

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI
YÖNETİM BİLİŞİM SİSTEMLERİ PROGRAMI
YÜKSEK LİSANS TEZİ

EMLAK SEKTÖRÜNDE MAKİNE ÖĞRENME
TEKNİKLERİ KULLANILARAK İŞ ZEKASI
UYGULAMASI GELİŞTİRİLMESİ

Hakan Can TAŞCI

Danışman

Dr. Öğr. Üyesi Can AYDIN

İZMİR - 2019

TEZ ONAY SAYFASI



YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “Emlak sektöründe makine öğrenme teknikleri kullanılarak iş zekası uygulaması geliştirilmesi” adlı çalışmanın, tarafımdan, akademik kurallara ve etik değerlere uygun olarak yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

Tarih/....../.....

Hakan Can TAŞCI

ÖZET

Yüksek Lisans Tezi

**Emlak Sektöründe Makine Öğrenme Teknikleri Kullanılarak İş Zekası
Uygulaması Geliştirilmesi**

Hakan Can TAŞCI

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

Yönetim Bilişim Sistemleri Anabilim Dalı

Yönetim Bilişim Sistemleri Programı

Bilgi ve iletişim teknolojilerinde yaşanan gelişmelerle birlikte, işletmelerin verilerini anlamlandırmaya olan gereksinimleri de gün geçtikçe artmaktadır. Bu bağlamda sürekli akan ve artan bu verinin etkin ve etkili bir biçimde yönetilebilmesi işletmeler için büyük önem arz etmektedir. Bu durum da, işletmelerin iş süreçlerini yeni teknolojilere adapte etmesini zorunlu kılmaktadır.

İş zekası uygulamaları, bu anlamda son zamanlarda en çok kullanılan çözümlerin başında gelmektedir. Bu büyük veriyi anlamlandırarak analiz etmek zorunda kalan işletmeler, makine öğrenmesi gibi teknikleri belirli operasyonlarına entegre etmeye başlamışlardır. Özellikle emlak sektörünün dinamik ve geniş yelpazede veri sağlayan yapısı, etkin ve gelişmiş raporlama araçlarının kullanımını da beraberinde getirmiştir. Bu çalışma, emlak sektöründe etkili analiz sonuçları elde edebilmek için kullanılan yeni nesil bir teknolojik yaklaşım önermektedir.

Bu çalışmanın amacı, bankalar, kredi finansman kuruluşları, inşaat şirketleri, aracılar, değerlendirme uzmanları ve danışmanlara; emlak fiyatları ve endeksleri, değerlendirme modülü, makro-ekonomik veriler ve trend analizleri gibi bilgiler sağlamak, piyasayı yakından takip etmek, sektör fırsatlarını tespit ederek, operasyonel, taktiksel ve stratejik kararlar alınmasına yönelik bir iş zekası uygulaması geliştirmektir.

İş zekası uygulaması tasarlanırken kullanılan yazılımlar ve frameworkler: PHP, Python ve Flask dır.

Anahtar Kelimeler: İş Zekası, Emlak, Fiyat Tahminlemesi, Makine Öğrenmesi, Flask .



ABSTRACT

Master Thesis

Developing Business Intelligence Application by Using Machine Learning Techniques in Real Estate Sector

Hakan Can TAŞCI

Dokuz Eylul University

Graduate School of Social Sciences

Department of Management Information Systems

Management Information Systems Program

With the developments in information and communication technologies, the needs of enterprises to interpret their data are increasing day by day. In this context, it is important for enterprises to manage this continuously flowing and increasing data effectively. This situation necessitates enterprises to adapt their business processes to new technologies.

In this sense, business intelligence applications are one of the most used solutions in recent years. Businesses that had to analyze and interpret this big data began to integrate techniques such as machine learning into their specific operations. Especially the dynamic and wide data structure of the real estate sector has brought about the use of effective and advanced reporting tools. This study proposes a new generation technological approach used to achieve effective analysis results in the real estate sector.

The aim of this work is to provide information to banks, credit financing institutions, construction companies, intermediaries, valuation experts and consultants; to provide information such as real estate prices and indices, valuation module, macro-economic data and trend analysis, to closely monitor the market, to identify sector opportunities and to develop a business intelligence application for operational, tactical and strategic decisions.

Software and frameworks used for designing Business Intelligence Application: PHP, Python and Flask.

Keywords: Business Intelligence, Real Estate, Price Prediction, Machine Learning , Flask .



**EMLAK SEKTÖRÜNDE MAKİNE ÖĞRENME TEKNİKLERİ
KULLANILARAK İŞ ZEKASI UYGULAMASI GELİŞTİRİLMESİ
İÇİNDEKİLER**

TEZ ONAY SAYFASI	ii
YEMİN METNİ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER	viii
KISALTMALAR	xi
TABLOLAR LİSTESİ	xii
ŞEKİLLER LİSTESİ	xiii
EKLER LİSTESİ	xv
GİRİŞ	1

**BİRİNCİ BÖLÜM
EMLAK SEKTÖRÜNDE BİLGİ SİSTEMLERİ GEREKSİNİMLERİ**

1.1. PROBLEMİN TANIMI	6
1.2. İLGİLİ ÇALIŞMALAR	7

**İKİNCİ BÖLÜM
MAKİNE ÖĞRENMESİ**

2.1. MAKİNE ÖĞRENMESİ KAVRAMI	9
2.2. ÖĞRENME ÇEŞİTLERİ	11
2.2.1. Gözetimli Öğrenme	11
2.2.2. Yarı- Gözetimli Öğrenme	11
2.2.3. Gözetimsiz Öğrenme	11
2.2.4. Pekiştirmeli Öğrenme	12
2.3. MAKİNE ÖĞRENMESİNDE KULLANILAN ALGORİTMALAR	12

2.3.1. Gradient Boosting (GBM – Gradyan Artırma)	12
2.3.2. Karar Ağaçları (DT)	13
2.3.3. K-En Yakın Komşu Algoritması (KNN)	14
2.3.4. Lasso Regresyon	15
2.3.5. Çoklu Lineer Regresyon (MLR)	15
2.3.6. Polinomsal Regresyon (PR)	16
2.3.7. Rassal Orman (RF- Random Forest)	16
2.3.8. Ridge Regresyon	17
2.3.9. Destek Vektör Regresyon (SVR)	18
2.3.10. Facebook Prophet	19
2.3.11. ARMA	19
2.2. MODEL PERFORMANS DEĞERLENDİRME YÖNTEMLERİ	20
2.2.1. Ortalama Mutlak Hata (MAE – Mean Absolute Error)	20
2.2.2. Ortalama Kare Hatası (Mean Squared Error -MSE)	20
2.2.3. Kök Ortalama Kare Hatası (Root Mean Square Error-RMSE)	21
2.2.4. R^2	21

ÜÇÜNCÜ BÖLÜM

İŞ ZEKÂSI

3.1. İŞ ZEKÂSI KAVRAMI	22
3.2. İŞ ZEKÂSI SİSTEMLERİNİN MİMARİ YAPISI	23
3.3. İŞ ZEKÂSI UYGULAMA ORTAMLARI	26
3.4. MAKİNE ÖĞRENMESİ VE İŞ ZEKASI İLİŞKİSİ	27
3.5. İŞ ANALİTİĞİ VE İŞ ZEKÂSI İLİŞKİSİ	28

DÖRDÜNCÜ BÖLÜM

İŞ ZEKASI TASARIMI VE UYGULAMASI

4.1. METODOLOJİ	30
4.2. VERİLERİN TOPLANMASI	35
4.3. PROBLEMİN TANIMLANMASI	37

4.4. VERİYİ ANLAMA	37
4.5. VERİ HAZIRLAMA (VERİ ÖNİŞLEME)	40
4.5.1. Eksik Verileri Tamamlama	40
4.5.2. Metin Değerlerin Numerik Hale Getirilmesi	41
4.5.3. Özellik Seçimi	43
4.5.4. Aykırı Değerlerin Tespiti ve Çıkarılması	44
4.5.5. Özellik Mühendisliği	46
4.5.6. Normalizasyon	46
4.6. MODEL KURMA	48
4.7. MODEL DEĞERLENDİRME VE SEÇİMİ	48
4.8. MODELİN UYGULAMAYA GEÇİRİLMESİ	66
4.9. ARAYÜZ GELİŞTİRİLMESİ	67
4.9.1. Problemin Tanımı	67
4.9.2. İhtiyaç Analizi	67
4.9.3. Sistem Tasarımı	67
4.9.3.1. Arayüz Tasarımı	68
4.9.2. Sistem Yazılımının Geliştirilmesi	72
SONUÇ VE ÖNERİLER	74
KAYNAKÇA	76
EKLER	

KISALTMALAR

API	Uygulama Programlama Arayüzü (Application Programming Interface)
GİB	Gelir İdaresi Başkanlığı
PHP	Nesne Tabanlı Bir Programlama Dili (PHP: Hypertext Preprocessor)
s.	Sayfa No
vb.	Ve benzeri
vd.	Ve diğerleri
TCMB	Türkiye Cumhuriyeti Merkez Bankası
TÜİK	Türkiye İstatistik Kurumu
YBS	Yönetim Bilişim Sistemleri (Management Information Systems)
SGYD	Sistem Geliştirme Yaşam Döngüsü
GBM	Gradient Boosting (Gradyan Artırma)
DT	Karar Ağaçları
KNN	K En Yakın Komşu
MLR	Çoklu Lineer Regresyon
PR	Polinomsal Regresyon
RF	Rassal Orman
SVR	Destek Vektör Regresyon
MAE	Ortalama Mutlak Hata
MSE	Ortalama Kare Hatası
RMSE	Kök Ortalama Kare Hatası
GUI	Grafik Kullanıcı Arayüzü

TABLÖLER LİSTESİ

Tablo 1: İş Zekası ve Makine Öğrenmesi Farkları	s. 28
Tablo 2: İş Zekası ve İş Analitiği Farkı	s. 29
Tablo 3: Veri Setleri ve Açıklamaları	s. 36
Tablo 4: Tahminleme Algoritmalarında Kullanılan Veri Setinin Özellikleri ve Açıklamaları	s. 39
Tablo 5: Tahminleme algoritmalarının karşılaştırılması $k = 5$ için	s. 49
Tablo 6: Tahminleme algoritmalarının karşılaştırılması $k=10$ için	s. 49
Tablo 7: Facebook Prophet Algoritması ile Başarı Oranları	s. 50



ŞEKİLLER LİSTESİ

Şekil 1: Maslow İhtiyaçlar Hiyerarşisi Teorisi Piramidi	s. 1
Şekil 2: Rassal Ormanın Yapısı	s. 16
Şekil 3: Rassal Ormanın Ağaç Yapısı	s. 17
Şekil 4: İş Zekası Sistemleri Mimari Yapısı	s. 24
Şekil 5: Uygulama İş Akışı	s. 30
Şekil 6: İş Zekası Geliştirme Modeli	s. 31
Şekil 7: CRISP-DM Modeli	s. 32
Şekil 8: Sistem Geliştirme Yaşam Döngüsü	s. 34
Şekil 9: İlan Verilerine Ait Genel Bilgi	s. 38
Şekil 10: USD/TRY Veri Setine Ait Genel Bilgi	s. 38
Şekil 11: İlan Verilerine Yönelik Eksik Veriler	s. 39
Şekil 12: Eksik Verilerin Kontrolü	s. 41
Şekil 13: İlan Veri Seti Korelasyon Matrisi	s. 43
Şekil 14: Fiyat ile En Fazla Korelasyona Sahip Özellikler	s. 44
Şekil 15: Aykırı Değer Tespiti Korelasyon Grafikleri	s. 45
Şekil 16: Aykırı Değerler Çıkarıldıktan Sonra Konut Fiyat Dağılımı	s. 45
Şekil 17: İlan Veri Seti Kelime Bulutu	s. 46
Şekil 18: İlan Veri Seti Normalize Edilmeden Öncesi Dağılım	s. 47
Şekil 19: İlan Veri Seti Normalizasyon Sonrası Dağılımı	s. 48
Şekil 20: GBM Algoritmasına Göre Özellik Önemleri	s. 51
Şekil 21: RF Algoritmasına Göre Özellik Önemleri	s. 52
Şekil 22: İnşaat Maliyet Endeksi Zaman Serisi Grafiği	s. 52
Şekil 23: İnşaat Maliyet Endeksi Trend Çizgisi ve Yıllık Dağılımı	s. 53
Şekil 24: İnşaat Ciro Endeksi Zaman Serisi Grafiği	s. 53
Şekil 25: İnşaat Ciro Endeksi Trend Çizgisi ve Yıllık Dağılımı	s. 54
Şekil 26: İnşaat Güven Endeksi Zaman Grafiği	s. 54
Şekil 27: İnşaat Güven Endeksi Trend Çizgisi ve Yıllık Dağılımı	s. 55
Şekil 28: Konut Kredisi Faiz Oranları Zaman Serisi Grafiği	s. 56
Şekil 29: Konut Kredisi Faiz Oranları Trend Çizgisi ve Yıllık Dağılımı	
Şekil 30: Buca Konut Satış Sayıları Zaman Serisi Grafiği	s. 57

Şekil 31: Buca Konut Satış Sayıları Trend Çizgisi ve Yıllık Dağılımı	s. 58
Şekil 32: İzmir TÜFE Zaman Serisi Grafiği	s. 58
Şekil 33: İzmir TÜFE Trend Çizgisi ve Yıllık Dağılımı	s. 59
Şekil 34: USD/TRY Zaman Serisi Grafiği	s. 60
Şekil 35: USD/TRY Trend Çizgisi ve Yıllık Dağılımı	s. 61
Şekil 36: Adatepe Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği	s. 62
Şekil 37: Adatepe Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı	s. 62
Şekil 38: Barış Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği	s. 63
Şekil 39: Barış Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı	s. 63
Şekil 40: Akıncılar Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği	s. 64
Şekil 41: Akıncılar Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı	s. 64
Şekil 42: Atatürk Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği	s. 65
Şekil 43: Atatürk Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı	s. 65
Şekil 44: ARMA Algoritması ile Karşılaştırma	s. 66
Şekil 45: Uygulama Arayüz Anasayfası	s. 68
Şekil 46: Arayüz Tahminler Sayfası	s. 69
Şekil 47: Uygulama Analiz Arayüzü	s. 70
Şekil 48: Arayüz Trend Analizleri Ekranı	s. 70
Şekil 49: Arayüz Dış Kaynak Analizleri Ekranı	s. 71
Şekil 50: Uygulama Çıkarımlar Arayüzü	s. 72
Şekil 51: Flask Frameworkü Örnek Kod Bloğu	s. 73

EKLER LİSTESİ

EK 1: Makine öğrenmesi için kullanılan model kodu

ek s.1



GİRİŞ

Son yıllarda ülkemizde sosyo-ekonomik sektörlerin başında konut ve gayrimenkul sektörü gelmektedir. Nüfus artışı, kentleşme, iç-dış göç ve konut tüketim alışkanlıklarından dolayı konut gereksinimine neden olmaktadır. Diğer bir açıdan konutların yüksek gelirli bir yatırım aracı olarak görülmesinin günümüzde konut arz ve talebinin dolayısıyla fiyatlarında artmasına neden olmuştur.

Maslow'un ihtiyaçlar hiyerarşisini incelediğimizde insan yaşamının sürdürülebilmesi için beşli bir kademelenme önerilmektedir (Şekil 1).

Şekil 1: Maslow İhtiyaçlar Hiyerarşisi Teorisi Piramidi



Kaynak: Maslow, 1970

Maslow'a göre güduları kademeli bir yapı içinde görmeli ve insanların alt basamaktaki gereksinmelerini halleder halletmez üst basamaktakilere yöneleceğini kabul eder. İlk basamak olan fizyolojik ihtiyaçlar insanların yaşamlarını devam ettirebilmeleri için gerekli olan hava, su, besin, barınma ve neslin devamını getirebilme gibi temel düzeyde olan ihtiyaçları alır. Dolayısıyla insanın kaliteli bir hayat sürdürebilmesi için barınma ihtiyacını karşılaması gerekir.

Emlak; ev, dükkân, mağaza, turistik mekânları, oteller, arsa, çiftlik gibi gayrimenkullerin genel ismidir. Emlak kelimesi genellikle konutu çağrıştırmakla beraber mülk kelimesinin çoğuludur. Ev, işyeri, arsa, tarla, bağ, bahçe gibi, tasarruflara

konu olabilen taşınmaz mallara emlak denir (İçli, 2008: 2). Gayrimenkul kelime anlamı başka anlamında “gayr” ile taşınmış anlamına gelen “menkul” kelimelerinin birleşmesiyle meydana gelmiştir. Yapı anlamında gayrimenkul karada gerekse su üzerinde yapıldığı madde ne olursa olsun sabit inşaatın hepsini kapsamaktadır. Ancak emlak diğer taşınmazları da kapsamaktadır. Emlak türlerini alttaki gibi sıralayabiliriz (İçli, 2008: 3):

- Gecekondü
- Apartman dairesi
- Müstakil ev
- Villa
- Mobil ev
- Yazlık ev
- Toplu konut
- Arsa
- Arazi

Konut; bireyin barınmasını sağlayacak kapalı bir alandan ziyade doğal afetlere dayanıklı yapıları ifade etmektedir. Kentlerin var olması için olan konutlar, sadece barınak değil; kişilerin ve ailelerin mahremiyetinin korunduğu, güvende hissettiği, günlük hayat içerisinde sağlıklı ve kaliteli bir yaşam sürmesi amacıyla dinlenip enerji depoladığı bir mekândır. Konut insanlar için bir güvencedir (Daşkiran, 2015:850). İlk çağlarda konutların temel işlevi mağaradakiler gibi insanları dış etkilere karşı korumakken günümüzde birden fazla oda olması bunların üzerine mutfak, banyo gibi eklentilerle insanların tüm ihtiyaçlarını karşılayabilecekleri mekânlara dönüşmüşlerdir. Hatta sitelerde temel ihtiyaçları dışında spor salonları, kafeler ve saunalar gibi sosyal ihtiyaçları da karşılamaya başlamıştır. Barınma ihtiyacının kalitesi hane halkının ekonomik gücü ve istekleri ile yakından ilgilidir. Ekonomik gücü homojen olmayan bir yapıya sahiptir. Diğer taraftan üretilen konutlar da sadece barınma ihtiyacını karşılamakla kalmayıp, insanları mutlu etmek ve isteklerini karşılamak için oldukça heterojendir (Bulut ve diğerleri, 2015:121). Konut sektörü, ihtiyaçlarının artmasından dolayı en hareketli piyasalardan biridir. Bu bağlamda konut bazı özelliklere sahiptir (Özkurt, 2007:160).

1. İnşaat sektörünün alt bileşeni konut sektörüdür.
2. Konut nihai olarak bir maldır.
3. Konut sektörü; hizmetler, sanayi ve tarım sektörü gibi bir çok sektörle ilgilidir.
4. Konut sektörü gelişmemiş ülkelerde bir spekülasyon aracı olarak kabul edilmektedir.
5. Konut sektörü istihdamı olumlu yönde etkilemektedir.

Konut satın alımını etkileyen faktörler nüfus, nüfusun büyüme hızı, gelir, hane halkı sayısı ve hane halkının özellikleri gibi demografik faktörler olduğu görülmektedir (Martin, 1966:503). Günümüzde konut satın alınırken sadece demografik faktörlere değil aynı zamanda sosyo-psikolojik faktörlere de bakılmalıdır. Dolayısıyla bireylerin konut satın alma talebi , demografik faktörlerinden, ekonomik durumlarından, beklentilerinden, ihtiyaçlarından ve zevklerinden etkilenmektedir (Durkaya, 2002: 10; Abar ve Karaaslan, 2013: 324).

Konut satın alma talebini etkileyen demografik faktörler incelendiğinde birçok alt etken ortaya çıkmaktadır. Yaş, cinsiyet, medeni durum, eğitim gibi birçok faktör bireyin konut satın alma talebi üzerinde etkili olduğu anlaşılmaktadır (Öztürk ve Fitöz, 2009: 28).

Konut satın alma talebini etkileyen Sosyo-psikolojik faktörler incelendiğinde sosyal güvenliğin az olduğu ülkelerde bireyler için bir güven aracı ortaya çıkmaktadır. Güven aracından kastı barınmadan ziyade yatırım aracı olarak görünmesidir (Öztürk ve Fitöz, 2009: 28). Konut satın alırken etkileyen diğer bir sosyo-psikolojik etken ise bireylerin başarılarını simgeleştirme ve sosyal statü kazanmasıdır (Karagöz vd., 2015: 248).

Konut satın alma talebini etkileyen başka bir faktörse fiyattır. Konut fiyatları kullanılan malzeme miktarı, çeşidi, kalitesi, konutun büyüklüğü, sunulan altyapı ve üstyapı olanakları, işçilik kalitesi, diğer yapım maliyetleri ve arsa maliyetinden etkilenmektedir (Öztürk ve Fitöz, 2009: 26). Konutun fiziksel özelliklerinin fiyatı etkilemesine ek olarak dünya ve ülke genelinde ekonomik yapı da konut fiyatları

üzerinde etkili olabilmektedir. Enflasyon ve faiz oranlarının konut fiyatları üzerinde etkili olduğunu göstermektedir (Durkaya, 2002: 15). Bütün bu değişkenler göz önünde bulundurulduğunda fiyatın bireyin konut satın alma talebi üzerinde oldukça etkili bir etken olduğu söylenebilir.

Konut satın alma talebini etkileyen başka bir faktörse gelirdir. Gelirin konut satın alma talebini dolaylı yoldan etkilediğini söylenmektedir. Blumenfeld (1944) yaptığı çalışmada gelir artışına bağlı olarak evliliklerin arttığını ve buna bağlı olarak konut satın alma talebinin arttığını belirtmektedir. Lee'ye (1963) göre bireyler yaşlarıyla orantılı bir şekilde konut için harcayacağı gelirlerin getirisini planlayarak bireyin konut satın alma talebini etkilemektedir. Durkaya (2002)'nın yaptığı çalışmada yüksek gelir dağılımı grubunda yer alan bireyler konut satın almalarında lüks konutları tercih ederken, düşük ve orta grupta yer alan bireyler gelirleri doğrultusunda taleplerde bulunabilmekte veya talepte bulunmamayı tercih edebilmektedir.

Konut satın alma talebini etkileyen gelir ile ilgili bir diğer konu kredi koşullarıdır. Mevcut kredi alabilme olanakları ve ödeme koşullarının dışında kredi faiz oranları da bireyin konut satın alma talebinde etkilidir. Yüksek faiz oranlarının kredi alabilme olanaklarını düşürdüğü ve bu bağlamda konut satın alma talebini olumsuz yönde etkilediği söylenebilir (Durkaya, 2002: 29). Yayar ve Gül (2014) yapmış oldukları çalışmada Mersin'de konut fiyatlarını etkileyen faktörleri incelediklerinde konutun kullanım alanı, mutfak büyüklüğü, pazara uzaklık, banyo sayısı, garaj, merkezi uydu sistemi, özel güvenlik ve asansör sayısı değişkenlerinin konut fiyatlarını artırdığı; öte yandan konutun bahçeye sahip olması, site içinde olması, toplu taşıma araçlarına uzak olması ve eski olmasının konut fiyatlarını azalttığı sonucuna ulaşmışlardır.

Konut sektörü iktisadi faaliyetlerden dolayı piyasanın gelişim süreçlerinde vazgeçilmez bir sektör olduğu ortaya çıkmaktadır (İlhan, 2008: 13). Ekonomistlere göre, konut yatırımlarının; yapı gereçleri, kullanılan malların üretim kesimine yansması ve yeni konutların döşenmesi gibi etkisinin fazla olmasından dolayı ulusal ekonomiye katkısı büyüktür (Keleş, 1983: 125). Bu nedenle konut piyasası başlangıçta istihdam olmak üzere ekonominin birçok alt sektörüne katkıda bulunmaktadır. Konut piyasası, başta demir çelik, çimento, boya, seramik, cam, ahşap, tuğla ve kiremit

olmakla birlikte fazla sayıda malzemeyi kapsayan diğer sanayi ürünlerinin önemli kaynağıdır (Bayraktar, 2007: 135). Konut sektörü sadece Türkiye’de değil, diğer dünya ülkeleri arasında da önemini gittikçe artırmaktadır.

