, 461, 482, 554, 563, 630, 720, 827, 905, 918, 1004, 1126, 1134

Vergi Vermeme	55, 81, 93, 110, 115, 143, 185, 219, 275, 276, 303, 573, 587, 600, 606, 618, 629, 708, 757, 762, 764, 768, 788, 805, 812, 835, 838, 841, 842, 878, 972, 1056
Yaylağa Müdahale	751, 771, 951, 958, 1023, 1186
Yol Kapama	844
Yurtluğa ve Ocaklığa Müdahale	356
Zeamet Mutasarrıflığı Tevcihi	1129, 1196
Zeamet Tahsili	127, 208, 509, 932, 1117
Zeamete Müdahale	10, 35, 51, 62, 75, 77, 116, 132, 141, 142, 163, 179, 180, 186, 231, 232, 246, 289, 290, 305, 323, 376, 402, 419, 431, 457, 458, 459, 460, 462, 511, 526, 527, 601, 614, 624, 628, 648, 671, 690, 709, 710, 734, 742, 753, 772, 775, 777, 781, 868, 895, 902, 921, 963, 964, 969, 1003, 1061, 1095, 1115, 1135, 1158, 1163, 1166, 1213, 1216
Zimmete Mal Geçirme	615
Öşr Tahsili	507, 910, 946, 1009
Öşr'e Müdahale	146, 506, 627, 668, 994, 1021, 1073
İftira	181, 207, 212, 215, 216, 393, 414, 736, 890, 1068, 1190
İhtida	229, 1159
Şahıs Güvenliğinin Sağlanması	306, 854