Çalışmanın amacı, bankalar, kredi finansman kuruluşları, inşaat şirketleri, aracılar, değerlendirme uzmanları ve danışmanlara; emlak fiyatları ve endeksleri, değerlendirme modülü, makro-ekonomik veriler ve trend analizleri gibi bilgiler sağlamak, piyasayı yakından takip etmek, sektör fırsatlarını tespit ederek, operasyonel, taktiksel ve stratejik kararlar alınmasına yönelik bir iş zekası uygulaması geliştirmektir. Ayrıca, çeşitli aktörleri içerisinde barındıran emlak piyasasındaki fiyat ve fiyata etki eden faktörleri dinamik olarak takip edebilmek hedeflenmektedir.

Araştırmanın temel konusu olan “Emlak sektöründe makine öğrenme teknikleri kullanılarak iş zekası uygulaması geliştirilmesi”, dört ayrı bölümden oluşmaktadır. Birinci bölümde araştırmayla ilgili genel bilgiler verilmekte olup, ikinci bölümde makine öğrenmesi kavramı, amacı, türleri ve tez kapsamında yararlanılan algoritmalar açıklanmaktadır. Üçüncü bölümde, iş zekası kavramı, iş zekası mimarisi, iş zekası uygulama ortamları ve iş analitiğiyle ilişkisine değinilmektedir. Dördüncü bölümde, sistemin tasarımı ve uygulanması anlatılmaktadır. Sonuç ve öneriler kısmında, çalışmadan temel çıkarımların neler olduğu ve sonraki çalışmalara yön verebilecek önerilere yer verilmektedir.

BİRİNCİ BÖLÜM

EMLAK SEKTÖRÜNDE BİLGİ SİSTEMLERİ GEREKSİNİMLERİ

Bilgi ve iletişim teknolojilerinin emlak sektöründe hızla gelişmesi farklı gereksinimleri ortaya çıkarmaktadır. Özellikle hali hazırda kullanılan yazılım çözümlerinin bu sektördeki işletmelerin ihtiyaçlarını karşılayamaması, emlak sektöründeki işletmeleri zamansal ve mali kayıplarla baş etmek zorunda bırakmaktadır. Bu anlamda kullanıcı dostu, işlevsel ve sade, işletmenin farklı amaçlara yönelik ihtiyaçlarını karşılayacak yazılımlara gereksinim duyulmaktadır.

Bu açıdan bakıldığında iş zekası yaklaşımının emlak sektöründeki işletmelere etkin ve etkili bir şekilde entegre edilmesi gerekmektedir. Çünkü, farklı seviyede yöneticilerin kararlarına destek amaçlı, yöneticilere anlamlı raporların sunulması bu işletmeler için büyük önem taşımaktadır. Öncelikle işletmenin iş zekası uygulamasına olan ihtiyacını ortaya koymak ve geliştirme aşamasına buradan başlamak verimli bir karar verme süreci için büyük önem taşımaktadır.

1.1. PROBLEMİN TANIMI

Günümüzde, işletmeler hizmet verdikleri sektörden bağımsız olarak bazı bilgilere anlık olarak ihtiyaç duymaktadırlar. Bu durumun eksikliği ise çoğu zaman işletmenin ekstra maliyetlere katlanmasına neden olmaktadır. Bu da işletmeleri veri yönetimiyle ilgili olarak yeni stratejiler belirlemeye teşvik etmektedir.

Bu nedenle, uygulamanın geliştirilmesine başlamadan önceki aşamada, emlak sektöründeki işletmelerinin iş zekası uygulaması gereksinimleri analiz edilmiş ve yapılan bu analiz sonucunda yazılımın çözüm bulabileceği sorunlar belirlenmiştir. Buna göre;

- Emlak kararlarında, özellikle farklı kaynaklardan toplanan ve biriken verinin anlamlı hale getirilmesi noktasında mevcut raporlama araçlarının yetersiz kalması,
- Karmaşık veri kümelerinden, ihtiyaç duyulan bilgiye anlık olarak, hızlı ve etkin bir şekilde ulaşma zorluğu sebebiyle özellikle emlak sektöründe zamansal ve mali kayıpların ortaya çıkması,

- Ülke ekonomisinin lokomotiflerinden biri olan emlak sektöründe yapılan hatalı yatırımların ekonomik krizlere neden olması, bunun sonucunda finansal istikrarın hem doğrudan hem de dolaylı olarak etkilenmesi
- Raporlama araçlarının kullanıcı dostu olmamasından dolayı, özellikle yönetici kademesinin karmaşık veriler içerisinden analiz yapmalarının gerekmesi, uygulamanın çözüm sağlayacağı problemler olarak belirlenmiştir.

1.2. İLGİLİ ÇALIŞMALAR

Pashardes ve Savva (2009), 1988-2008 yılları arasında Kıbrıs'ta makro ve mikro değişkenlerin konut fiyatlarını nasıl etkilediğini araştırmışlardır. Müstakil ve yarı müstakil, yatak odası sayısı, binanın büyüklüğü, en yakın şehir merkezine uzaklık ve yer gibi çeşitli konut tipleri üzerinde 4872 gözlem örneği kullanmışlardır. Hangi faktörlerin artan emlak fiyatına en fazla katkıda bulunduğu dair soru ile ilgili olarak, malzeme ve işçilik, nüfus ve kişi başına gayri safi yurtiçi hasıla maliyetinin bir analizini yapmışlar ve artan fiyatlara nasıl katkıda bulunduklarını belirlemişlerdir. Ayrıca, ikamet özelliklerinin, ev tipinin/büyüklüğünün ve en yakın şehir merkezine olan mesafenin, bireysel düzeyde fiyat üzerinde belirleyici bir etkiye sahip olduğu sonucuna varmışlardır.

Apergis ve Rezitis (2003), Yunanistan'da konut kredisi oranları, enflasyon, istihdam ve para arzı gibi belirli makroekonomik değişkenlerin dinamik etkilerini analiz etmiştir. Yukarıdaki makroekonomik değişkenlerin konut fiyatına etkisinin modellenmesinde bir hata düzeltme vektörü olarak Auto Regressive modeli kullanılmıştır. Ayrıca, paranın bir evin fiyatı üzerinde önemli bir etkisi olmadığı belirtilmiştir.

Filho ve Bin (2005), araştırmalarında parametrik olmayan model kullanarak konut hedonik fiyat fonksiyonu modeli oluşturmak için parametrik özellikleri ileri süren regresyon kullanmıştır. Kullanılan konut özellikleri; alan, lot büyüklüğü ve bina yaşı olmakla birlikte, merkezi iş bölgesi çevresel özelliklerinde, en yakın sulak alan ile park alanına yakınlığın satış fiyatları üzerinde olumsuz etkisi olduğu; sanayi bölgeleri konut fiyatlarına zarar verirken, ticari bölgelere en yakın merkezi iş bölgesi için ise olumlu bir etkisi olduğu vurgulanmıştır.

Baldominos ve diğerkleri (2018) arařtırmalarında, Madrid'in Salamanca bölgesindeki küçük bir alana odaklanmışlardır. Arařtırmanın amacı, emlak piyasasında fırsatların gerçek zamanlı olarak belirlenmesine yardımcı olabilecek bir makine öğrenmesi uygulaması geliřtirmektir. Odak noktası çevrimiçi olarak yayınlanmış olan ev listeleridir. Çalışmada çok deęişkenli regresyon kullanılmıştır. Ayrıca, destek vektör regresyonu, k-en yakın komşu, karar ağacı ve çok katmanlı algılayıcı gibi parametrelere başvurulmuştur. Çıktı olan fiyat, ikili, kategorik ya da süreklilik arzeden parametreler kullanılarak tahmin edilebilmektedir.Özellikler ise , asansör, sandık odası, bahçe, otopark olup olmadığı, apartman dairesi ya da villa olma durumu ve yüzme havuzu olması, dubleks ya da çatı katı olması gibi listelenebilmektedir. Kategorik parametreler ise yeri, kat sayısını, oda sayısını ve banyo sayısını içermekteir. Süreklilik arzeden parametreler arasında ise cadde adı, cadde numarası, yapım yılı, otopark fiyatı ve toplum maliyetler bulunmaktadır. Ayrıca Sürekli parametrelerde inşa alanı ve kullanılabilir alana da yer verilmiştir.Sonuçlar için, daha iyi performans gösteren model, karar ağacı olarak belirlenmiştir. Nicel olarak, arařtırmacılar en küçük ortalama mutlak hatanın 338.715 Euro, en iyi ortalama mutlak hatanın ise 94.850 Euro olduęu sonucuna erişmişlerdir. Bu hatalar, klasik lineer regresyon modelinin sağladığı hatalardan önemli ölçüde daha küçüktür, bu nedenle bu durumda kullanılan daha karmaşık makine öğrenme algoritmalarının avantajı vurgulanmaktadır.

İKİNCİ BÖLÜM

MAKİNE ÖĞRENMESİ

2.1. MAKİNE ÖĞRENMESİ KAVRAMI

Makine öğrenmeleri yapay zekânın bir alt boyutudur. Verilerden önemli davranışlar ve kurallar çıkartarak geleceğe yönelik tahminler yapmamızı kolaylaştırır (Alpaydin, 2004). Bilimsel çalışmalarda ve endüstride kullanılırlar.

Makine öğrenmesinin kesin bir tanımı yoktur. Simon (2013)'a göre makine öğrenmesi bilgisayarları programlamadan sonra onlara öğrenme yeteneği kazandıran çalışma alanıdır. Makine öğrenmesinde amaç deneyimlerden öğrenmek ve gelecekteki örneklemeler için genelleme yapabilmektir (Bishop ve diğerleri, 2006). Bir başka araştırmada makine öğrenmesi verinin anlaşılabilir aksiyonlara dönüştürülmesini sağlayan, bilgisayar algoritmalarının geliştirilmesine dayalı çalışma alanı olarak tanımlanmaktadır (Cui, Wong ve Lui, 2006: 598).

Makine öğrenmesi, veri madenciliği, istatistik, veri analizi, yapay zekâ ve matematik güçlü bir şekilde birbirleriyle ilişkilidir. İstenmeyen e-postaların tanınması, dolandırıcılığın ortaya çıkarılması, basamak tanıma, konuşma tanıma ve işleme, yüz tanıma, ürün tavsiyesi, medikal teşhis, stok ticareti, müşteri analizi, şekil tanıma, eğlence sektörü, otomotiv sektörü, eğitim, moda, fintech, insan kaynakları, hukuk makine öğrenmesinin kullanım alanlarındandır.

Makine öğrenmesiyle çok benzer olan veri madenciliğinde, büyük verilerden farklı anlayışlar elde edilmesi amaçlanmakta, genelde bu iki benzer alanın ne kadar örtüştüğü tartışılmaktadır. Temelde bu iki alanın arasındaki fark; makine öğrenmesi tanımlanmış bir görevin sonuçlanması beklenirken, veri madenciliği veri gurubu içerisinde gizli olan ve o verinin sırlarını barındıran bilgilere ulaşmayı hedeflemektedir (Gerlein ve diğerleri, 2016: 193). Örneğin, makine öğrenmesi ile bir robota nasıl araba kullanılacağı öğretilirken, veri madenciliği ile hangi arabaların daha güvenli olduğu bulunabilmektedir (Lantz, 2013: 7).

Öğrenme işlemini gerçekleştirecek olan insan ya da makinelerin izleyeceği süreç aşağıdaki üç temel adımda gerçekleşmektedir (Shwartz ve David, 2014: 19):

- *Girdi veriler:* Verilerin gözlenmesi, hafızaya alınması ve daha sonraki değerlendirme için gerekli olan anımsama bu bölümde sağlanır.
- *Çıkarım:* Verinin daha geniş anlamda sunuma imkân sağlayacak şekile dönüştürülmesini sağlayan bölümdür.
- *Genelleştirme:* Aksiyon alınmasına temel oluşturacak çıkarım yapılmış verilerin kullanılmasıdır.

Hedef olarak belirtilen amaca ulaşabilmek adına makine öğrenmesi için aşağıda sıralanan adımların takip edilmesi gerekmektedir (Zocca ve diğerleri, 2017: 9-10):

- *Verilerin Toplanması:* Analiz yapılmadan önce gerekli olan verilerin elektronik ortama kurallarına uygun şekilde aktarılması gerekmektedir. Söz konusu toplanan veriler, makinenin üzerinde öğrenme gerçekleştireceği temeli oluşturmaktadır.
- *Verilerin İncelenmesi ve Hazırlanması:* Makine öğrenmesi algoritmasının kaliteli olması verinin ne kadar kaliteli olduğuna dayanmaktadır. Bu aşamada ise insan müdahalesi gerekmektedir. Analiz yaparken bireylerin veriyi ve yapısını anlaması gerekir ve gerektiği takdirde düzeltmeler yapması gerekmektedir.
- *Modelin Veri Üzerinde Eğitilmesi:* İlk olarak veri setinden öğrenilmesi gerekenler net bir şekilde belirlenmelidir. Makine öğrenmesiyle gerçekleşmesi beklenen amaç doğrultusunda veriye uygun olan hangi model algoritma kullanılacağı belirlenmekte ve sunulmaktadır.
- *Modelin Performansının Değerlendirilmesi:* Kullanılmakta olan modelin türüne uygun olan test verilerinin ne kadar isabetli bir model seçildiği sınanır.
- *Modelin Performansının Artırılması:* Test verileriyle sınanan model, beklenen performansı sergileyemediği takdirde, modelin değiştirilmesi, ilave veriler eklenmesi ya da her ikisini de kapsayacak değişiklikler yapılması gerekebilmektedir.

2.2. ÖĞRENME ÇEŞİTLERİ

Eğitim metodolojisi ve veri setinin karakteristik yapısına göre makine öğrenmesi dört temel kategoriye ayrılmaktadır. Bunlar gözetimli öğrenme, yarı-gözetimli öğrenme, gözetimsiz öğrenme ve pekiştirmeli öğrenmedir.

2.2.1. Gözetimli Öğrenme

Makine eğitilirken önceden işaretlenmiş olan veri setlerini alarak tahmin yapılmasını ya da sınıflandırma kararı verilmez. (Russell ve Norvig, 2003: 650).

Gözetimli öğrenme girdi ve çıktı değerlerinin birlikte olduğu veri setleri için gerçekleştirilebilen öğrenme biçimidir. Hangi girdi değerlerinin hangi çıktı değerine karşılık geldiğinin bilinmesi veri setinin etiketli olduğu anlamına gelir. Bu öğrenme tekniğinde algoritmalar girdi ve çıktı değerleri arasındaki ilişkiyi ifade etmesi beklenen fonksiyonlar üretmeye çalışır. Bu fonksiyon yardımıyla yeni bir örnek ve örneği oluşturan girdiler verildiğinde hangi çıktı değerinin oluşacağını tahmini yapılır. Üretilen fonksiyonun çıktıları eğri uydurmak için ya da girdilerin sınıflarının tahmin edilmesi için kullanılabilir (Mohri, 2012).

Doğrusal regresyon modeli, lojistik regresyon, destek vektör makinaları, karar ağaçları, topluluk öğrenme yöntemleri ve yapay sinir ağları gözetimli öğrenme algoritmalarına örnek olarak verilebilir (Niculescu-Mizil ve diğerleri, 2005:627).

2.2.2. Yarı- Gözetimli Öğrenme

Yarı-gözetimli öğrenmede yeni örneklerin etiketlerinin tahmin edilmesi için hem etiketlenmiş hem de bir miktarda etiketlenmemiş veri kullanılır (Apaydın, 2010).

2.2.3. Gözetimsiz Öğrenme

Sadece girdi veri setlerinden oluşan öğrenme biçimidir. Etiketlenmiş bir veri yoktur. Dolayısıyla öğrenmede değerlendirme yapılması zordur. Bu kategoride amaç,

benzer özelliklere göre elementlerin gruplandırılmasıdır. Sonuçların isabeti, yapılan kümeleme işleminde, her bir birimin dâhil edildiği gruba ne ölçüde benzediği ve diğer gruplardan ne ölçüde farklı olduğu ile ölçülmektedir (Berry ve Linoff, 2004: 57).

Temel bileşen analizi, çok boyutlu ölçekleme ve kümeleme analizi gözetimsiz öğrenme metodlarına örnek olarak verilebilir (Apaydın, 2010).

2.2.4. Pekiştirmeli Öğrenme

Pekiştirmeli öğrenmede öğrenme sistemine yerine getirmesi için bir hedef verilmekte ve bu hedefi deneme-yanılma yoluyla gerçekleştirmesi beklenmektedir. Sistem deneme-yanılma işlemlerini en çok ödül sağlayan eylemlerine göre düzenlemekte ve verilen hedefe nasıl ulaşacağını kendisi öğrenmektedir (Mohri, 2012).

Makine öğrenmesinde, yukarıda da açıklandığı üzere farklı yaklaşımlar ve yöntemler (algoritmalar) bulunmakla birlikte, makinenin öğrenmesine esas olacak unsurların temelde veri setinin özelliklerine ve öğrenilmesi amaçlanan konuya göre değiştiğini ifade etmek mümkündür.

2.3. MAKİNE ÖĞRENMESİNDE KULLANILAN ALGORİTMALAR

2.3.1. Gradient Boosting (GBM – Gradyan Artırma)

Gradyan artırma, hem regresyon hem de sınıflandırma problemleri için çok yararlı bir algoritmadır. En büyük avantajlarından biri, karışık veri türlerini doğal olarak ele alması ve ayrıca aykırı değerlere oldukça dayanıklı olmasıdır. Ek olarak, diğer birçok algoritmadan daha iyi tahmin gücü vardır; bununla birlikte, sıralı mimarisi paralel teknikler için uygun değildir ve bu nedenle büyük veri kümelerine iyi ölçeklenememektedir. Gradyan yükseltme tipik olarak, düşük işlevli bir topluluğa dayalı bir tahmin modeli oluşturmak için karar ağaçları kullanır ve maliyet fonksiyonunda bir optimizasyon algoritması kullanır. (Raschka, Julian ve Hearty, 2016:599)

Gradyan artırma zayıf öğrenicileri bir araya getirir ve kayıp fonksiyonunun negatif gradyanı ile maksimum olarak ilişkilendirilen yeni bir temel öğrenen yaratır. Biri bu yöntemi regresyon veya sınıflandırma problemlerine uygulayabilir ve farklı veri setlerinde iyi sonuç verir. (Wiswanathan, 2016: 774).

2.3.2. Karar Ağaçları (DT)

Karar ağaçları, Makine öğrenimi alanındaki en güçlü ve yaygın kullanılan modelleme tekniklerinden biri olarak bilinir. Karar ağaçları, doğal olarak veri sınıflandırma ve tahmininde kullanılabilecek kuralları sağlar. Aşağıda bir Karar ağacı oluşturmaktan türetilen bir kural tanımına bir örnek verilmiştir: Eğer (dizüstü bilgisayar modeli x ise) ve (y tarafından üretilir) ve (z yaşındaysa) ve (bazı sahipleri k ise) o zaman (pil ömrü n saat).

Yakından izlendiğinde, bu kurallar basit, okunabilir ve anlaşılır formatlarda ifade edilir. Ek olarak, bu kurallar daha sonra başvurmak üzere bir veri deposunda saklanabilir.

Karar ağaçları, kökünden bir yaprağa kadar uzanan bir ağaç yapısını temsil ederek örnekleri sınıflandırır. En önemlisi, yüksek düzeyde bir Karar ağacının iki temsili vardır, bir düğüm ve düğümleri birbirine bağlayan bir yay. Bir karar vermek için, akış kök düğümlerinde başlar, bir yaprak düğümüne ulaşana kadar yaylara gider ve sonra bir karar verir. Ağacın her düğümü bir özneliğin testini belirtir ve dallar özneliğin alabileceği olası değerleri belirtir.

Bir Karar ağacı temsili özelliğinden bazıları şunlardır:

- Her yaprak olmayan düğüm (örneğin bir karar düğümü), öznelik değerinin temsiliğini belirtir.
- Her dal, değer temsiliğinin kalanını gösterir
- Her yaprak (veya terminal) düğümü, hedef niteliğinin değerini temsil eder
- Başlangıç düğümü kök düğümü olarak adlandırılır.

Karar Ağaçlarının Faydaları

Karar ağaçları kullanmanın avantajlarından bazıları burada listelenmiştir:

- Karar ağaçları hızlı ve kolaydır ve çok az deneme gerektirir
- Sağlamdır
- Anlaması ve yorumlaması kolaydır
- Karar ağaçları karmaşık veri hazırlığı gerektirmez
- Hem kategorik hem de sayısal verileri kullanabilirler. (Gollapudi ve Laxmikanth ,2016: 159)

2.3.3. K-En Yakın Komşu Algoritması (KNN)

K'-en yakın komşu parametrik olmayan bir sınıflandırma tekniğidir. Hem sınıflandırma hem de regresyon problemleri için kullanılabilir. Sınıflandırma durumunda, bilinmeyen bir örnek (q) verildiğinde, bir KNN sınıflandırıcısı, q'ye en yakın olan k eğitim örnekleri için kalıp uzayını arar. Bu k eğitim örnekleri, q'nun “en yakın komşuları”dır. Daha sonra q sınıfı basit oy çoğunluğu ile veya KNN'sinden uzak ağırlıklı oylama ile belirlenebilir. Regresyon durumunda, veri kümesindeki q örneklerine en yakın olan k örnekleri de bulunur. O zaman q değerini tahmin etmek için onların değerlerini kullanılmalıdır. Pek çok yaklaşım olsada, en basit yaklaşım komşuların değerlerinin ortalamasını almaktır. Öyleyse, KNN'nin iki adımı vardır; Birincisi en yakın komşuların belirlenmesi, ikincisi ise bu komşular kullanılarak sınıfın (sınıflandırmada) veya gerçek değer (regresyonda) belirlenmesidir. (Han ve Kamber 2001; Hardle 1989; Korhonen ve Kangas,1997).

KNN regresyon yöntemi, KNN sınıflandırıcısı ile yakından ilgilidir. K için bir değer ve bir tahmin noktası x_0 verildiğinde, KNN regresyonu, ilk önce x_0 ile temsil edilen x_0 'a en yakın olan K eğitim gözlemini tanımlar. Daha sonra x_0 'daki tüm eğitim cevaplarının ortalamasını kullanarak $f(x_0)$ değerini tahmin eder. Diğer bir deyişle, formülü şu şekildedir (James ve diğerleri, 2013: 59).

$$\hat{f}(x_0) = \frac{1}{K} \sum_{x_i \in N_0} y_i.$$

2.3.4. Lasso Regresyon

Lasso regresyonu Ridge regresyonuna benzer ancak deęişken seçim yapabilir. (Tibshirani, 1996: 267) Lasso regresyon, denklem'deki regresyon katsayılarının mutlak deęerlerinin toplamına dayanan bir kısıtlamaya maruz kalan karelerin kalıntı toplamını en aza indirir. (Padgett, 2008: 1869). Lasso Regresyonu řu řekilde tanımlanır:

$$\hat{\beta}_{lasso} = \underset{\beta}{argmin} \left\{ \sum_{i=1}^N \left(d_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\}$$

Lasso regresyonunda ceza $\sum_{j=1}^p |\beta_j|$ 'ye uygulanır ve Lasso cezası biçiminden dolayı Lasso regresyonunun deęişken seçimi ve regresyon katsayılarının büzölmesi; yani, en önemli deęişkenler korunurken, önemsiz deęişkenler modelden kaldırılır. Bias deęeri, Lasso regresyonunda Ridge regresyonuna kıyasla daha kontrol edilebilir (Zou, 2005: 301).

2.3.5. Çoklu Lineer Regresyon (MLR)

Çoklu doğrusal regresyon analizi, istatistiksel yöntemleri kullanarak bir dizi bağımlı ve bağımsız deęişken arasındaki ilişkiyi tanımlayan iyi bilinen bir yaklaşımdır. Bağımlı deęişken ile bağımsız deęişken sayısı arasındaki ilişkiler formundadır.

$$Y_i = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_kX_k + \epsilon_i$$

Bir “i” ardışık gözlem seti için, öngörülen Y deęişkeninin, “a” katsayıları ile eşleşen bir “a” dengesi, “a₀”, katsayıları ile eşleşen bir “k” tahmin deęişkenleri dizisi ve kalan hata ϵ . “a” deęerleri genellikle en küçük kareler yöntemi ile elde edilir. Regresyon denklemi prediktif modda kullanıldığında, ϵ (gerçek ve öngörülen deęerler arasındaki fark) ihmal edilir, çünkü beklenen deęeri sıfırdır. (Elaoud ve dięerleri, 2017: 442)

2.3.6. Polinomsal Regresyon (PR)

Polinom regresyonu, polinom fonksiyonunu kullanarak açıklayıcı değişken ile bağımlı değişken arasındaki ilişkiyi kurar.

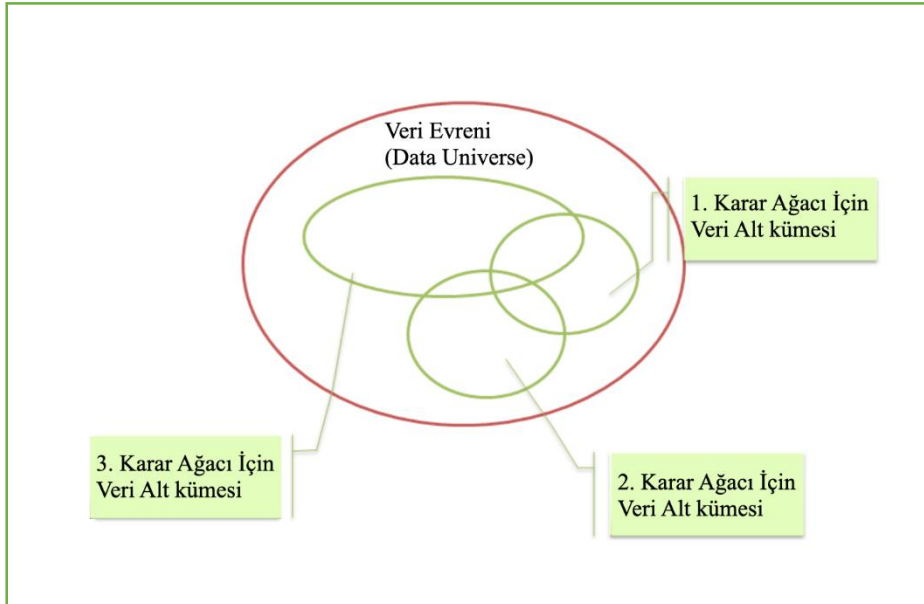
$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n + \epsilon$$

2.3.7. Rassal Orman (RF- Random Forest)

Bu özel ağaçlar çok fazla boyut olduğunda kullanılır. Boyutluluk sorunun temel önceliği, yüksek boyutlu verilerin karmaşıklığı getirmesidir.

Rastgele orman, daha fazla sayıda Karar ağacı ekleyerek Karar ağaçlarını genişletir. Bu Karar ağaçları, rastgele veri seçimi (örnekler) ve rastgele bir öznelilikler kümesi seçimi ile oluşturulmuştur. Aşağıdaki diyagram, karar ağaçlarının her birini oluşturmak için rastgele veri kümelerinin seçimini göstermektedir.

Şekil 2: Rassal Ormanın Yapısı



Kaynak: Gollapudi ve Laxmikanth, 2016.

Birden fazla karar ağacının yapılması için gereken diğer bir değişken girişi, buradaki şemada gösterilen özelliklerin rasgele altkümeleridir:

Şekil 3: Rassal Ormanın Ağaç Yapısı

	Değişken		
1	X1		Karar ağacı 1 için x1, x2, x3 değişkeni
2	X2		
3	X3		Karar ağacı 2 için x3, x4, x5 değişkeni
4	X4		
5	X5		
...	...		Karar ağacı 3 için Xa, Xb, Xn değişkeni
N	Xn		

Kaynak: Gollapudi ve Laxmikanth, 2016.

Her bir ağaç, rastgele veri kümesi ve rastgele değişken kümesi kullanılarak oluşturulduğundan, bu ağaçlara Rastgele ağaçlar denir. Ayrıca, bu tür Rastgele ağaçların çoğu Rastgele bir orman tanımlar. Rastgele ağacın sonucu iki radikal inanca dayanır. Birincisi, ağaçların her birinin verinin azami kısmı için doğru bir tahmin yapmasıdır. İkincisi, farklı yerlerde hatalarla karşılaşmaktadır. Böylece, bir karar vermek için karar ağaçları arasında bir sonuç anketi alınır. (Gollapudi ve Laxmikanth 2016:180)

2.3.8. Ridge Regresyon

Ridge regresyon yönteminde en küçük kareler yönteminde izlenen aşamalar birden fazla tekrarlanmaktadır. Ridge yönteminin en küçük karelerden farklılığı k^* ridge parametresinin varlığıdır. 0 ile 1 arasında değer olan her k^* için hesaplanan parametre kestirimleri arasında, aranan kriterlere sahip olanları belirlenir. Albayrak (2005)

Ridge şu şekilde formülize edilmektedir.

$$\hat{\beta}_R = (X'X + kI)^{-1}X'Y$$

Eşitlikte β^* ; $(p-1) \times 1$ boyutlu ridge regresyon katsayıları vektörü, I ; $(p-1) \times (p-1)$ boyutlu birim matris ve k sabit değer ve genelde $0 \leq k \leq 1$ aralığında bulunur. Ridge parametresi olan k değeri en küçük kareler yönteminden daha küçük hata kareler ortalaması elde edilmesini sağlar (İpek, 2001).

Ridge regresyon yönteminin kullanım yerleri şöyle sıralanabilir.

- * Çoklu regresyon modelinde bağımsız değişkenler birbirleri ile bağlantılı olduklarında E.K.K. β tahmin edicisinden daha küçük varyanslı β tahmin edicilerinin

elde edilmesinde,

- * Güçlü çoklu bağlantı etkisi ile regresyon katsayılarında oluşan kararsızlıkların grafik üzerinde gösterilmesinde,

- * Modeldeki gereksiz değişkenlerin çıkarılmasında (İmir, 1986) .

2.3.9. Destek Vektör Regresyon (SVR)

Destek vektör makineleri (SVM), genelleştirme hatasını en aza indiren Yapısal Risk Azaltma (SRM) ilkesini kullanarak Vapnik (1995) tarafından sunulan çekirdek bazlı öğrenme algoritmasıdır. Başlangıçta, sınıflandırma problemlerini çözmek için SVM geliştirilmiştir. Bununla birlikte, Vapnik'in "duyarsız" kayıp işlevinin devreye girmesiyle birlikte SVM, regresyon tahmini sorunlarını çözmek için genişletilmiştir. Bu, destek vektör regresyonu (SVR) olarak bilinir. SVR, bu özellik alanındaki doğrusal bir regresyon problemini çözmek için, lineer olmayan haritalama ile girdi verisini x daha yüksek boyutlu bir özellik alanına eşler. Bu dönüşüm bir çekirdek işlevi kullanılarak yapılabilir. Var olan en yaygın çekirdek işlevleri doğrusal çekirdek, polinom çekirdeği, Gauss (RBF) çekirdeği ve sigmoid (MLP) çekirdeğidir. Bu çalışmada Gaussian RBF çekirdeği kullanılmıştır, çünkü daha iyi tahmin performansı sağlar (Smola 1998). Bir SVR fonksiyonunun genel formu şöyle yazılabilir:

$$f(\mathbf{x}) = \mathbf{w} \cdot \phi(\mathbf{x}) + b$$

burada $\mathbf{w} \in \mathbb{R}^n$, $b \in \mathbb{R}$. $\mathbb{R}^n$ 'deki nokta ürününü belirtir ve $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 'den yüksek boyutlu alan $\mathbb{R}^n$ ye (yani, $n > m$) doğrusal olmayan bir dönüşümdür. Alan kısıtlamaları nedeniyle, SVR'nin ayrıntılı açıklamaları Burges (1998), Cristianini ve Shawe-Taylor (2000), Gun (1998), Smola (1998) 'da bulunabilir.

2.3.10. Facebook Prophet

Facebook Prophet modeli Facebook Temel Veri Bilimi tarafından geliştirilmiştir. (Taylor, S.J., ,2017). Prophet, doğrusal olmayan bir modele dayanarak zaman serilerini geleceğe yönelik tahminlemektedir. Prophet algoritması, büyümeyi, mevsimsellik ve tatilleri hesaplayarak y_t 'yi öngörür.

$$\hat{y}_t = g_t + s_t + h_t + \epsilon_t$$

Bunun yanı sıra g_t terimi, serinin nasıl büyüdüğünü ve devam eden büyüme için beklenen değerleri ve serilerin davranışlarına bağlı olarak mevsimsellik değişikliğini, tatil günlerinin etkilerini, hata terimini temel alarak hesaplayan büyüme fonksiyonudur.

2.3.11. ARMA

Tahmine yönelik istatistiksel yaklaşım, önceki gözlemleri kullanarak bir gözlemin değerini tahmin etmek için stokastik modellerin oluşturulmasını içerir. Bu genellikle rasgele girdilere sahip doğrusal stokastik fark denklem modelleri kullanılarak gerçekleştirilir. Bu modellerin bugüne kadarki en önemli sınıfı, üç boyutlu otoregresif hareketli ortalama (ARMA) modelleridir.

$$x_t = \gamma + \sum_{i=1}^q \phi_i x_{t-i} + \sum_{j=1}^q \theta_j e_{t-j} + e_t$$

E (e_t & $x_{t-1}, x_{t-2}, \dots$) = 0 olduğu varsayılır. Bu koşul, örneğin, sıfır ortalama, bağımsız ve özdeş bir şekilde dağılmış ve geçmiş x_t 'lerden bağımsız olduğunda tatmin

ediciştir. Bunun boyunca sonlu bir varyansın (nz) olduđu varsayılmıştır. Sıfır ortalama işlem için x_t kesigi y sıfırdır. İşaretlemeyi basitleştirmek için, x_t 'nin $x_t - E(z_t)$ ile değıştirilmesine eşdeğer olan $y = 0$ varsayılır.

2.2. MODEL PERFORMANS DEĞERLENDİRME YÖNTEMLERİ

2.2.1. Ortalama Mutlak Hata (MAE – Mean Absolute Error)

Ortalama mutlak hata (MAE) artıkların ortalamasıdır (yani tahmini hedef ile gerçek hedef sonuçları arasındaki fark). Bu önlemin alternatif formülasyonları, ağırlık faktörleri olarak nispi frekansları içerebilir. MAE, verilerin ölçüldüğü aynı ölçekte (yani birim) çalışır. Dolayısıyla sınırlama, farklı ölçekler / birimler kullanarak seri arasında karşılaştırma yapmanın mümkün olmamasıdır. Bu matematiksel olarak şu şekilde tanımlanabilir:

$$MAE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{l=0}^{n_{samples}-1} (y_l - \hat{y}_l)^2$$

- $\hat{y}_l$: Tahmini hedef çıktı
- y_l : Karşılık gelen (doğru) hedef çıktı
- $n_{samples}$: Örneklerin sayısı (Haroon, 2017:67).

2.2.2. Ortalama Kare Hatası (Mean Squared Error -MSE)

Ortalama kare hatası (MSE), MAE'ye benzer, ancak artık hedef karesi ile tahmini hedef çıktısı ve karşılık gelen gerçek çıktı arasındaki mutlak farkı almak yerine artık karenin alınmış olmasıdır. Değer her zaman pozitifdir ve 0'a yakın değerler daha iyidir. MSE'nin matematiksel tanımı şöyledir:

$$MSE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{l=0}^{n_{samples}-1} (y_l - \hat{y}_l)^2$$

$\hat{y}_i$: i. örneklerinin tahmini hedef çıktısı

y_i : Karşılık gelen (doğru) hedef çıktı

$n_{samples}$: Örneklerin sayısı (Haroon, 2017:67).

2.2.3. Kök Ortalama Kare Hatası (Root Mean Square Error-RMSE)

MSE'nin karekökü alınması, hedef çıktınıninkilerle aynı birimleri tutan kök ortalama karesi hatasını (RMSE) verir.

2.2.4. R^2

Belirleme katsayısı olarak da adlandırılan R^2 , gelecekteki kaç örneğin regresyon modeli tarafından tahmin edilebileceği ölçüsünü tanımlar. Dahası, aynı zamanda keşiflerden tahmin edilebilir olan yanıt değişkenindeki varyans oranını da gösterir. Değer 1'den eksi sonsuzluğa kadar değişmektedir, burada 1 en uygundur. Matematiksel temsil aşağıdaki gibidir.

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{l=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2}{\sum_{l=0}^{n_{samples}-1} (y_i - \bar{y})^2}$$

$\hat{y}_i$: Örneklerin tahmini hedef çıktısı

y_i : Karşılık gelen (doğru) hedef çıktı

$n_{samples}$: Örneklerin sayısı (Haroon , 2017:67)

$$\bar{y} = \frac{1}{n_{samples}} \sum_{l=0}^{n_{samples}-1} y_i$$

ÜÇÜNCÜ BÖLÜM

İŞ ZEKÂSI

3.1. İŞ ZEKÂSI KAVRAMI

İş dünyasında bulunan kurumlar bilişim teknolojilerinin gelişimiyle birlikte sürdürülebilir olması adına sahip oldukları verileri en üst düzeyde kullanımı ve maksimum fayda elde edilebilmeleri eğilimindedirler. Bu bağlamda ele aldığımızda iş zekâsı sistemleri ile kurumlar ellerinde bulunan verileri analiz edip etkin kararlar vererek kaliteli ve hızlı kararların verilmesini sağlamaktadır. Günümüzde artan bilgi hacmiyle karar verme adına iş zekâsı sistemlerinden faydalanılabilmektedir (Lin, Tsai, Shiang, Kuo, Tsai, 2009). İş zekâsı sisteminin nasıl olduğunu ve sistemden hangi faydalar elde edileceğini bilinmesi sistemin yürütülmesi adına çok önemli yer almaktadır.

Literatürde iş zekâsı adına yıllar içinde birçok araştırmacı tarafından tanımlar geliştirilmiştir. Gartner, 1989 yılında iş zekâsı terimi hakkında araştırma yapmıştır. Gartner'in araştırması içerisinde yer alan anahtar kavramları Dresner açıklamıştır. Dresner'e göre iş zekası "Kurumsal kullanıcıların daha iyi iş kararları vermelerine olanak tanıyan toplama, birleştirme, analiz etme ve veriye erişim sağlama çözümlerini sağlayan yazılımlar ve çözümlerdir" şeklinde tanımlamıştır (Gibson ve diğerleri, 2004). İş zekâsı kavramı ilk kez 1958 yılında Luhn tarafından IBM journal isimli dergide ortaya çıkmıştır. İş ve zekâ kavramlarını ayrı ayrı iki yapı olarak ele almıştır. Luhn'e göre iş "Bilim, teknoloji, ticaret, sanayi, hukuk, devlet, savunma gibi belirli bir amaç için sürdürülen faaliyetlerin toplamı" olarak ifade edilirken, zekâyı "İstenilen bir amaca yönelik harekete yönlendiren bir şekilde sunulan olguların karşılıklı ilişkilerini kavrama yeteneği" olarak ifade etmiştir (Chee ve diğerleri, 2009). İki kavramı bir araya getirildiğinde iş zekâsı bilgiyi yaymak için kullanılan otomatikleştirilmiş bir sistem olduğu vurgulanmaktadır. Kullanılan veri tabanlarını işle alakalı olarak araç olarak kullanılmasını öngörmüştür (Agrawal, 2008). Başka bir araştırmaya göre iş zekâsı kurumların karar verme becerilerine yardımcı olarak verileri bilgiye dönüştürüp rekabet sırasında avantaja dönüşmesini sağlayan kavram, süreç ve yazılımların bütünüdür (Brohman ve diğerleri, 2000). Lin (2009)'e göre iş zekâsı kurumlar

tarafından biriktirilen yapılandırılmış ve yapılandırılmamış tüm verilerin analizini yaparak avantaj elde edilen araçlardır. Uygulanan bu araçlar iş sürecinde toplanılan büyük verileri anlamlı bir hale sokarak kurumların ya da kişilerin daha net kararlar almasını sağlar (Lin vd., 2009).

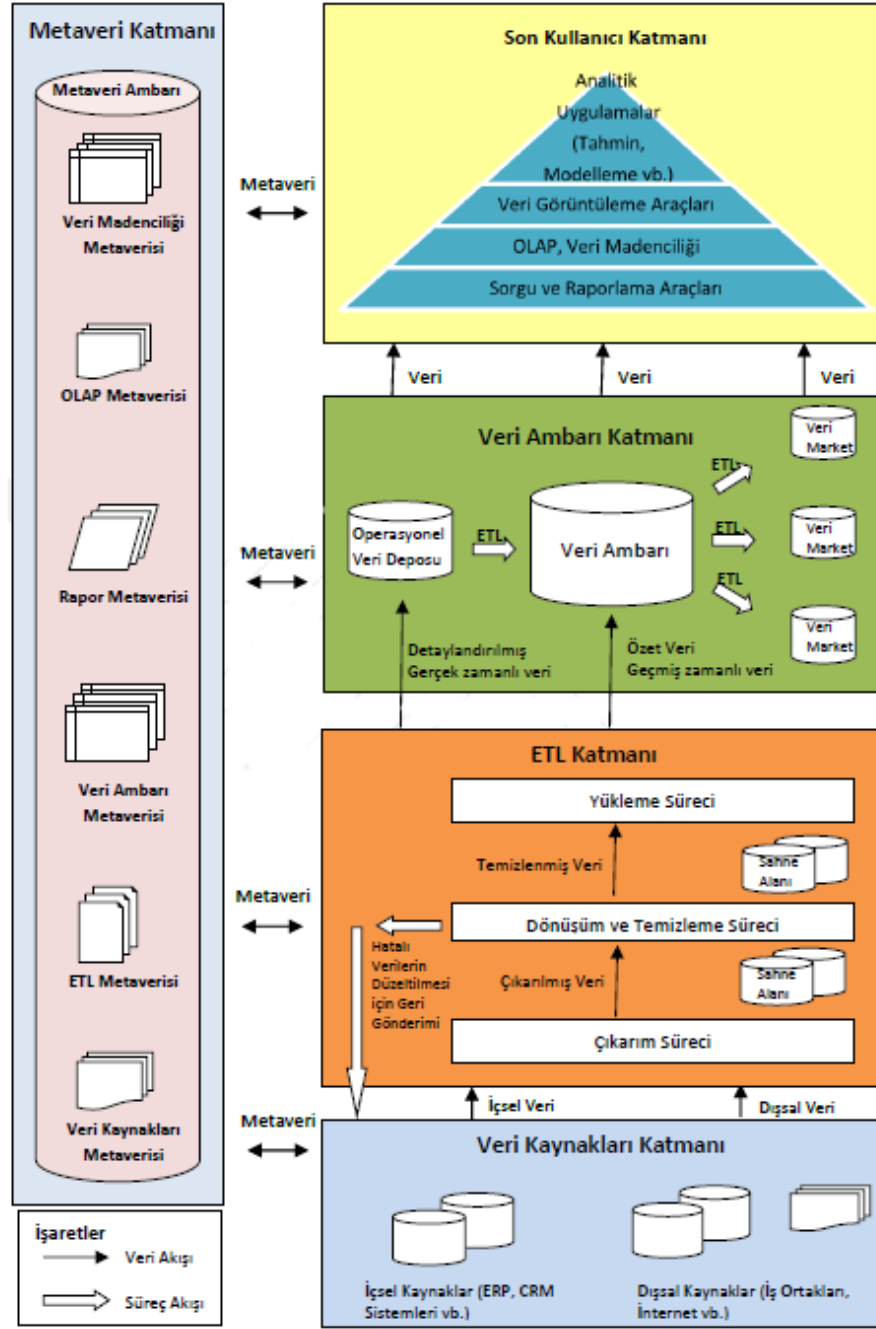
Genel olarak iş zekâsı uygulamaları verilerin işletmeler tarafından sorgulama ve verilerin çok boyutlu analizlerinin yapılmasını sağlar (Baars, Kemper, Siegel, 2008).

3.2. İŞ ZEKÂSİ SİSTEMLERİNİN MİMARİ YAPISI

İş zekâsı sistemlerini ele aldığımızda verileri toplama, analiz etme ve bu verilerden raporlar sunarak kurumların kararlarını belirlemede etkin bir role sahip olduğunu gözlemlemekteyiz. Bu bağlamda iş zekâsı sistemlerinin karar verme aşamasına kadar geçirdiği tüm işlemlerin ve süreçlerin bir bakıma iş zekâsı sistemlerinin mimarisini oluşturmaktadır (Kalelkar, Churi, Kalelkar, 2014). İş zekâsının toplanan veriler ile yaptığı işlemleri incelediğimizde verilerin değerli bilgiler haline getirmesiyle kendine has bir özelliği olduğu görülmektedir. Bu bağlamda iş zekâsı sistemleri mimari yapısı önemli bir role sahiptir.

İş zekâsı sistemlerinin mimari yapısı genel hatları itibarıyla açıklanmıştır (Şekil 1). Bu mimari yapı metaveri katmanı, veri kaynakları katmanı, ETL katmanı, veri ambarı katmanı ve son kullanıcı katmanı olmak üzere birbirleriyle ilişkili beş ayrı yapıdan oluşmaktadır.

Şekil 4: İş Zekası Sistemleri Mimari Yapısı



Kaynak: Ong, Siew ve Wong: 2011:1

Şekil 4’de görüldüğü üzere iş zekâsı sistemlerinin mimari yapılarını oluşturan katmanları kısaca şu şekilde tanımlayabiliriz:

- **Veri Kaynakları Katmanı**

Veri kaynağı iş zekası uygulamalarının temelini oluşturan bileşendir. Organizasyonlar için verilerin kaynakları çok önemlidir. Şirketlerin birçoğu operasyon

işlemlerinde ERP, SCM, satış odaklı işlemler için CRM ya da perakende satış işlemleri için POS gibi yazılımlar kullanırlar. Bu sistemler veri kaynağını üretir (Loshin, 2003). Veriler farklı kaynaklardan toplanırlar. İçsel kaynaklar organizasyon içerisindeki operasyonel verilerin sağlandığı sistemlerdir. Dışsal kaynaklar ise genellikle kuruluş dışı olan verilerin elde edildiği platformlardır. Bu kaynaklardaki veriler de genellikle internet, iş ortakları, veri sağlayıcılar, devlet veya market araştırmaları verileri gibi verilerden oluşmaktadır. Veri kaynaklarından elde edilen verilerin bir kısmı yapılandırılmış veriler olabilirken bir kısmı ise yarı yapılandırılmış veya yapılandırılmamış verilerden oluşabilmektedir (Loshin, 2003).

- **ETL Katmanı**

ETL katmanı üç ana sürecin bir araya gelip işlemesiyle ortaya çıkmaktadır. Bu süreçler; çıkarım, dönüşüm ve yüklemeden meydana gelmektedir. İçsel ve dışsal kaynaklardan toplanan veriler belirli bir hedefi gerçekleştirmek için istenilen verilerin bir araya getirilip belirli bir analiz sürecinden geçtikten sonra yeni bir forma kavuşmaktadırlar (Ishaya ve Folarin, 2012). Diğer adı veri entegrasyonu olarak geçen bu katmanda ilk olarak verilerin çıkarım işlemleri gerçekleşir. Çıkarım işleminde elde edilen ham verilerin istenilen kadar alınır. Bu veriler genellikle entegre edilmemiş, eksik veya tekrarlı veriler olabilmektedir. Çıkarılan veriler dönüşüm ve temizleme işlemlerini gerçekleştirmeden önce sahne alanı ismi verilen bir depolama alanına gönderilerek burada geçici olarak tutulurlar (Ong ve diğerleri, 2011). ETL sürecinin en temel görevi; gereksiz bilgilerin elenmesi ve gerekli bilginin kaliteli olarak depolanmasını sağlamaktır (Pazarçeviren ve diğerleri, 2015).

- **Veri Ambarı Katmanı**

Veri ambarı da iş zekâsı sistemleri için önemlidir. Veri ambarları ihtiyacı olunan verileri daha iyi analiz etme şansına sahiptirler. Veri ambarları iş zekâsı sistemleri dahilinde kurumların belirli bir alanda karar verme ihtiyacını karşılayacak olan verilerini destekleyecek yapıya sahiptir. (Ranjan, 2009). Veri ambarları kurumlar için verilere kolay ulaşabilmeyi sağlarlar (Inmon, 2005). Örneğin bir üreticinin siparişler, ürünler, tedarikçiler veya malzeme verileri veri ambarında belirli bir veri kümesini meydana getirerek verileri özelleştirme imkânı sağlayarak yöneticilere karar verme şansı vermektedir. Bu bakımdan ele aldığımızda veri ambarları iş zekâsı sistemlerinin olmazsa olmazları arasındadır.

- **Son Kullanıcı Katmanı**

Son kullanıcı katmanı ise iş zekâsı sistemlerinin en üst katmanıdır. Bu aşamada kullanıcı ile sistem arasında verilerin belirli formda görülmesi sağlanmaktadır. Kalelkar ve diğerlerinin (2014)'nın yaptığı çalışmada bu aşamadaki son kullanıcı araçlarını “son kullanıcılara bilgi görüntüleyen programlardır. Son kullanıcı araçları, farklı kullanıcı türleri için aynı bilgilerin birden çok görünümünü sağlayabilir.” şeklinde tanımlamaktadırlar. Son kullanıcıların bu aşamada verilerin daha derin analiz edebilecek şekilde donatılmış raporlar, tablolar ve grafikleri görüntülemesini sağladıklarını belirtmektedir (Chan ve diğerleri, 2011:1). Dolayısıyla bu aşamada kullanıcılar ve veriler arasında önemli bir ağ kurulduğu belirtilmektedir. Son kullanıcı katmanında kullanılan araçlar sorgu ve raporlama araçları, OLAP, veri madenciliği, veri görüntüleme araçları, analitik uygulamalardır.

- **Metaveri Katmanı**

İş hayatında iş zekâsı sistemleri çerçevesi içerisinde yöneticilerin doğru kararlar alabilmesi adına yeterli ve etkin verilere sahip olmaları gerekmektedir. Bu durumda verilerin işletmeler için önemi çok büyüktür. İş zekâsı sistemlerinden anlamlı veriler elde ederken mimaride kullanılan metaveri (metadata) kavramı da mimari yapı için vazgeçilmez bir unsurdur. Metaveri esasında verilerden elde edilen veriler olarak bilinmektedir (Chee, Yeoh ve Gao, 2011:16). Metaveriler gerçekte olan verilerin nerede kullanıldığını, hangi veri kaynağına sahip olduğunu, veri üzerinde hangi değişikliklerin yapıldığını ve hangi bilgilerle ilişkili olduğu gibi verileri tutabilmektedir. Metaveriler iyi tutulduğu zaman tekrardan bilgi üretilmesi gerekirse veriler yeniden ortaya atılmaktadır.

İş zekâsı sistemleri mimari yapısında da son kullanıcı katmanı, veri ambarı katmanı, ETL katmanı, veri kaynakları katmanlarında oluşan metaveriler depolarda saklanmaktadır (Ong, 2011:1).

3.3. İŞ ZEKÂSI UYGULAMA ORTAMLARI

İş zekâsı uygulama yazılımları değişik ortamlarda sunulabilmektedir. Bu ortamlar mobil ortamlar ve masaüstü yazılımlardır. Genel olarak bir iş zekâsı

uygulamasının aktif raporlamasında (dashboard) aranan temel özellikler aşağıdaki şekilde sıralanabilir:

Kullanılabilirlik: Kişilerin ihtiyacı olan özelliği kendi başlarına bulmalarının kolay olmasıdır.

İçerik Zenginliği: Basit bir rapor alınmasının ötesinde kişilere tavsiyelerde bulunan yardım menüleri ve destek dokümanları ile yönlendirilmesi beklenmektedir.

Sadelik: Karmaşık ekranın kişiler için zorlaşmaması gereklidir. Sistemde gerek kişilerin kendi ekranlarını tasarlamaları gerekse kendi kontrollerini koyabilmesi beklenmektedir.

Güncellik: İş zekâsı portalinin devamlı güncel olması gerekmektedir.

İş zekâsı: yönetsel kararların doğru ve hızlı alınabilmesi için verilerin depolanmasını, verilere kolay ulaşılabilmesini, verilerin analiz edilmesini, verilerle planlama yapılabilmesini, verilerle performans değerlendirmesinin yapılabilmesini sağlayan teknoloji ve uygulamaların bütünüdür (Köklü, 2018:121).

3.4. MAKİNE ÖĞRENMESİ VE İŞ ZEKASI İLİŞKİSİ

Makine öğrenmesi modelleri, verilerin içindeki gizli örüntüleri ortaya çıkarmak için çok etkili çözümler sunmaktadır. Veri uzmanları bu teknikleri uzun yıllardır karmaşık ve uzman iş problemlerini çözmek için kullanmaktadırlar. Son yıllardaki işlem gücündeki gelişmeler, bu karmaşık matematiksel modellerin geliştirilmesini ve yürütülmesini daha erişilebilir hale getirmiştir. Pahalı ve ileri teknoloji donanım gerektiren modeller artık herkesin kullanabileceği platformlarında çalışabilmektedir.

Günümüzde makine öğrenmesi iş zekasını beslemekte ve iş zekasının gücünü artırmak gibi önemli bir işleve sahiptir. Her ne kadar bu iki kavram birbirleriyle ilişkili olsada, ikisi arasında belli başlı farklılıklar bulunmaktadır. Tablo 1 de bu farklılıklara değinilmiştir.

Tablo 1: İş Zekası ve Makine Öğrenmesi Farkları

Özellik	İş Zekası	Makine Öğrenmesi
Yapılan İş	İşletmeyi istenen yola ulaştırabilmek için yapılan metodolojik işlemleri kapsar.	Makinenin mevcut verilerden öğrenmesini sağlar
Teknoloji	İş fırsatlarını tanımlar.	Veri merkezli öğrenme ve karar verme sistemleri geliştirilir.
Veri İşleme	Ham verileri yararlı bilgilere dönüştürür.	Tahminleme amaçlı modeller geliştirmek için veri madenciliği teknikleri uygular.
Algoritmaların Kullanımı	Bir algoritmaya bağımlı değildir ve beceriye dayanır	Büyük ölçüde algoritmalarından yararlanır.

Kaynak: Educba, 2019.

3.5. İŞ ANALİTİĞİ VE İŞ ZEKÂSI İLİŞKİSİ

İş zekası ve iş analitiği kavramları her ne kadar iç içe geçmiş gibi dursa da, karıştırılmakta ve çoğu zaman birbirlerinin yerine kullanılarak anlam karmaşası yaratmaktadır. Bu bağlamda ikisinin farkının ortaya konulması, yapılacak çalışmalar için büyük önem taşımaktadır. İş analitiğinin (Business Analytics- BA) iş zekası yaklaşımından en temel farkı, iş analitiğinin problemleri çok daha geniş bir perspektifte ele alması ve bunun yanında yapısal kararlarla birlikte yapısal olmayan kararları da problemin çözüm sürecine dahil etmesidir. İş zekası uygulamaları daha çok, belli bir metodoloji kullanan ve veri bilimi gibi yaklaşımlardan yararlanan uygulamalar olarak nitelendirilebilir; fakat iş analitiği daha geniş kapsamda bir metni, ses dosyasını ya da daha farklı verileri analiz sürecine katabilmektedir. Özetle, İş zekası ve iş analitiği kavramları işlenen veri/bilgi kaynakları bakımından birbirlerinden ayrılmaktadırlar. Ayrıca, iki yaklaşımın çıktıları ele alındığında, çıktılar arasında da farklılıklar bulunmaktadır. İş zekasının hedeflediği çıktılar genellikle planlanmış ve sürdürülebilir çıktılardır (Şeker, 2013:605). Örneğin, bir işletme için

belirli amaca yönelik oluşturulan yönetim paneli (dashboard) iş zekası yaklaşımının bir çıktısı olarak ifade edilebilmektedir. İş analitiği ise tüm bu çıktıları kapsamasının yanında, karar sürecini de tüm bu sistemin içine dahil edebilmektedir.

İş analitiği uygulamaları temelde yoğun biçimde istatistiksel modellerden yararlanmakta ve veri bilimi çalışmalarının bir sonucu olan araçları kullanmaktadır. Bu bağlamda, tanımlayıcı analitik yöntemler (descriptive analytics) kullanılarak verinin analiz süreci ve buna bağlı farklı amaçlara yönelik raporların hazırlanması veya tahminleyici analitik yöntemler (predictive analytics) kullanılarak bir karar destek sisteminin geliştirilmesi iş analitiği uygulamalarına örnek olarak gösterilebilir.

Sonuç olarak, iş zekası daha çok sorgulama, raporlama ve çevrimiçi analitik işleme (OLAP) gibi konular üzerine yoğunlaşırken, iş analitiği uygulamaları kararların otomatize edilmesi aşamasında farklı sektörlerde kullanılmaktadır. Özetle, iş analitiği uygulamaları daha geniş araçlar sunmakta ve yeni fırsatlar ortaya çıkarmaktadır. İş zekası ve iş analitiği kavramları arasındaki farklar Tablo 2 de özetlenmektedir:

Tablo 2: İş Zekası ve İş Analitiği Farkı

İş Zekasının İlgilendiği Sorunlar	İş Analitiğinin İlgilendiği Sorunlar
• Ne oldu? • Kaç tane? • Olayların tekrar etme sıklığı nedir? • Hangi sistemler bu olaylardan ne kadar etkilenmektedir? • Maliyeti nedir? • Gelirler nelerdir? • Problem nerede oluyor? • Neler yapılması gerekir?	• Bu olay neden oluyor? • Sebebi Nedir? • Bu gidiş devam edere ne olur? Bu olaydan sonra nasıl bir devam bekleyebiliriz? • Bu durumda alınabilecek en iyi pozisyon (aksiyon) nedir?

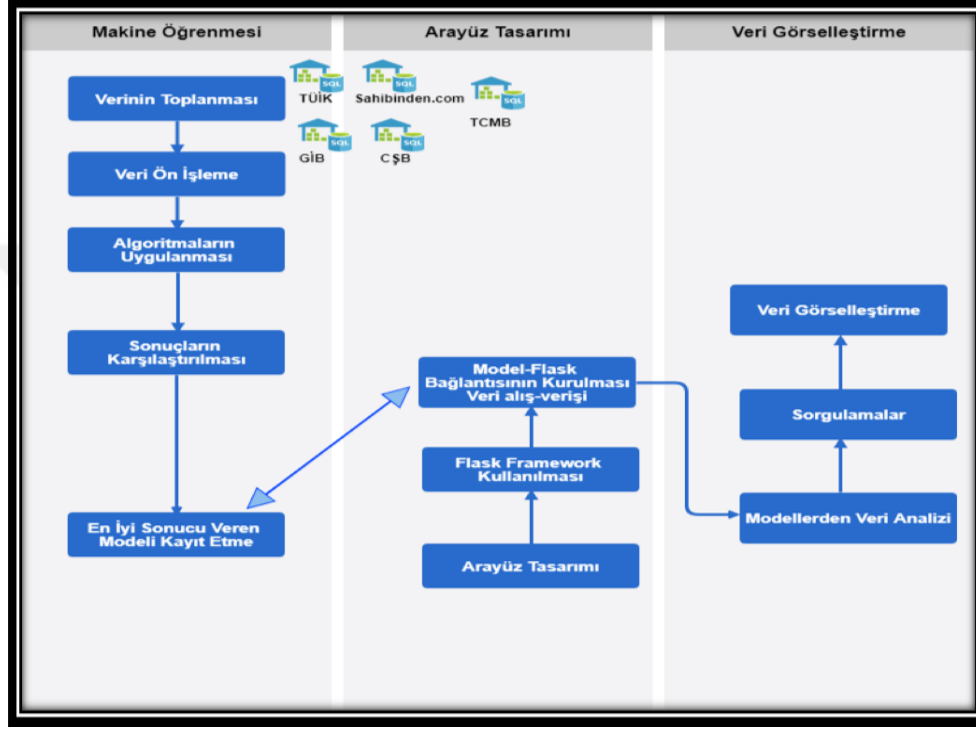
Kaynak: Şeker, 2016: 2.

DÖRDÜNCÜ BÖLÜM

İŞ ZEKASI TASARIMI VE UYGULAMASI

Çalışmada, uygulama geliştirilirken Şekil 5'deki iş akışının adımları takip edilmiştir.

Şekil 5: Uygulama İş Akışı



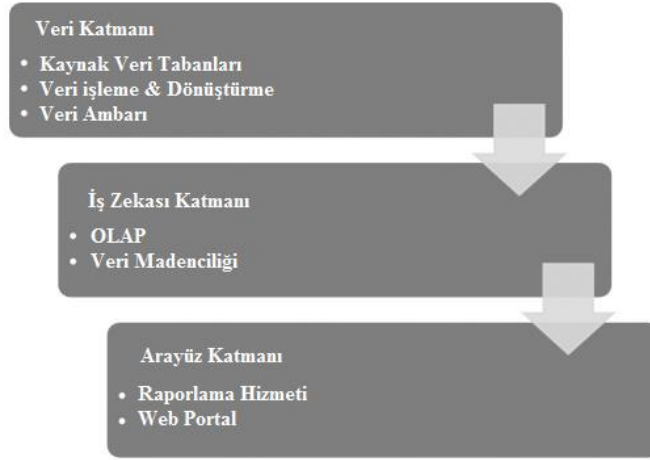
Kaynak: Yazar tarafından derlenmiştir.

Uygulamanın geliştirilmesinde ilk adım olarak makine öğrenmesi teknikleri uygulanmıştır. Süreç olarak CRISP-DM (Cross-industry Standard Process for Data Mining) adımları izlenmiştir.

4.1. METODOLOJİ

Uygulama kapsamında üç farklı yöntem kullanılmıştır. İlk kullanılan model Şekil 6 da açıklanmaktadır.

Şekil 6: İş Zekası Geliştirme Modeli



Kaynak: Deveci Kocakoç ve Erdem, 2010: 171.

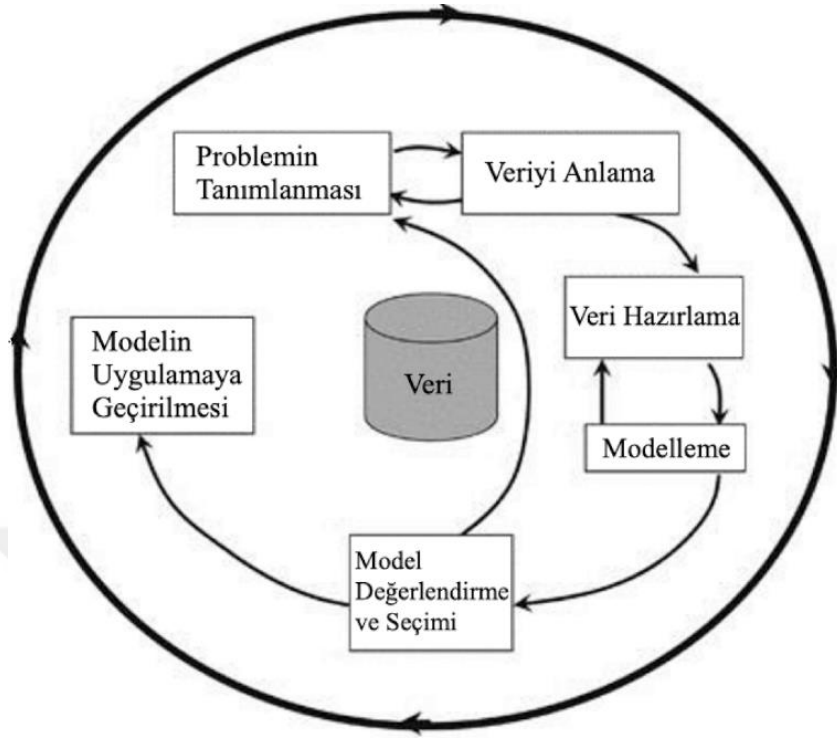
Veri katmanında kaynak veritabanları olarak online ilan siteleri, TÜİK, Çevre ve Şehircilik Bakanlığı, sektöre yönelik raporlar baz alınmıştır. Bu kaynaklardan elde edilen veriler işlenerek iş zekası uygulaması için veri ambarı haline getirilmiştir. İş zekası katmanında, çapraz sorgulama ve makine öğrenmesi teknikleri uygulanmıştır. Arayüz katmanında ise, Python frameworkü Flask ile HTML tabanlı bir raporlama arayüzü geliştirilmiştir. İkinci olarak CRISP-DM Modelinden yararlanılmıştır. CRISP-DM modelinin adımları, makine öğrenmesi temelinde uygulanmıştır.

CRISP-DM Modeli:

Veri Madenciliği için Çapraz Endüstri Süreç Modeli (Cross-Industry Standard Process for Data Mining – CRISP-DM) (Shearer ,2000)

CRISP DM Modeli Şekil 7’de şematize edilmektedir.

Şekil 7: CRISP-DM Modeli



Kaynak: (Kartal, 2015; Shearer , 2000)

CRISP DM modelinin altı aşaması ve bunlara uygulama kapsamında nasıl yer verildiği aşağıdaki gibi listelenmektedir:

- Problemin Tanımlanması (Business Understanding)

Problemin tanımlanması aşamasında; öncelikle amaç belirlenmiştir. Buna ek olarak, “Çözülmek istenen problem nedir?”, “Makineden öğrenmesi istenen şey nedir?”, “Görevi nedir?” sorularına yanıt aranmıştır ve bu bağlamda veri setlerinde emlak sektörüne yönelik, sonuca etki edebilecek nitelikler araştırılmıştır.

- Veriyi anlama (Data Understanding)

Veriyi anlama aşamasında, veri setindeki nitelikler, bu niteliklerin açıklamaları ve veri tipleri açıkça ifade edilmiştir.

- Veri Hazırlama (Data Preparation)

Veri hazırlama aşamasında, veri seti ayrıntılı bir biçimde analiz edilip, aykırı değerler ve tekrar eden gözlemler tespit edilmiş ve ayıklanmıştır. Ayrıca, kayıp değerler belirlenip özellik (feature) lere uygun olarak tamamlanmıştır ve logaritma yöntemi ile normalizasyon gerçekleştirilmiştir.

- Model Kurma (veya Modelleme) (Modeling)

Model kurma aşamasında, veri seti genellikle tahminlemeye yönelik olduğundan, farklı regresyon algoritmaları denenmiştir. Bunun yanı sıra, aynı öğrenme algoritmalarına yönelik farklı modellerden de yararlanılmıştır.

- Model Değerlendirme ve Seçimi (Evaluation)

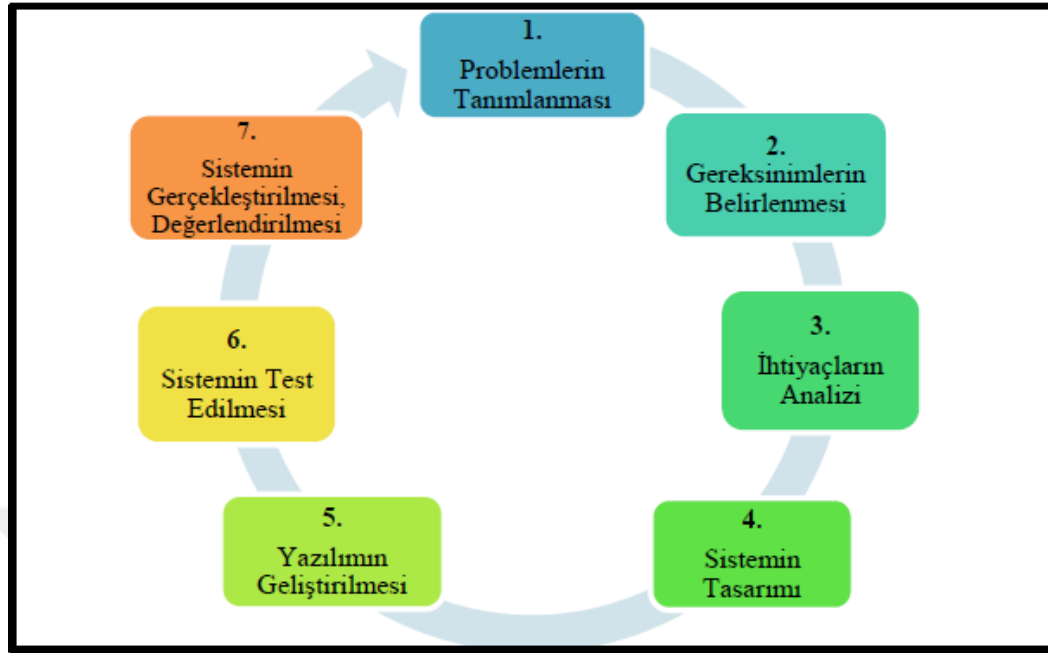
Model değerlendirme ve seçimi aşamasında, k-kat çapraz geçerleme (k-fold cross validation) uygulanıp, k katsayısı olarak literatürde sıkça yer verilen 5 ve 10 denenmiştir.

- Modelin Uygulamaya Geçirilmesi (Deployment)

Modelin uygulamaya geçirilmesi aşamasında, problemin çözümü için en iyi performans gösteren model seçilmiş olup, oluşturulacak iş zekası uygulamasına entegre için hazır hale getirilmiştir.

İş zekası uygulaması geliştirilirken, uygulamanın her aşamasında Sistem Geliştirme Yaşam Döngüsü (SGYD) adımları takip edilmiştir. 7 aşamalı bir yaklaşım izlenmiştir. Buna göre, birinci ve ikinci aşamalarda, yani planlama ve analizde sistem ihtiyaçları ve amaçlar belirlenmiştir. Aynı zamanda hedefler ve kilometre taşları belirtilmiştir. Tasarım ve uygulama sürecinde, grafik kullanıcı arayüzleri (Graphical User Interface- GUI) tasarlanmış ve sistemin kodlaması yapılmıştır. Test aşamasında, yazılım geliştirme aşamasında oluşturulan kod, kullanıcı ve sistem testlerinden geçirilmiştir. Son olarak sistem, canlı ortama taşınmıştır.

Şekil 8: Sistem Geliştirme Yaşam Döngüsü



Kaynak: Tecim, 2015: 3.

Problemlerin, fırsatların ve amaçların tanımlanması aşamasında, geliştirmek amaçlanan iş zekası uygulamasının arayüzünün operasyonel, taktiksel ve stratejik bağlamda sektöre yönelik iş akışları incelenmiştir. İşlevsellik ve eksikler belirlenerek yazılımın kapsamı tahminlenmiştir.

Bu sürecin sonunda yöneticilerin ve kullanıcıların ihtiyaç duyabilecekleri bilgi gereksinimleri ortaya çıkartılmıştır. Bunun için özellikle karar vericilerin hangi bilgilere gereksinim duyduğu, mevcut sistemin analizinden yararlanılarak detaylı bir biçimde ortaya konmuştur. Bu aşamada iş zekası amaçlı kullanılan sistemlerdeki bilgi akışı incelenmiştir ve buna yönelik olarak ek geliştirmeler yapılmıştır.

Sistem ihtiyaçlarının analizi aşamasında, kullanıcılar için farklı seviyelerde yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış verilere ihtiyaç duyulmuştur. Bu veriler ile elde edilmek istenen bilgiler analiz edilmiş ve buna yönelik olarak sistem tasarımı adımına geçilmiştir.

Sistem tasarımında geliştirilecek iş zekası uygulaması için farklı veri ambarlarından gelecek verilerin hatasız bir şekilde uygulamaya entegre edilmesi prosedürü tasarlanmıştır. Kullanıcı dostu ve mobil uyumlu bir arayüz tasarımı ile tüm cihazlardan erişim kolaylığı amaçlanmıştır.

Yazılımının geliştirilme aşamasında, kullanılabilecek en etkin yazılım dili Python olarak belirlenmiştir. Ayrıca ihtiyaç duyulan kütüphaneler ve frameworkler sisteme dahil edilip geliştirme safhası tanımlanmıştır.

Sistem geliştirilirken iki farklı sistem testi yapılması düşünülmüştür. Birincisi, uygulamanın doğruluk ve tutarlılık testidir. İkincisi ise geliştirilen uygulamanın test kullanıcıları tarafından kullanılıp geri bildirim alınması planlanmıştır.

Sistemin gerçekleştirilmesi ve değerlendirilmesi aşamasında, sistemin tamamlanıp son kullanıcıya ulaşması hedeflenmektedir. Bu aşamadan sonra sistemin işlevselliği ve avantajları üzerine hedef kitleden geribildirim alınması düşünülmektedir.

4.2. VERİLERİN TOPLANMASI

Uygulama amacıyla seçilen bölgeye yönelik etki edebilecek veriler uzman görüşleri ve literatür yardımıyla belirlenmiş olup ilgili veriler farklı veri kaynaklarından elde edilmiştir. Veri toplama sürecinde iki yöntem uygulanmıştır. İlk yöntemde internette yer alan emlak ilanlarını ve onlara ait veriler geliştirilen bir Hypertext Preprocessor (PHP) uygulaması ile elde edilmiştir. İkinci yöntemde emlak sektörüne ilişkin faktörlere yönelik veri paylaşımı yapan Türkiye İstatistik Kurumu (TÜİK), Türkiye Cumhuriyeti Merkez Bankası (TCMB), Çevre ve Şehircilik Bakanlığı ve Gelir İdaresi Başkanlığı (GİB) tarafından sağlanan veriler, veri seti halinde elde edilmiştir. Toplamda makine öğrenmesi teknikleri uygulanacak 13 adet veri seti elde edilmiştir ve edilen veri setlerine ait açıklamalar Tablo 3 de görülmektedir.

Tablo 3: Veri Setleri ve Açıklamaları

Veri Seti	Açıklama
İlan Verileri	İki aylık süreçte online yayınlanan ilan verileri
İnşaat Ciro Endeksi	TÜİK tarafından yayınlanan inşaat sektörüne ait aylık ciro endeksleri
İnşaat Maliyet Endeksi	TÜİK tarafından yayınlanan inşaat sektörüne ait maliyet endeksleri
Buca Satılan Konut Sayıları	TÜİK tarafından yayınlanan İzmir ili Buca ilçesi aylık satılan konut sayıları
İnşaat Güven Endeksi	TÜİK tarafından yayınlanan inşaat sektörüne yönelik aylık güven endeksleri
Konut Faiz Oranları	TCMB tarafından sağlanan ağırlıklı ortalama haftalık konut kredisi faiz oranları
İzmir TÜFE	TÜİK tarafından yayınlanan İzmir iline ait Tüketici Fiyat Endeksi
USD/TRY Verileri	TCMB tarafından yayınlanan günlük USD/TRY gün sonu kapanış fiyatları
Bina Yapım Maliyetleri	GİB tarafından sağlanan yıllık bazda bina metrekare yapım maliyetleri
Adatepe Mahallesi Metrekare	İzmir ili Buca ilçesi Adatepe mahallesine ilişkin aylık olarak ortalama metrekare fiyatları
Akıncılar Mahallesi Metrekare	İzmir ili Buca ilçesi Akıncılar mahallesine ilişkin aylık olarak ortalama metrekare fiyatları
Atatürk Mahallesi Metrekare	İzmir ili Buca ilçesi Atatürk mahallesine ilişkin aylık olarak ortalama metrekare fiyatları
Barış Mahallesi Metrekare	İzmir ili Buca ilçesi Barış mahallesine ilişkin aylık olarak ortalama metrekare fiyatları

Kaynak: Yazar tarafından derlenmiştir.

İlan verileri hariç diğer veriler 2015 baz yılı alınarak elde edilmiştir.

4.3. PROBLEMİN TANIMLANMASI

Bu adımda makineden öğrenmesi istenilen şey nedir? (Görevi nedir?) soruları sorulmuştur. Makineden yapılması istenilen şey elde edilen veri setleri içerisinde belirlenen hedef parametrenin tahminlenmesi amaçlanmıştır. İlk olarak veri setleri oluşturulmuştur. Veri setleri incelenmiş bunun sonucunda istenilen sonuca etki edebilecek nitelikler araştırılmıştır. Makine öğrenmesi algoritmaları için veri setlerinin yapıları oluşturulmuştur. Öğrenme stratejisi kararı verilmiş ve veri setlerine göre uygulanacak algoritmalar ve algoritmalara ait parametreler belirlenmiştir. Elde edilen sonuçlar karşılaştırılarak en başarılı sonucu veren model kayıt edilmiştir.

4.4. VERİYİ ANLAMA

Veriyi anlama aşamasında, veri setlerine ait özellikler, gözlemler, eksik veriler, uç noktalar vb. incelenmiştir.

İlan verilerinde;

- Şekil 9 da görüldüğü gibi 21 özellik ve 6634 satır yer almaktadır. Fiyat hedef özelliğdir ve regresyon işlemi sonucu tahmin edilecek değerleri tutmaktadır.
- Şekil 11 de görüldüğü gibi veri setinde eksik bilgiler mevcuttur.
- Şekil 9 da görüldüğü gibi veri setinde metin değerler vardır.
- Her bir satır bir eve ait özellikleri ifade etmektedir.

Diğer veri setlerinde;

- Şekil 10 da görüldüğü gibi özellikler, tarih-değer sıralamasında olup iki özellik mevcuttur. Değer hedef özelliğdir zaman serisi algoritmaları sonucu geleceğe yönelik tahmin edilecek değerleri tutmaktadır.
- Veri setlerinde eksik bilgi ve değer bulunmamaktadır.

Şekil 9: İlan Verilerine Ait Genel Bilgi

```
Data columns (total 21 columns):
baslik          6635 non-null object
fiyat           6634 non-null object
no              6634 non-null float64
tarih           6634 non-null object
m2net           6627 non-null float64
m2burut         6634 non-null float64
odasayisi       6634 non-null object
binayasi        3600 non-null object
bulundugukat    6634 non-null object
katsayisi       6634 non-null float64
isitma          6634 non-null object
banyo           6634 non-null object
balkon          6634 non-null object
site            6634 non-null object
kredi           6634 non-null object
kimden          6634 non-null object
konum           6634 non-null object
aciklama        6634 non-null object
il              6634 non-null object
ilce            6634 non-null object
mahalle         6634 non-null object
dtypes: float64(4), object(17)
```

Kaynak: Yazar tarafından derlenmiştir.

Şekil 10: USD/TRY Veri Setine Ait Genel Bilgi

```
Data columns (total 2 columns):
Tarih          1097 non-null object
Fiyat          1097 non-null float64
dtypes: float64(1), object(1)
```

Kaynak: Yazar tarafından derlenmiştir.

Şekil 11: İlan Verilerine Yönelik Eksik Veriler

```
binayasi      3035
m2net         8
mahalle       1
katsayisi     1
fiyat         1
no            1
tarih         1
m2burut       1
odasayisi     1
bulundugukat  1
isitma        1
ilce          1
banyo         1
balkon        1
site          1
kredi         1
kimden        1
konum         1
aciklama      1
il            1
baslik        0
dtype: int64
```

Kaynak:Yazar tarafından derlenmiştir

Tahminleme algoritmalarının kullanıldığı veri setine ait özellikler Tablo 4’te açıklanmaktadır.

Tablo 4: Tahminleme Algoritmalarında Kullanılan Veri Setinin Özellikleri ve Açıklamaları

Özellik (Features)	Açıklama
Fiyat	Konuta ait ilanda yer alan fiyat bilgisi
Tarih	Konutun ilana verildiği tarih bilgisi
M2net	Konutun sahip olduğu net metrekare alanı
M2brut	Konutun sahip olduğu brüt metrekare alanı
Odasayisi	Konutta bulunan oda sayısı
Binayasi	Konutun bulunduğu binanın yaşı
Bulundugukat	Konutun binadaki bulunduğu katı
Katsayisi	Konutun yer aldığı binanın toplam kat sayısı

İsıtma	Konutta yer alan ısıtma sistemi
Banyo	Konutta yer alan banyo sayısı
Balkon	Konutta yer alan balkon sayısı
Site	Konutun bir site içinde yer alıp almadığı bilgisi
Kredi	Konutun banka kredisine uygun olup olmadığı
Kimden	Konutun kim tarafından ilana verildiği
Konum	Konutun enlem ve boylam olarak konum bilgisi
İl	Konutun bulunduğu şehir
İlce	Konutun bulunduğu ilçe
Mahalle	Konutun bulunduğu mahalle

Kaynak: Yazar tarafından derlenmiştir

4.5. VERİ HAZIRLAMA (VERİ ÖNİŞLEME)

Bu bölümde verilerin makine öğrenmesi teknikleri kullanımına uygun hale gelebilmesi için uygulanan işlemler yer almaktadır. Veri önışleme aşamasında Python yazılım dili, ayrıca Python için geliştirilen Pandas kütüphanesinden ve bir makine öğrenmesi kütüphanesi olan Scikit-learn den yararlanılmıştır. Veri hazırlama kısmı, eksik verilerin kontrol edilmesi ve tamamlanması, varsa numerik olmayan verilerin numerik hale getirilmesi, özellik seçimi, varsa aykırı değerlerin tespiti ve çıkarılması, özellik mühendisliği ve normalizasyon olmak üzere altı aşama ile tamamlanmıştır. Tablo 4’te yer alan veri setleri için sadece ilan veri seti tahminleme algoritmaları için kullanılmıştır, diğer veri setleri için ise zaman serisi algoritmaları kullanılmıştır. Tahminleme algoritması kullanılacak veri seti için bütün veri önışleme aşamaları yapılmıştır. Zaman serisi algoritmaları kullanılacak diğer veri setlerinde önışleme sadece normalizasyon aşamasında gerçekleşmektedir.

4.5.1. Eksik Verileri Tamamlama

Şekil 19 da görüldüğü gibi veri setinde eksik veri bulunup bulunmadığının kontrolü gerçekleştirilmiş olup eksik veri içeren özellikler tespit edilmiştir. En çok eksik veri “binayası” özelliğinde mevcuttur. İnceleme ve kontrol sonucunda bina yaşı

“0” olanların boş olduğu görülmüştür. Burada Pandas kütüphanesinin “fillna” özelliğinden faydalanıp boş veriler “0” ile doldurulmuştur. m2net özelliğinde yer alan eksik sekiz veri ise scikit-learn kütüphanesinin Imputer özelliğinden faydalanıp ortalamalara göre tamamlanmıştır. Her özellikte yer alan bir eksik veri ise fazla satırı ifade etmektedir, silinerek veri setindeki eksik veriler tamamlanmıştır.

Şekil 12 de yapılan kontrol üzerine veri setinde eksik veriler tamamlanmış olup eksik veri bulunmamaktadır.

Şekil 12: Eksik Verilerin Kontrolü

```
mahalle      0
ilce         0
fiyat        0
no           0
tarih        0
m2net        0
m2burut      0
binayasi     0
bulundugukat 0
katsayisi    0
isitma       0
banyo        0
balkon       0
site         0
kredi        0
kimden       0
konum        0
aciklama     0
il           0
baslik       0
dtype: int64
```

Kaynak: Yazar tarafından derlenmiştir.

4.5.2. Metin Değerlerin Numerik Hale Getirilmesi

Şekil 9 da görüldüğü gibi özelliklere ait değerler kontrol edilmiş olup numerik olmayan metin değerler mevcuttur. Makine öğrenmesi teknikleri uygulamak için değerlerin numerik hale getirilmesi gerekmektedir. Bu alanda tekrar Scikit-learn preprocessing kütüphanesine başvurulmuş olup ilk olarak LabelEncoder özelliğinden

faýdalanılmıştır. LabelEncoder metin deęerleri sınıflandırarak sırasıyla 0,1,2,.. şeklinde numaralandırmaktadır. Fakat veri setinde yer alan mahalle özelliğinde mahalleler arası ast-üst ilişkisi olmadığından yine scikit-learn preprocessing kütüphanesinden OneHotEncoder özelliğinden faýdalanılmıştır. OneHotEncoder özelliği ile mahalleler aralarında bir ast-üst ilişkisi olmayacak şekilde numerik hale getirilmiştir.

Metin deęer içeren bazı özelliklerin numaralandırılması şu şekildedir;

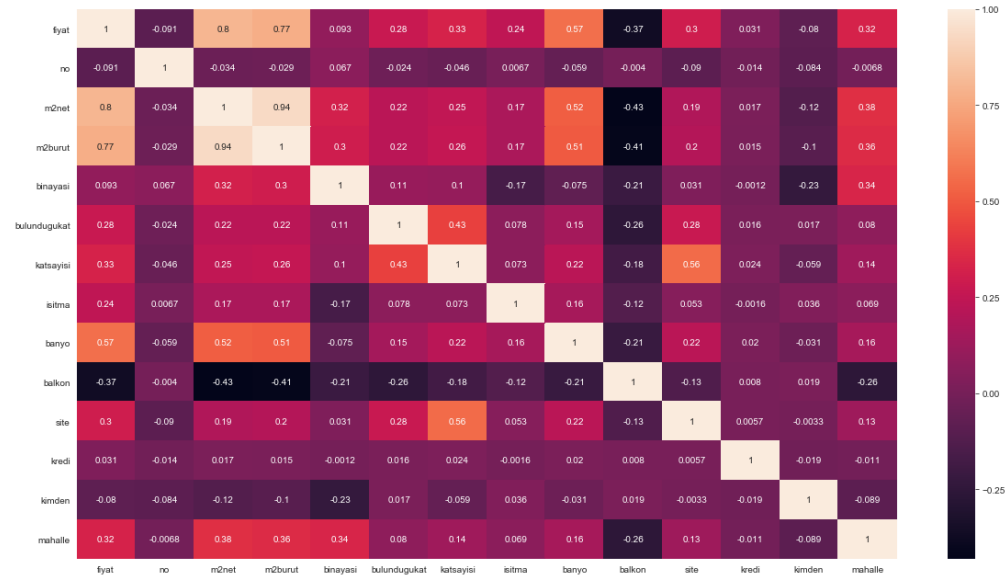
- Odasayisi = Konutun sahip olduęu oda sayısını ifade eden bu özellikte 1+0 ise 1, 1+1 ise 2 , 2+1 ise = 3 , 3+1 ise 4 , 4+1 ise 5 şeklinde oda ve salon toplamaları alınmıştır. 2+0 şeklinde ki ilanlara ait numaralandırma ise 2 oda ve ortak alana sahip olduğundan farklı deęerlendirilmiş olup 2,5 olarak olarak numaralandırılmıştır.
- Binayasi= Konutun bulunduęu bina yaşını ifade eden bu özellikte, 0,1,2,3,4 yaşlarındaki binalar numerik haldedir. Numerik halde olmayan 5-10 arası, 11-15 arası, 16-20 arası, 21-25 arası, 26-30 arası ve 31 ve üzeri seçenekleri ise sınıflandırılarak numerik hale getirilmiştir.
- Bulundugukat = Konutun bulunduęu katı ifade eden bu özellikte giriş kat, bahçe katı, yüksek giriş, kot 1 gibi numerik olmayan veriler mevcuttur. Zemin ile birinci kat arasında olduğ u için sıfır ile bir arasında deęerlendirilip numerik hale getirilmiştir.
- Isıtma = Konutun ısınma tipini ifade eden bu özellikte ısıtma sistemi olmayan yok seçeneği 0 olarak deęerlendirilmiş olup soba = 1, klima = 2, kat kaloriferi = 3, doğalgaz sobası = 4 , merkezi sistem = 5 , merkezi pay ölçer = 6 , doğalgaz kombi = 7 ve yerden ısıtma = 8 olarak numaralandırılmıştır.
- Site = Konutun site içerisinde yer alıp almadığını ifade eden bu özellikte hayır = 0, evet = 1 olarak numaralandırılmıştır
- Kredi = Konutun krediye uygun olup olmadığını ifade eden bu özellikte hayır = 0, evet = 1 olarak numaralandırılmıştır.

4.5.3. Özellik Seçimi

Bu bölümde veri setinde yer alan 21 özelliğin arasındaki ilişkiyi ölçmek için bir korelasyon matrisi oluşturulmuştur. Daha sonra hedef nitelik olan fiyat ile diğer nitelikler arasındaki ilişkiyi ölçmek için bir korelasyon grafiği oluşturulmuştur.

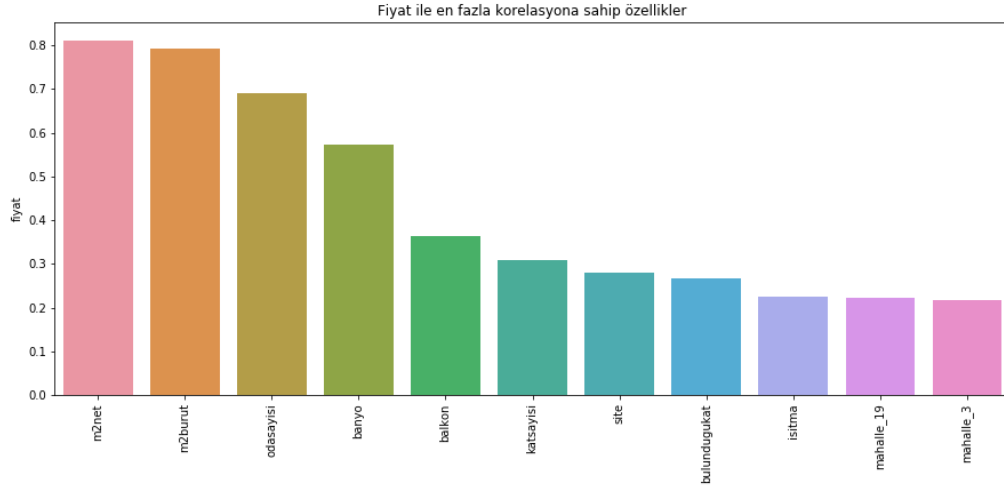
Şekil 13 ve 14 de görüldüğü gibi fiyatı en çok etkileyen özellikler net ve brüt m2 ve oda sayısı olarak görülmüştür. Bu grafiklerden elde edilen sonuçlara göre hedef nitelik olan fiyata etkisi olmayan veya dikkate alınacak değerde bir etkisi olmayan konum, ilan no, il, ilçe, ilan tarihi, başlık ve açıklamayı ifade eden özellikler veri setinden çıkartılmıştır.

Şekil 13: İlan Veri Seti Korelasyon Matrisi



Kaynak: Yazar tarafından derlenmiştir

Şekil 14: Fiyat ile En Fazla Korelasyona Sahip Özellikler



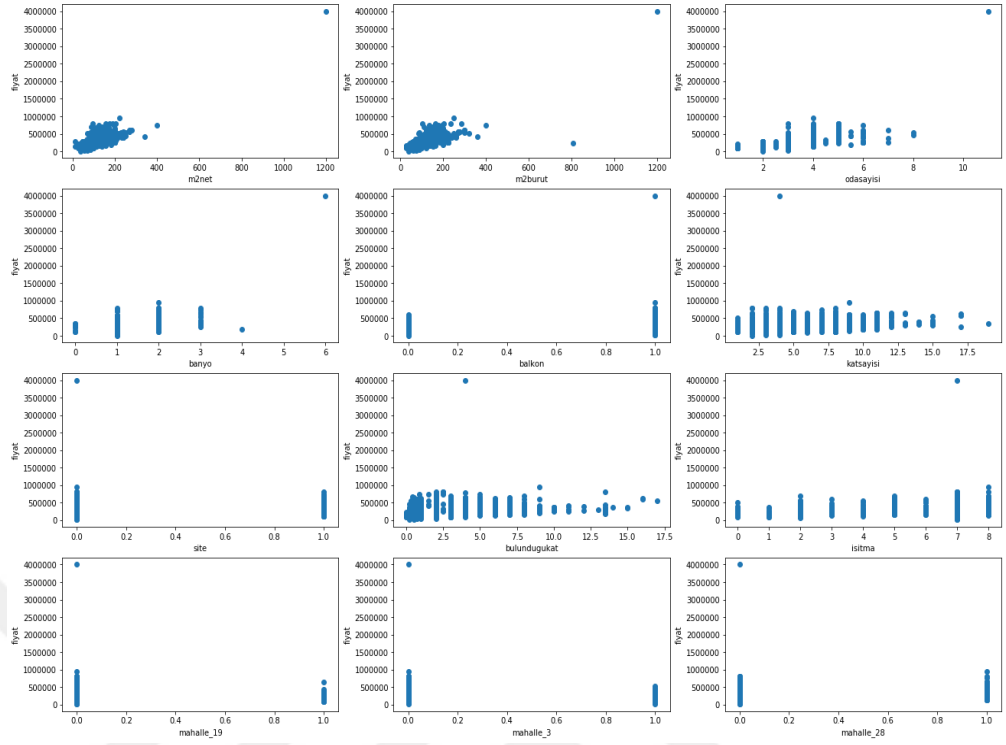
Kaynak: Yazar tarafından derlenmiştir.

4.5.4. Aykırı Değerlerin Tespiti ve Çıkarılması

Bu bölümde makine öğrenmesi teknikleri için uygulanacak metotlara etki edebilecek aykırı değerlerin belirlenmesi amaçlanmıştır. Kutu grafikleri çizilerek veri setinde gözle görülebilen bir farkla ayrılmış aykırı değerler tespit edilip veri setinden ayıklanmıştır.

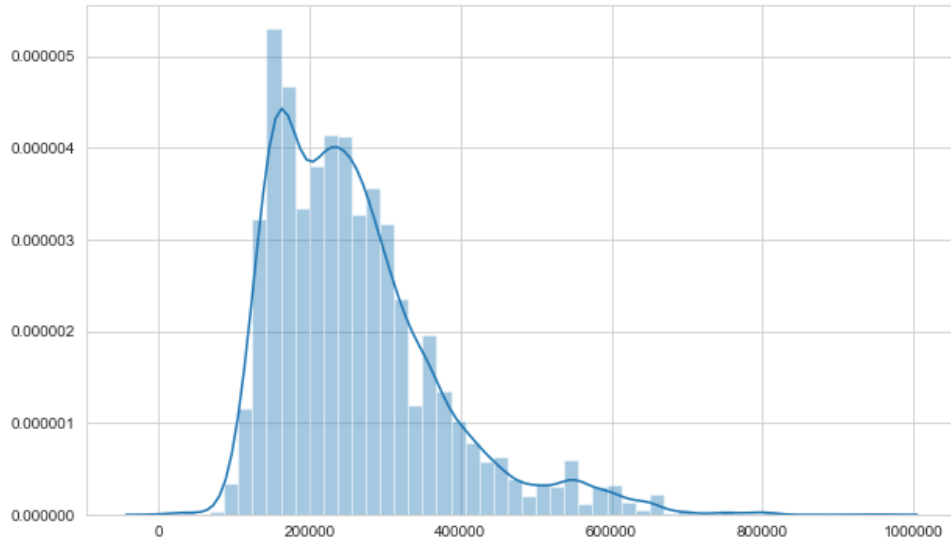
Şekil 15 de görüldüğü üzere 1200m2 bir alana sahip konut aykırı olarak grafiklerde gözle görülmektedir. Veri setinden çıkarılmıştır. Şekil 16'da ise aykırı değerler veri setinden çıkarıldığında elde edilen sonuçlara göre konutların fiyat dağılımları yer almaktadır.

Şekil 15: Aykırı Değer Tespiti Korelasyon Grafikleri



Kaynak: Yazar tarafından derlenmiştir

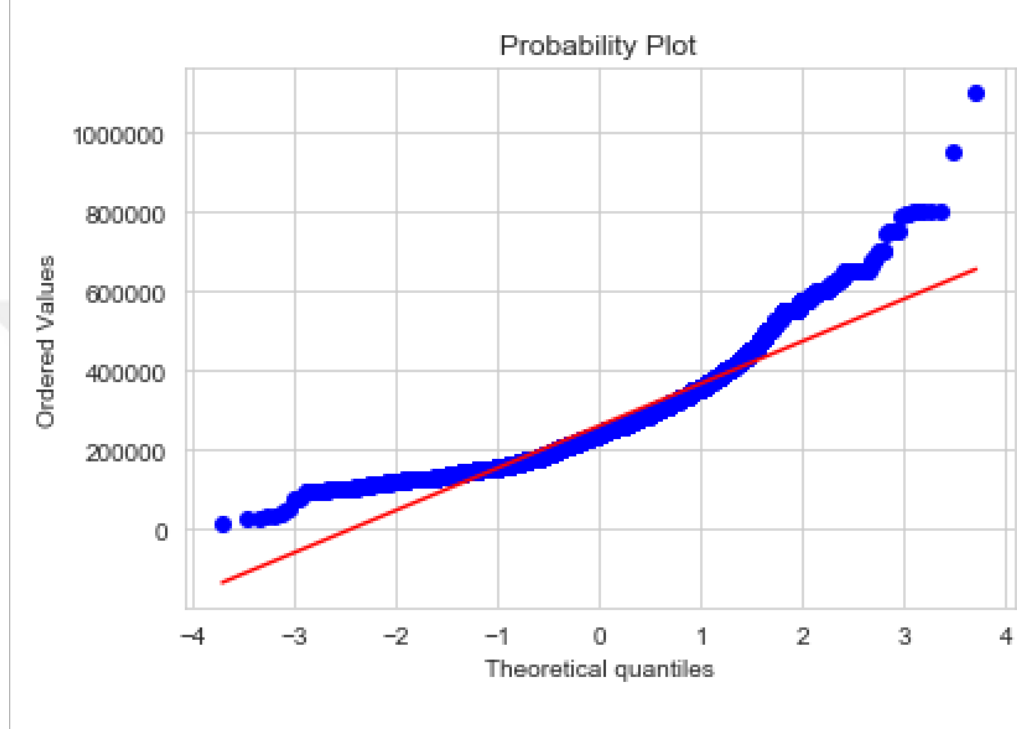
Şekil 16: Aykırı Değerler Çıkarıldıktan Sonra Konut Fiyat Dağılımı



Kaynak: Yazar tarafından derlenmiştir

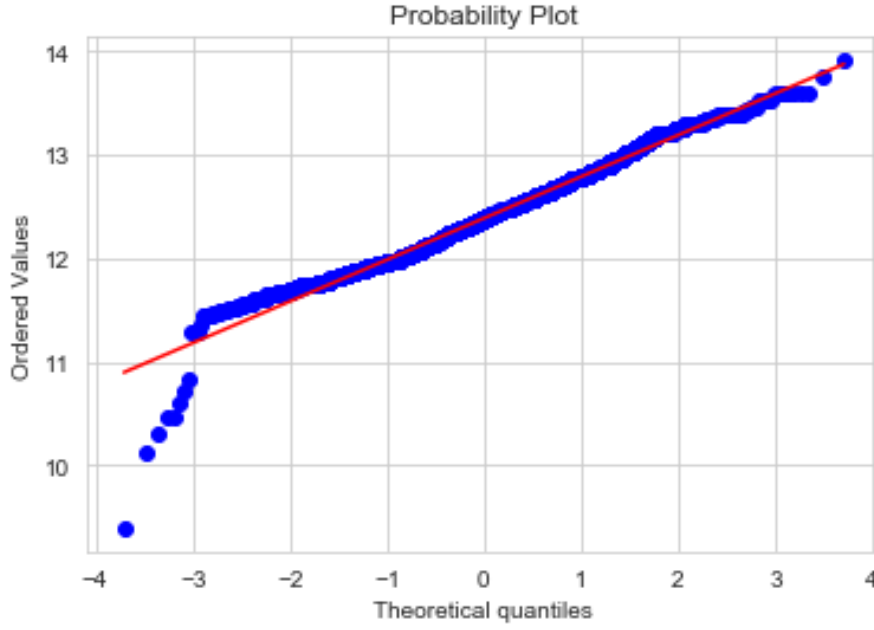
şekil 18 de görüldüğü gibidir. Normalizasyon metodu olarak logaritmik dönüşüm uygulandığında Şekil 19 da doğrusal bir hale daha yakın hale gelmiştir. Aynı işlem zaman serisi algoritmaları uygulanan veri setleri içinde gerçekleştirilmiştir

Şekil 18: İlan Veri Seti Normalize Edilmeden Öncesi Dağılım



Kaynak: Yazar tarafından derlenmiştir

Şekil 19: İlan Veri Seti Normalizasyon Sonrası Dağılımı



Kaynak: Yazar tarafından derlenmiştir

Normalizasyon öncesi ve sonrası modelin başarısı fark etmektedir.

4.6. MODEL KURMA

Modelin kurulması aşamasında Bölüm 2 de bahsedilen MLR, PR, Lasso, Ridge, KNN varyasyonları, DT, SVR, RF ve GBM tahminleme algoritmaları kullanılmıştır ve zaman serileri ile geleceğe yönelik tahminleme için ARMA ve Facebook Prophet algoritmaları kullanılmıştır. Modelin eğitilmesi için k-kat çapraz geçerleme yöntemi kullanılmış olup literatürde çokça geçen $k=10$ ve $k=5$ yaklaşımları ile denenmiştir.

4.7. MODEL DEĞERLENDİRME VE SEÇİMİ

Uygulanan algoritmalar MAE, MSE, RMSE, R2 ve Scikit-learn kütüphanesi başarı skoruna göre değerlendirme ölçütleri ile değerlendirilmiş olup sonuçlar Tablo 5 de belirtilmiştir;

Tablo 5: Tahminleme Algoritmalarının Karşılaştırılması (k = 5 için)

	MAE	MSE	RMSE	R2	Başarı Skoru
MLR	0.14827	0.04106	0.20265	0.74844	0.77696
PR 2. Derece	0.14394	0.03879	0.19697	0.76237	0.78952
Ridge	0.15109	0.04361	0.20885	0.73266	0.76783
Lasso	0.15090	0.04377	0.20922	0.73169	0.76817
KNN-Uniform	0.14882	0.04146	0.20363	0.74602	0.76383
KNN-Distance	0.11621	0.03224	0.17957	0.80260	0.81835
GBM	0.10622	0.02423	0.15568	0.85160	0.85626
DT	0.15877	0.04711	0.21705	0.71151	0.73307
RF	0.10661	0.02521	0.15877	0.84575	0.86170
SVR	0.14289	0.03869	0.19671	0.76312	0.78858

Kaynak: Yazar Tarafından Derlenmiştir**Tablo 6:** Tahminleme Algoritmalarının Karşılaştırılması (k=10 için)

	MAE	MSE	RMSE	R2	Başarı Skoru
MLR	0.14813	0.04078	0.20194	0.75059	0.76339
PR 2. Derece	0.14371	0.03829	0.1957	0.76589	0.77706
Ridge	0.15067	0.04235	0.20581	0.74094	0.760429
Lasso	0.15046	0.04241	0.20595	0.74061	0.760113
KNN-Uniform	0.14826	0.04115	0.20286	0.74852	0.74858
KNN-Distance	0.11211	0.03086	0.17567	0.81164	0.81351
GBM	0.10367	0.0234	0.1559	0.85172	0.86056

DT	0.1589	0.04807	0.21926	0.7058	0.71808
RF	0.10353	0.02432	0.15594	0.85191	0.86623
SVR	0.14226	0.03834	0.1958	0.7659	0.78264

Kaynak: Yazar tarafından derlenmiştir

Tablo 7: Facebook Prophet Algoritması ile Başarı Oranları

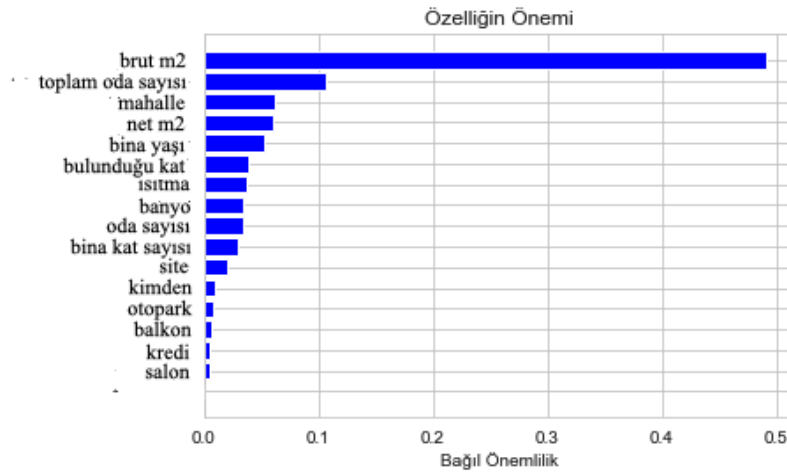
	MAE	MSE	RMSE	R2
İnşaat Ciro Endeksi Veri seti	0.02442	0.00106	0.03263	0.98445
İnşaat Maliyet Endeksi Veri Seti	1.51545	4.35411	2.08665	0.99454
Buca Konut Satış Sayıları Veri Seti	0.06605	0.00714	0.08450	0.83538
İnşaat Güven Endeksi Veri Seti	0.03147	0.00153	0.03917	0.95451
Konut Faiz Oranları Veri Seti	0.14882	0.04146	0.20363	0.74602
USD/TRY Veri Seti	0.03111	0.00142	0.03776	0.90144
Tüfe İzmir Veri Seti	0.00158	5.57247	0.00236	0.99975
Adatepe Mahallesi Metrekare Fiyatları Veri Seti	0.02191	0.0007	0.02654	0.95129
Akıncılar Mahallesi Metrekare Fiyatları Veri Seti	0.01463	0.00038	0.01966	0.98036

Atatürk Mahallesi Metrekare Fiyatları Veri Seti	0.02099	0.00081	0.02860	0.93945
Barış Mahallesi Metrekare Fiyatları Veri Seti	0.01928	0.00055	0.02361	0.97086

Kaynak: Yazar tarafından derlenmiştir

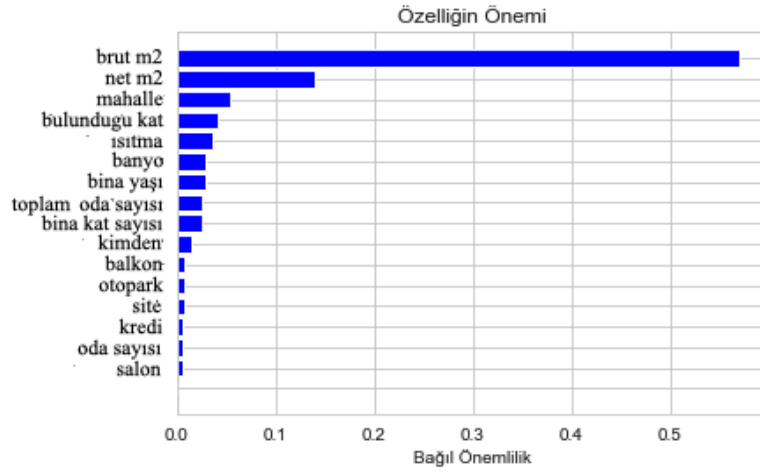
Şekil 20 ve Şekil 21 de en başarılı sonuç veren iki algoritmanın veri setindeki özelliklere göre bağıl önem derecesi görülmektedir. GBM algoritmasına göre en önemli özellik brüt metrekare sonrasında toplam oda sayısı gelmektedir. RF algoritmasına göre en önemli özellik yine brüt metrekare iken sonrasında net metrekare gelmektedir.

Şekil 20: GBM Algoritmasına Göre Özellik Önemleri



Kaynak: Yazar tarafından derlenmiştir

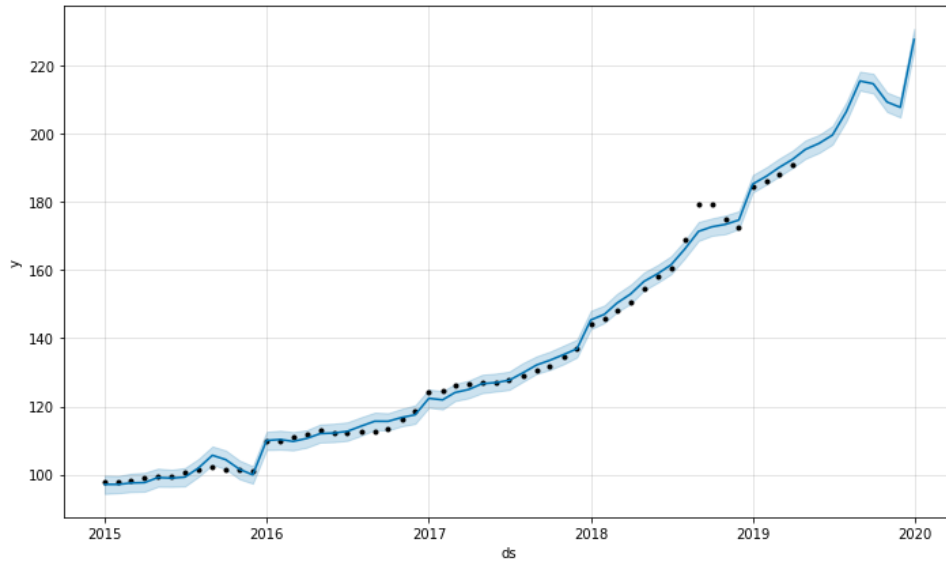
Şekil 21: RF Algoritmasına Göre Özellik Önemleri



Kaynak: Yazar tarafından derlenmiştir

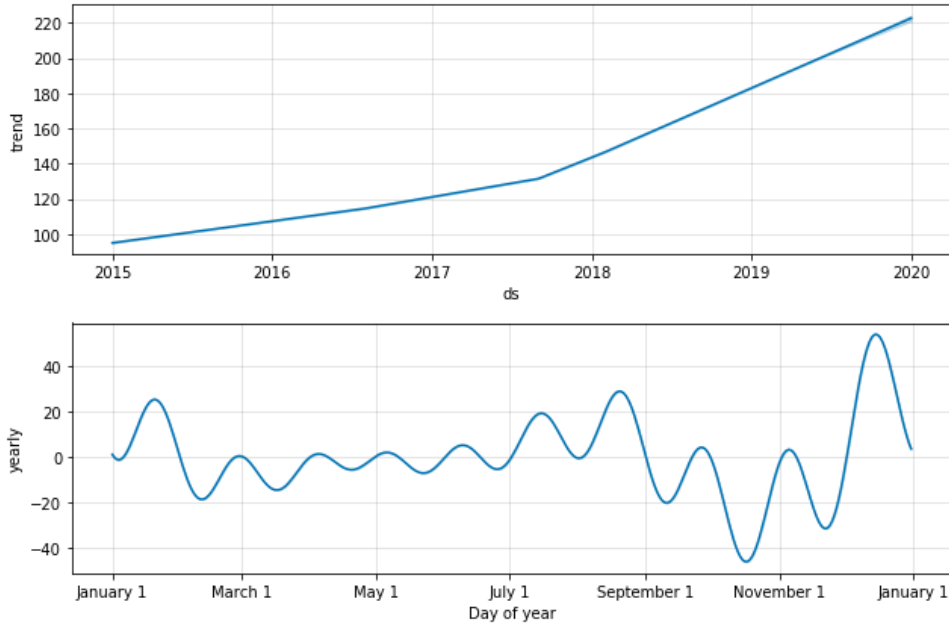
Şekil 22 de inşaat maliyet endeksi veri setine ait 2015 yılından itibaren zaman serisi ve gelecek tahminlemesi yer almaktadır. Şekil 23 de veri setine ait trend çizgisi ve yıllık dağılımı yer almaktadır. Yıllık dağılımda en düşük eylül kasım ayları arasında, en yüksek ise kasım ocak ayları arasında olduğu görülmektedir. Bu grafikler ile geçmişe yönelik analiz ve geleceğe yönelik maliyet endeksi tahminlenmesi amaçlanmıştır.

Şekil 22: İnşaat Maliyet Endeksi Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir

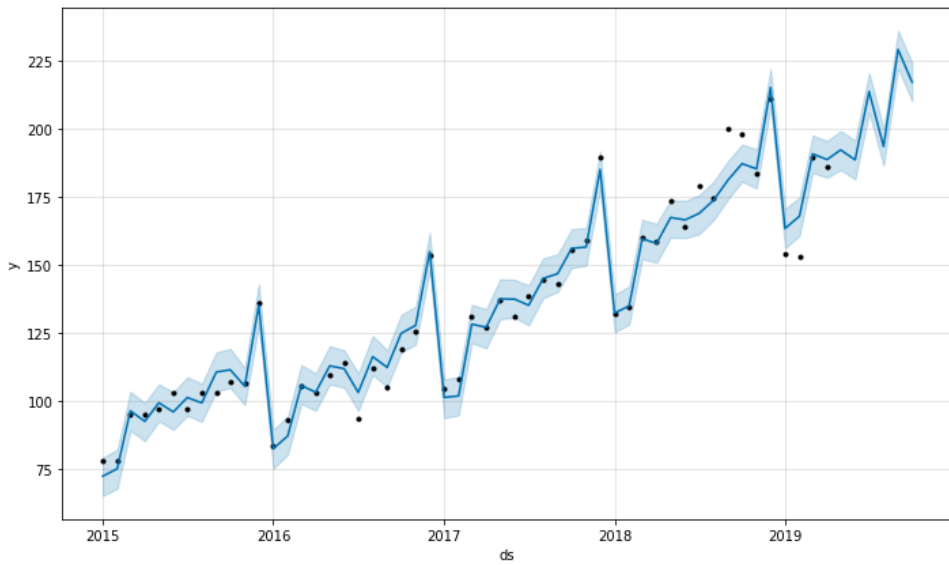
Şekil 23: İnşaat Maliyet Endeksi Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir

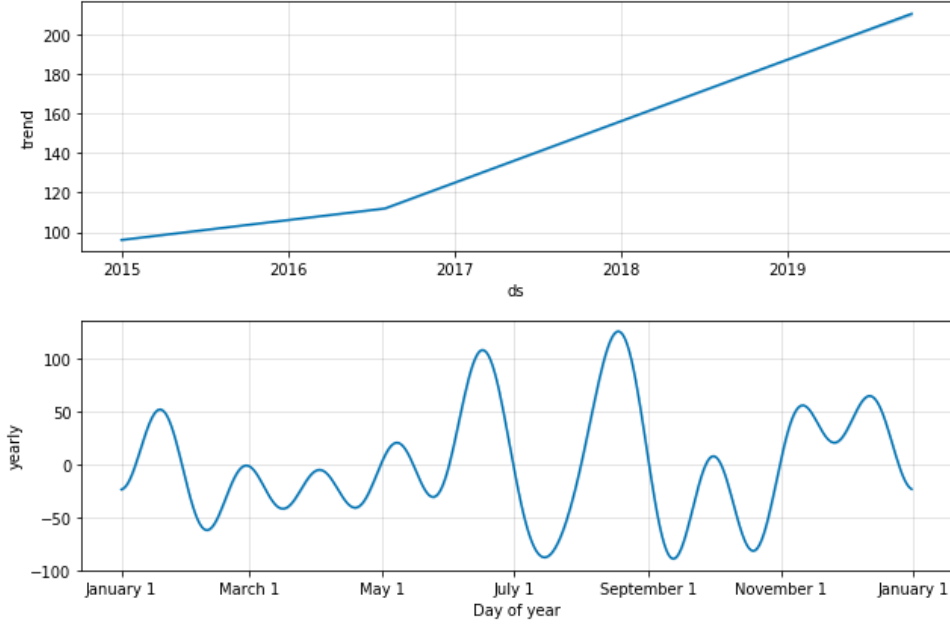
Şekil 24 ve Şekil 25 de inşaat ciro endeksine ait zaman serisi analizi, trend çizgisi ve yıllık dağılım gösterilmiştir. Her yılbaşından önce ciro endeksinin arttığı, 2016 yılının ortalarından itibaren trend çizgisinin daha dik eğime sahip olduğu ayrıca yıllık dağılımda tepe ve dip noktalar ay bazında görülmektedir.

Şekil 24: İnşaat Ciro Endeksi Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir

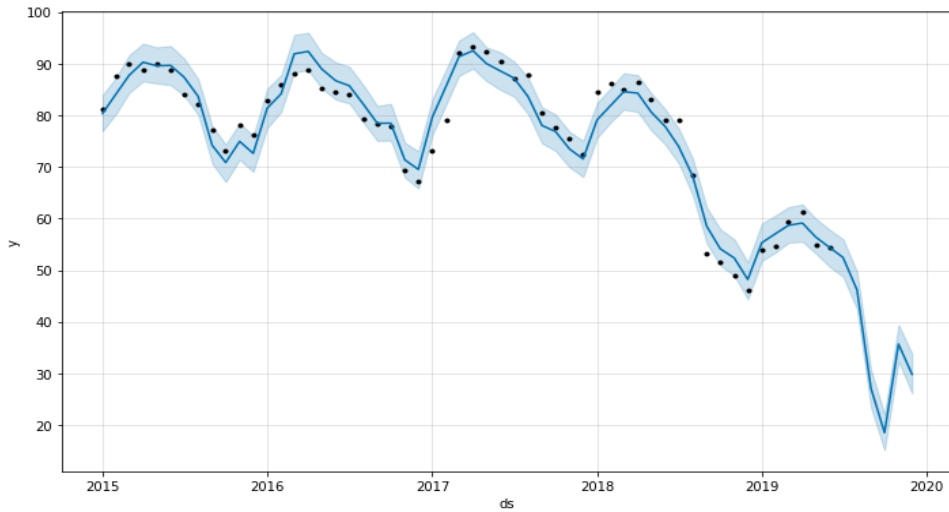
Şekil 25: İnşaat Ciro Endeksi Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir

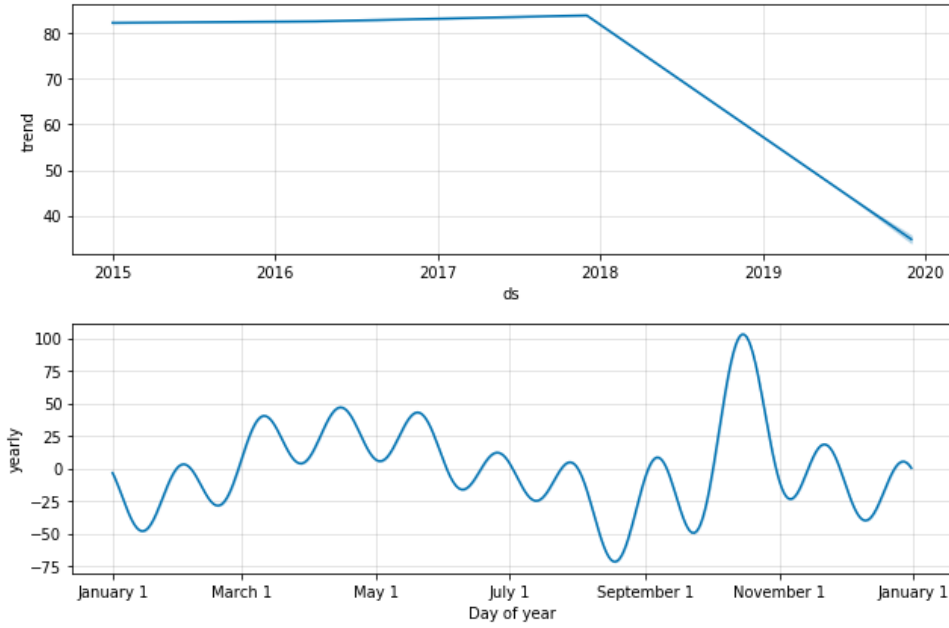
Şekil 26 ve Şekil 27’de inşaat güven endeksi için zaman serisi grafiği, geleceğe yönelik tahminlemesi, trend çizgisi ve yıllık dağılımı yer almaktadır. 2017 yılının sonlarından itibaren düşüş eğilimine geçtiği görülmektedir yıllık bazda ise en yüksek seviyeye eylül ve kasım ayları arasında ulaşmaktadır. Bu grafikler ile sektörel analizlerin yapılması ve tahminlenmesi amaçlanmıştır.

Şekil 26: İnşaat Güven Endeksi Zaman Grafiği



Kaynak: Yazar tarafından derlenmiştir.

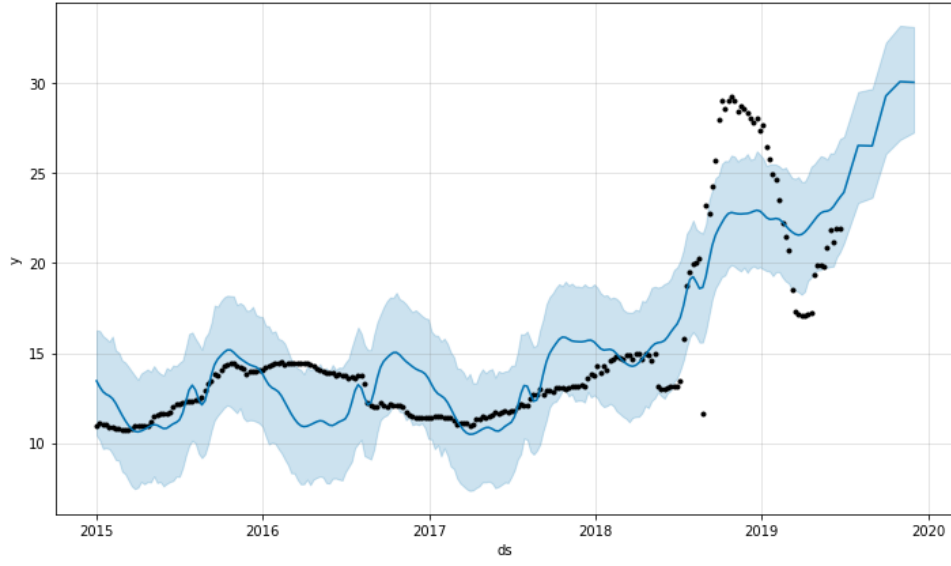
Şekil 27: İnşaat Güven Endeksi Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

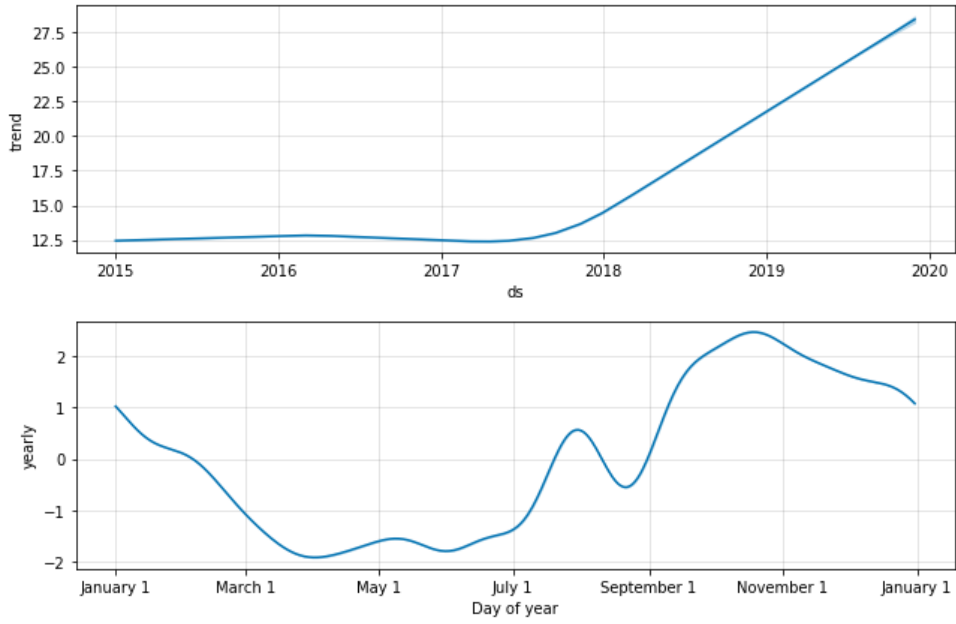
Şekil 28 ve Şekil 29 da bankaların uyguladığı konut kredisi faiz oranlarının zaman serisi grafiği, tahminlemesi, trend çizgisi ve yıllık dağılımı yer almıştır. 2017 ortaları itibari ile yükselen bir eğilime geçmiştir ayrıca yıllık bazda ise Mart ve Temmuz ayları arasında en düşük, Eylül ve Kasım ayları arasında ise en yüksek seviyelere ulaşmıştır. Bu grafikler ile konut kredisi kullanacak ya da kullandıracak kişi ve kurumlar için zaman analizleri ve tahminlemelerinin yapılması amaçlanmıştır.

Şekil 28: Konut Kredisi Faiz Oranları Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir

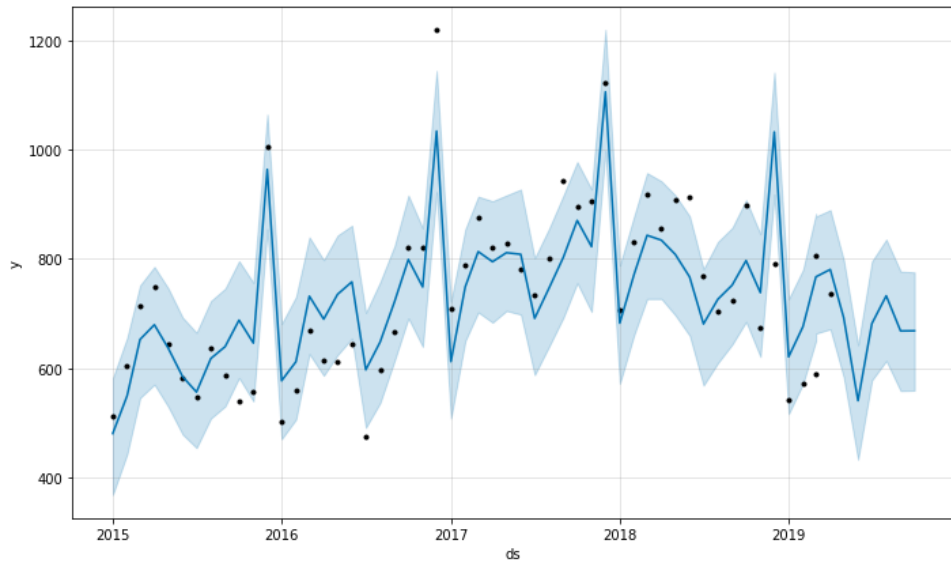
Şekil 29: Konut Kredisi Faiz Oranları Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir

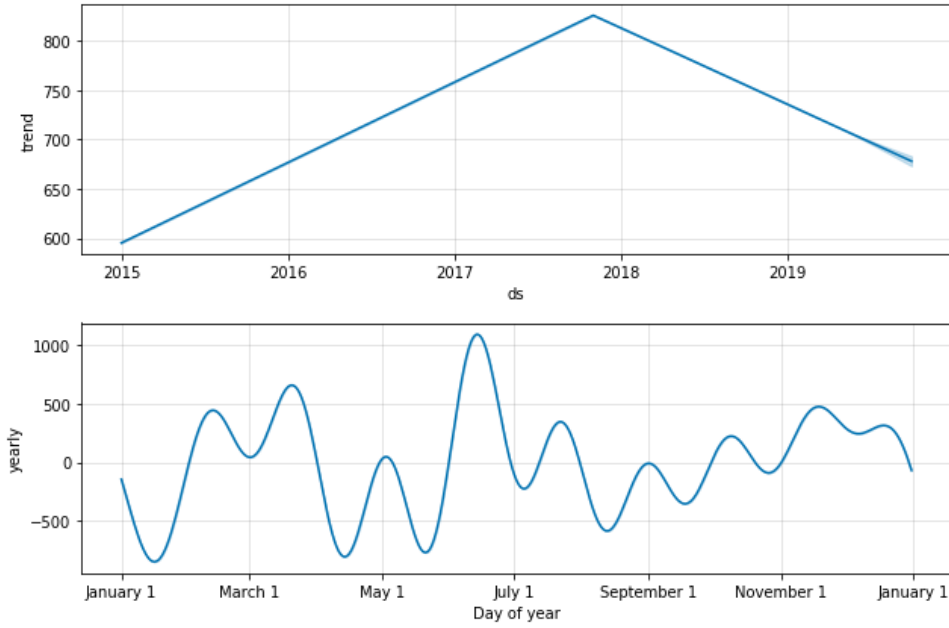
Şekil 30 ve Şekil 31 da İzmir ili Buca ilçesindeki konut satış sayıları zaman serisi grafiği, geleceğe yönelik tahminleme, trend çizgisi ve yıllık dağılımlar yer almıştır. Her yeni yıl öncesi bir yükseliş görülmektedir, trend analizinde ise 2017 sonlarına doğru süren artış yerini azalan eğime bırakmıştır. Yıllık analizde ise en yüksek konut satış yoğunluğunun haziran ve temmuz ayları arasında olduğu görülmektedir. Bu grafikler ile konut satışı için bir planlama yapılabilmesi amaçlanmıştır.

Şekil 30: Buca Konut Satış Sayıları Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir

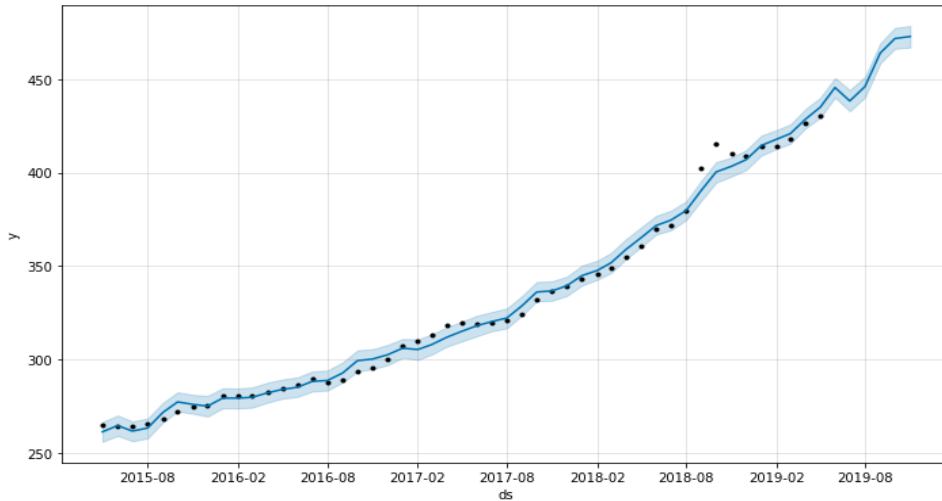
Şekil 31: Buca Konut Satış Sayıları Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir

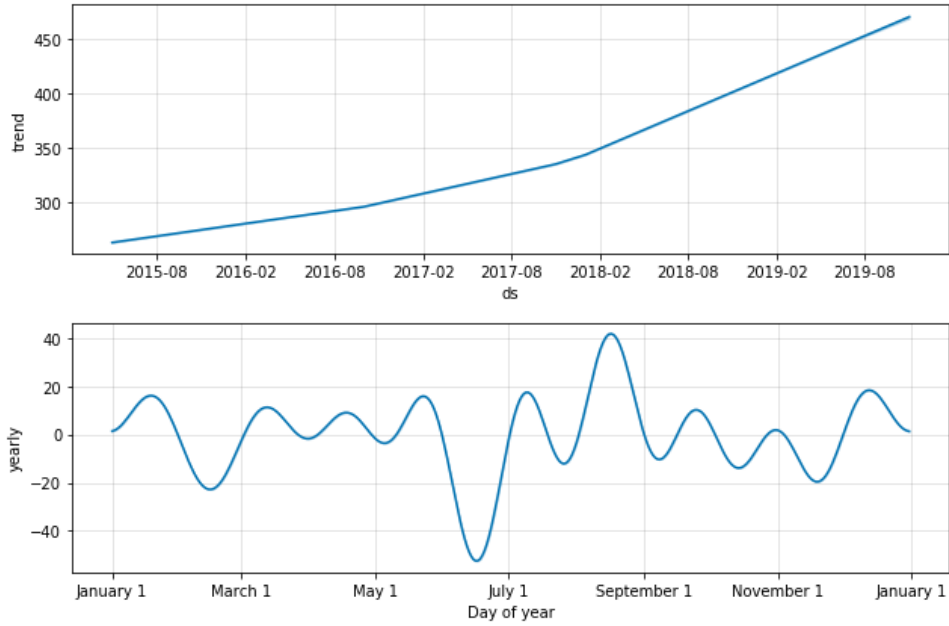
Şekil 32 ve Şekil 33, İzmir iline ait TÜFE zaman serisi, geleceğe yönelik tahminler, trend çizgisi ve yıllık dağılımı göstermektedir. Sürekli artışta olan bir eğim mevcuttur ayrıca yıllık dağılımlarda en düşük seviyeler Haziran- Temmuz ayları arasında en yüksek seviyeler ise Ağustos-Eylül arasında seyretmektedir. Bu grafikler ile ödemelere dayalı işlemler için analizler yapılması amaçlanmıştır.

Şekil 32: İzmir TÜFE Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir.

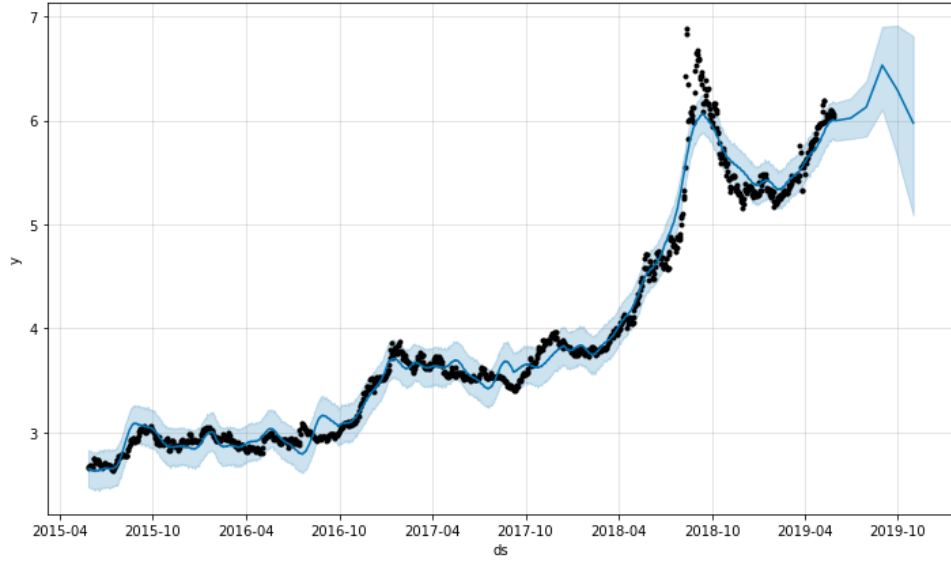
Şekil 33: İzmir TÜFE Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

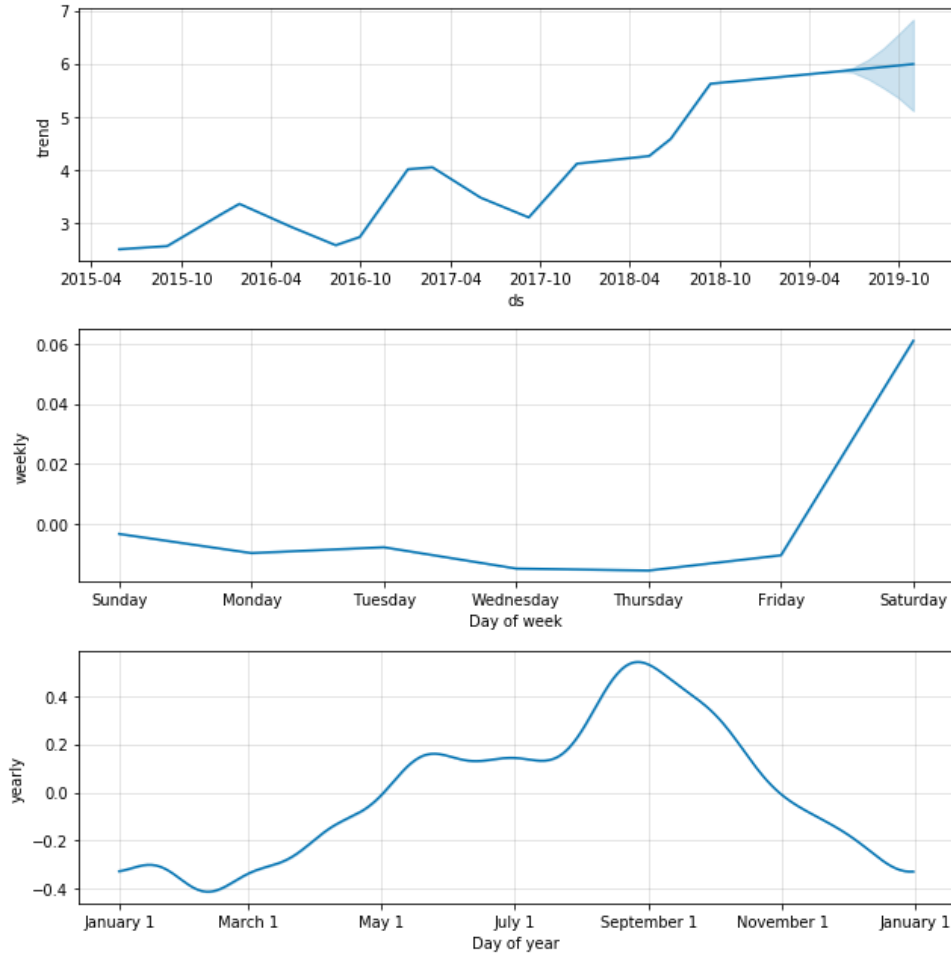
Şekil 34 ve Şekil 35, USD/TRY parametrelerinin zaman serisi analizi, geleceğe yönelik öngörü, trend analizi, haftalık ve yıllık dağılımları göstermektedir. Haftalık analizin amacı işlem saatleri dışında olması ve yüksek oranlı kur uygulamasını göstermektir. Yıllık bazda en yüksek seviye Ağustos ve Eylül ayları arasında yoğunlaşmaktadır. Bu grafiklerde döviz kuru ile yapılan işlemlerde bir öngörüye sahip olmak ve dövizin etkilerinin analizinin yapılabilmesi amaçlanmıştır.

Şekil 34: USD/TRY Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir.

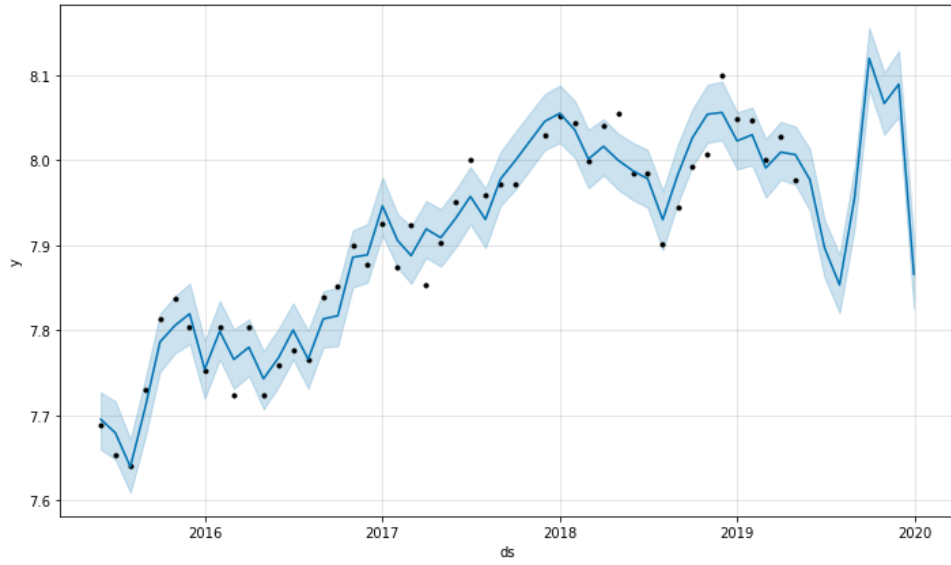
Şekil 35: USD/TRY Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

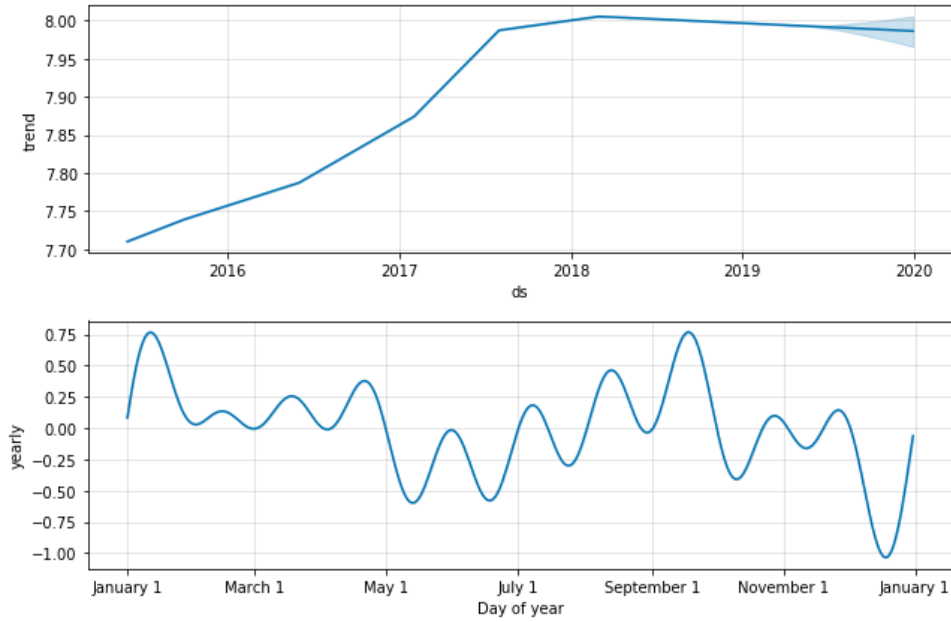
Şekil 36 ve Şekil 37 İzmir ili Buca ilçesi Adatepe mahallesine ait konutların metrekare fiyatlarına ilişkin zaman serisi, geleceğe yönelik tahminleme, trend analizi ve yıllık dağılımları yer almaktadır.

Şekil 36: Adatepe Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir.

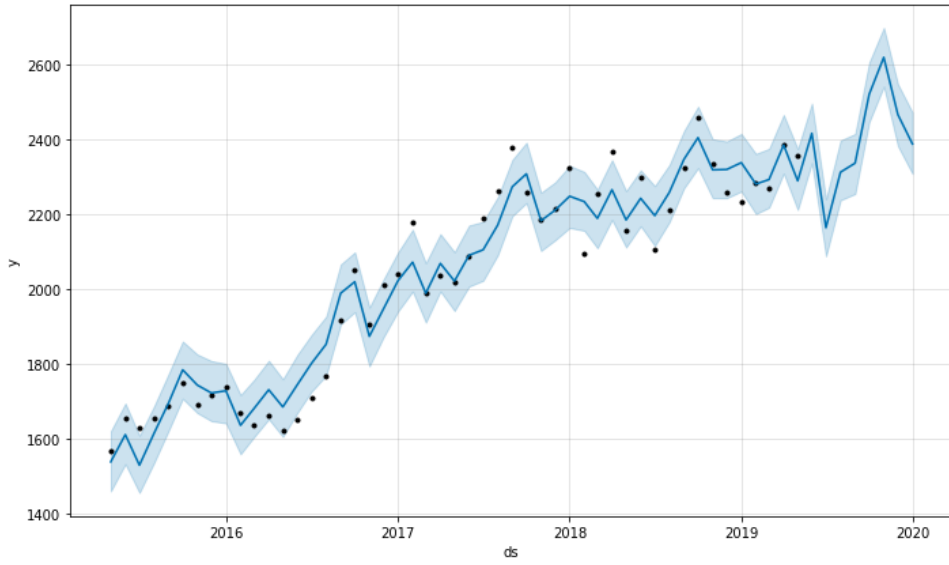
Şekil 37: Adatepe Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

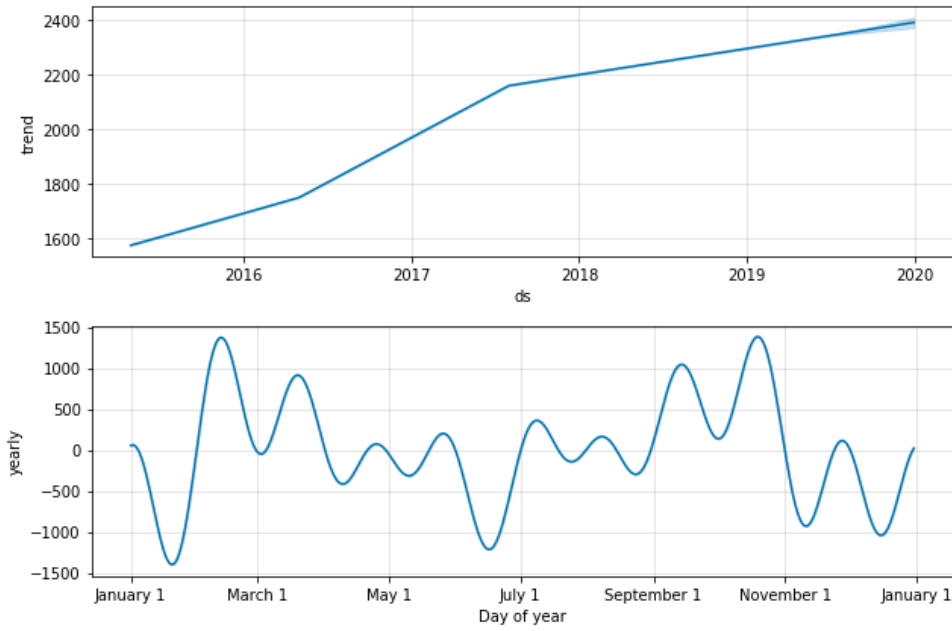
Şekil 38 ve Şekil 39 İzmir ili Buca ilçesi Barış mahallesine ait konutların metrekare fiyatlarının zaman serisi grafiği, geleceğe yönelik tahminleme , trend analizi ve yıllık dağılımları yer almaktadır.

Şekil 38: Barış Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir

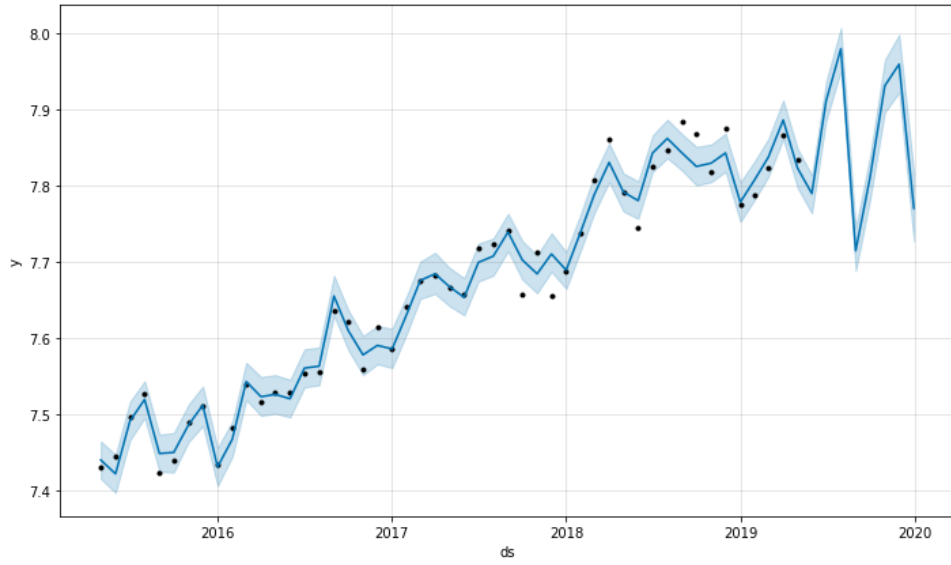
Şekil 39: Barış Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

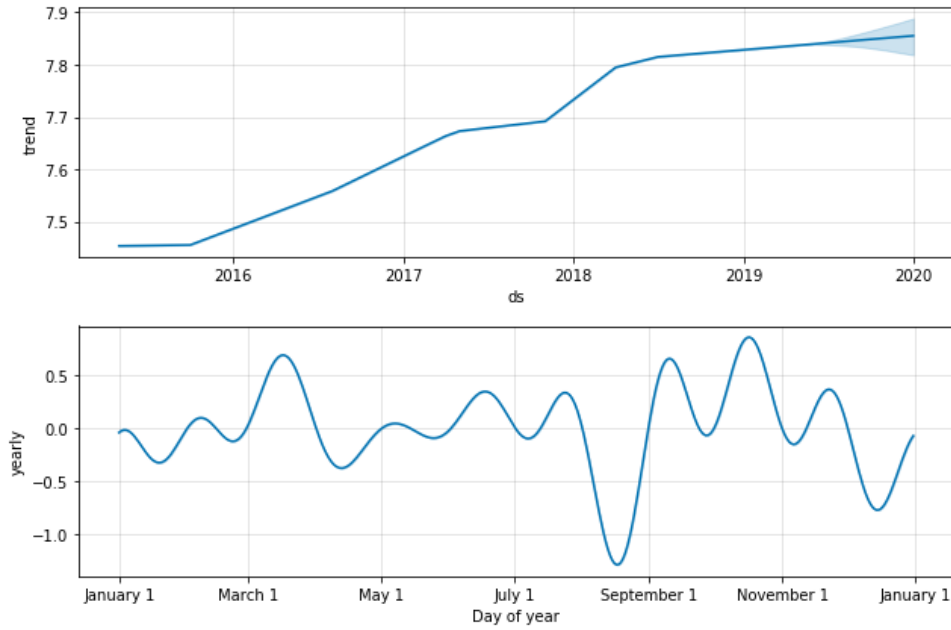
Şekil 40 ve Şekil 41 İzmir ili Buca ilçesi Akıncılar mahallesine ait konutların metrekare fiyatı bazında zaman serisi analizi, geleceğe yönelik tahminleme, trend analizi ve yıllık fiyat dağılımı yer almaktadır.

Şekil 40: Akıncılar Mahallesi Metrekare Fiyatı Zaman Serisi Grafı



Kaynak: Yazar tarafından derlenmiştir

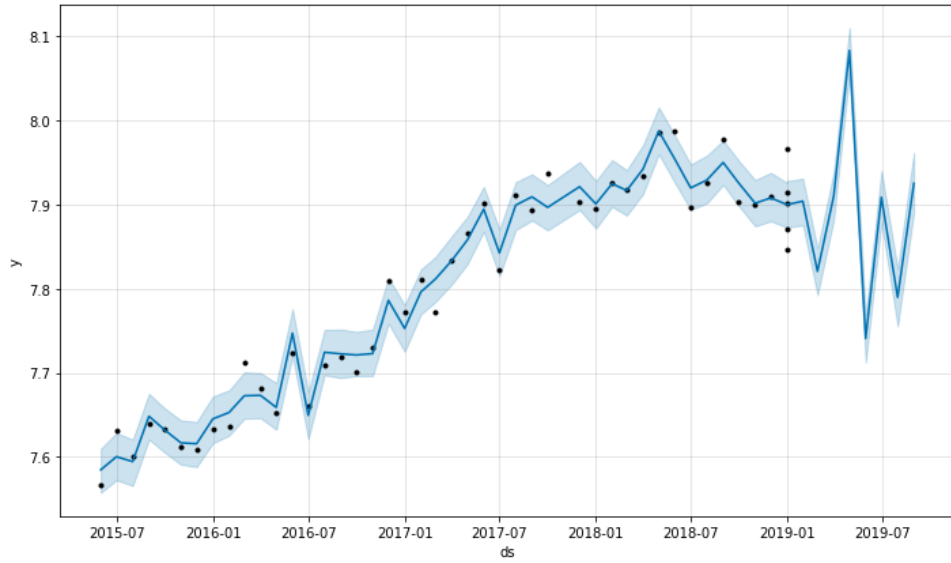
Şekil 41: Akıncılar Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

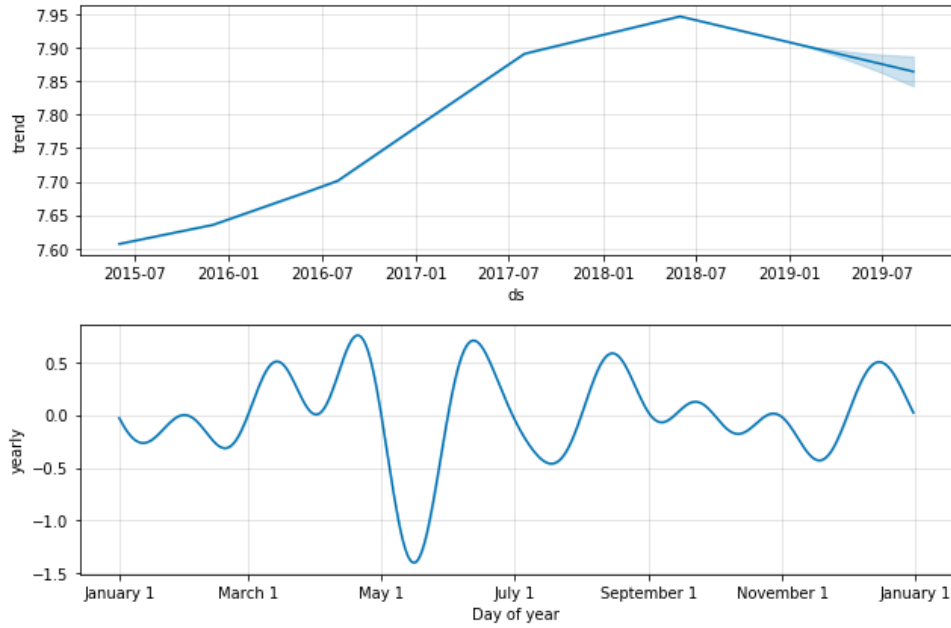
Şekil 42 ve Şekil 43 İzmir ili Buca ilçesi Atatürk mahallesinde bulunan konutların metrekare fiyatlarının zaman serisi analizi, geleceğe yönelik fiyat tahminleri, trend çizgisi ve yıllık fiyat dağılımlarının grafiklerini içermektedir.

Şekil 42: Atatürk Mahallesi Metrekare Fiyatı Zaman Serisi Grafiği



Kaynak: Yazar tarafından derlenmiştir.

Şekil 43: Atatürk Mahallesi Metrekare Fiyatı Trend Çizgisi ve Yıllık Dağılımı



Kaynak: Yazar tarafından derlenmiştir.

Şekil 44 ARMA algoritması ile çizilen mahallelere yönelik karşılaştırma grafiğidir. Bu grafik ile farklı renklerle her bir mahalleyi temsil ederek metrekare fiyat bazında değişimlerin zaman serisini oluşturmak ve tek tabloda karşılaştırmak amaçlanmıştır.

Yukarıdaki şekillerde yer alan zaman serileri Facebook Prophet algoritması kullanılarak oluşturulmuştur. Her veri seti için ayrı ayrı zaman serileri, geleceğe yönelik öngörüler, trend analizleri ve yıllık dağılımlar mevcuttur. Her bir grafik tek başına bir anlam ifade edebileceği gibi karşılaştırmalar ile de farklı anlamlar çıkarılması amaçlanmıştır. Bir iş zekası uygulaması için çapraz sorgulamalar düşünüldüğünde grafiklerde yer alan 2018 sonu döviz kuru artışı ile kredi faiz oranlarının arttığı, inşaat güven endeksinin düştüğü, konut satış sayılarında azalma gerçekleştiği şeklinde çıkarımlar yapılması amaçlanmıştır.

Şekil 44: ARMA Algoritması ile Karşılaştırma



Kaynak: Yazar tarafından derlenmiştir.

4.8. MODELİN UYGULAMAYA GEÇİRİLMESİ

Tahminleme algoritmalarında $k=5$ ve $k=10$ için GBM ve RF çok yakın sonuçlar sağlamıştır. Fakat en başarılı sonucu veren RF algoritması model olarak belirlenip kayıt edilmiştir.

Zaman serisi algoritmalarında ise Facebook Prophet yazılımsal kolaylığı, sunduğu grafikler ve trend analizi gibi özellikleri ile zaman serileri için tercih edilmiştir. Ayrıca ARMA algoritması da özellikle karşılaştırma yapılabilmesi amacı ile iş zekası uygulamasında kullanılmıştır.

4.9. ARAYÜZ GELİŞTİRİLMESİ

Uygulamanın bu kısmında arayüz geliştirilmesi yapılmıştır. Yazılımsal olarak Python frameworkü olan flask kullanılmıştır. Arayüz tasarımı için HTML/CSS kullanılmıştır. Süreç olarak sistem geliştirme yaşam döngüsü adımlarından faydalanılmıştır.

4.9.1. Problemin Tanımı

Kullanılan makine öğrenmesi tekniklerinin sonuçlarını incelemek, karşılaştırmak, çapraz sorgular yapabilmek ve sayısal verilere ulaşabilmek için bir arayüz gerekmektedir. Arayüz ile ihtiyaç duyulan bilgiye hızlı ve etkin bir şekilde ulaşılması sağlanmalıdır.

4.9.2. İhtiyaç Analizi

İhtiyaç analizi aşamasında sektörde kullanılan benzer uygulamalar incelenmiş ve bu sistemlerin eksik kalan tarafları analiz edilmiştir. Kullanılan sistemlerin tek bir alana yönelmiş olması ve sektöre etki eden farklı bilgi kaynaklarından yararlanılmaması fark edilmiştir. Farklı veri kaynaklarından alınan veriler ile yapılacak analizlerin karar vericiler tarafından daha etkin ve verimli kararlar verebilmesi düşünülmüştür.

4.9.3. Sistem Tasarımı

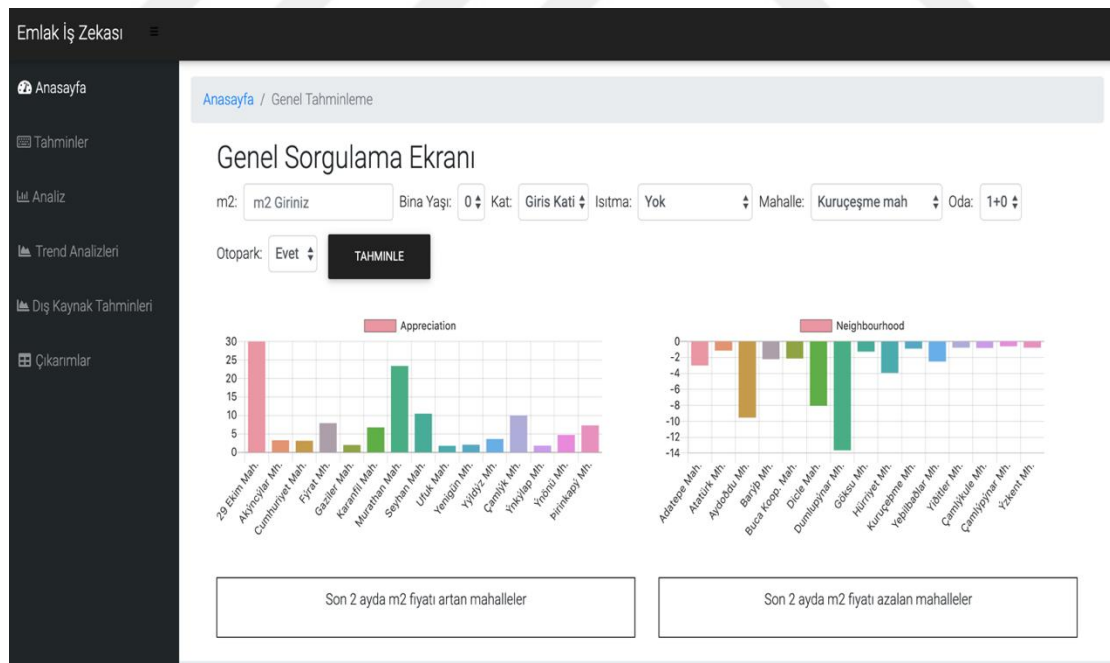
Sistem tasarımı aşamasında, piyasada var olan özellikle ticari sistemler incelenmiş olup, bu sistemlere ait dikkat çekici ve fonksiyonel özelliklerin yanı sıra kolay anlaşılabilir grafikler elde edilmiştir. Makine öğrenmesi teknikleri kullanıldığı için sistem Flask frameworkü sayesinde arka planda sürekli Python dosyalarını çalıştıracak şekilde tasarlanmıştır. Tahminleme algoritması için kullanılan veri seti otomatik olarak güncellenmektedir. Zaman serisi için kullanılan veri setleri ise

güncellendikten sonra arayüz ekranında yer alan bilgiler Flask frameworküne uygulanacak bir yeniden başlatma ile güncel halini almaktadır. Sistemde sorgulamalar, örnek çıkarımlar, bölge analizleri ve genel analizler yer almaktadır.

4.9.3.1. Arayüz Tasarımı

Uygulamanın web arayüzü geliştirilmesi esnasında kullanıcı dostu, kolay anlaşılabilir ve mobil uyumlu bir tasarım tercih edilmiştir. Şekil 45 de görüldüğü üzere uygulamanın ana sayfasında kullanıcının hızlı bir şekilde daha çok kısa vadeli alınabilecek kararlar için amaca yönelik bilgiye ulaşması amaçlanmıştır. İlk ekranda belirlenen kriterlere göre yapılan sorguda konut fiyatına, kredi planına ve aynı özellikte farklı mahallelerde olan konutların fiyatlarına ulaşılabilir. Bunun yanı sıra son iki aya ait mahalle bazında artan ve azalan metrekare fiyatları grafik halinde kullanıcıya sunulmuştur.

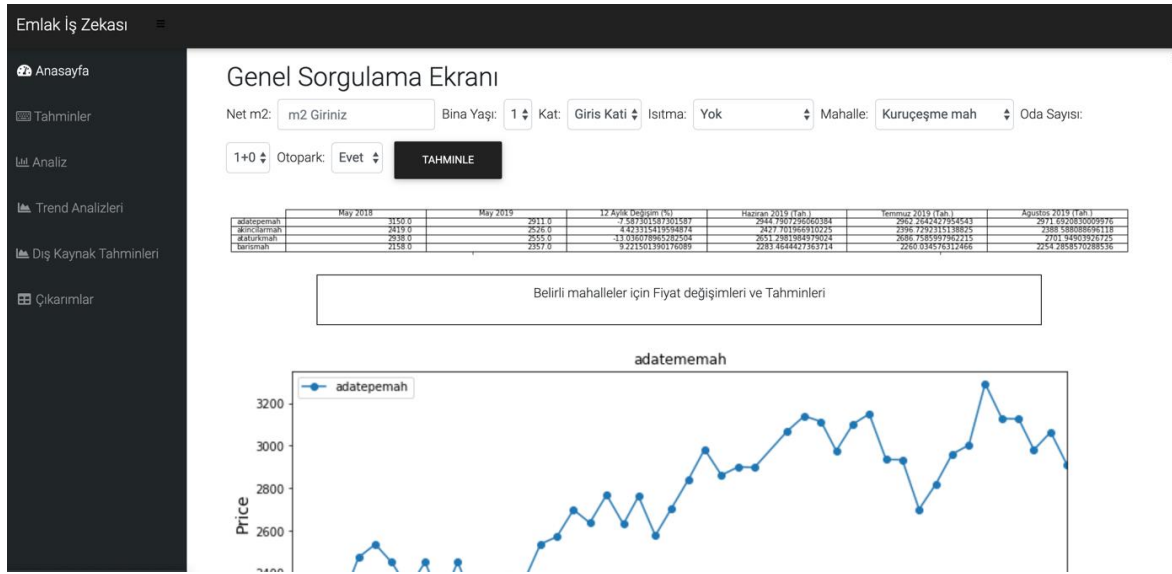
Şekil 45: Uygulama Arayüz Anasayfası



Kaynak: Yazar tarafından derlenmiştir

Şekil 46 de görüldüğü gibi tahminler ekranında ARMA algoritması ile elde edilen sonuçlar yer almaktadır. Belirlenen mahallelerin 2015 yılından 2019 yılı haziran ayına kadar olan metrekare fiyat grafikleri ayrıca gelecek üç aylık tahmini metrekare fiyat bilgilerinin yanı sıra son bir yıla ait metrekare bazında değişim oranları da yer almaktadır.

Şekil 46: Arayüz Tahminler Sayfası



Kaynak: Yazar tarafından derlenmiştir

Şekil 47 de görüldüğü üzere seçilen bölgeye yönelik çeşitli analizler grafiksel olarak yer almıştır. Bölgede yer alan konutların fiyata, metrekareye, oda sayısına, bina yaşına, bulunduğu kata, bina katlarına göre dağılımları bulunmaktadır. Ayrıca site-fiyat ilişkisi, balkonlu konutların sayıları, site içinde olan konutların sayıları, kim tarafından satıldığı, mahallelere göre satılan ilan sayıları, mahallelere göre fiyat aralıkları, ısıtma sistemine göre konut sayısı dağılımları gibi anlam ifade eden grafikler yer almaktadır. Bu sayede bölge trendleri hakkında bilgi vermek ayrıca kullanıcı için bölgeyi daha iyi tanıması eğilimleri farketmesi amaçlanmıştır.

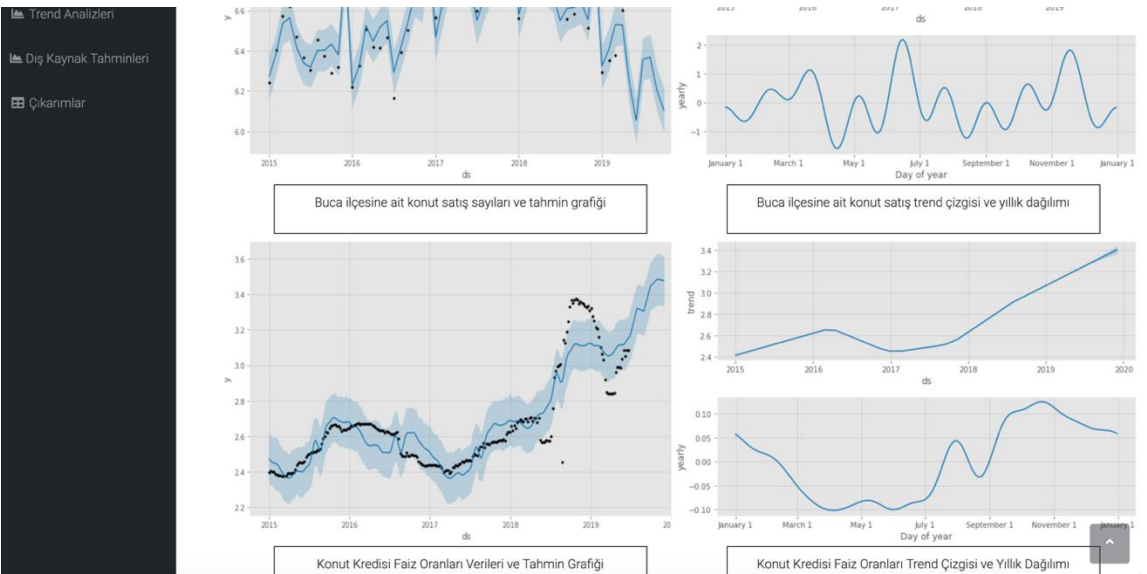
Şekil 47: Uygulama Analiz Arayüzü



Kaynak: Yazar tarafından derlenmiştir

Şekil 48 de görüldüğü üzere bölge ve genele yönelik Bölüm 4.6 da bahsedilen Facebook Prophet algoritması ile elde edilmiş zaman serileri ayrıca trend analizleri yer almaktadır. 2015 yılından 2019 yılı haziran ayına kadar olan grafiklerin yanında 2015 itibari ile var olan trend çizgisi ayrıca aylık dağılım grafikleri ayrı ayrı olarak her veri seti için gösterilmiştir.

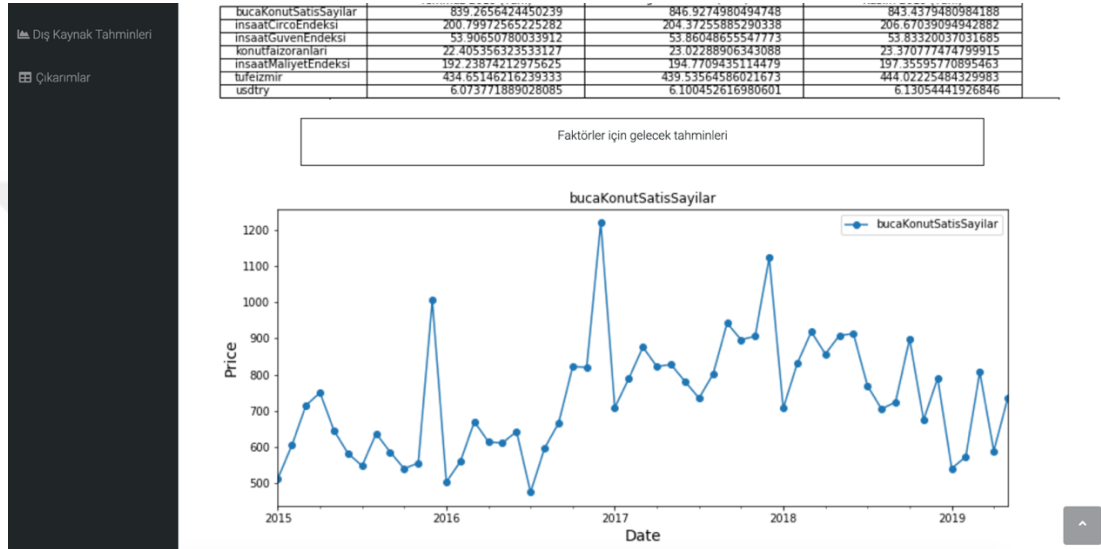
Şekil 48: Arayüz Trend Analizleri Ekranı



Kaynak: Yazar tarafından derlenmiştir

Şekil 49 de görüldüğü üzere Buca konut satış sayıları, inşaat maliyet endeksi, inşaat güven endeksi, inşaat ciro endeksi, konut faiz oranları, TÜFE İzmir ve USD/TRY veri setlerine uygulanan zaman serisi algoritmalarının grafiksel gösteriminin yanı sıra gelecek üç aya yönelik sayısal öngörüler tablo halinde yer almaktadır.

Şekil 49: Arayüz Dış Kaynak Analizleri Ekranı



Kaynak: Yazar tarafından derlenmiştir

Şekil 50 de görüldüğü üzere kullanıcılara kural bazlı senaryolar eşliğinde sistemi etkin ve efektif bir şekilde nasıl kullanacakları hakkında bilgi vermek amaçlanmıştır. Var olan veri setlerinden yapılan tahminleme ve zaman serileri baz alınarak belli bir yatırımın maliyetleri, beklenen kazanç oranları, aylık bazda beklenen maliyet artış ve azalışları, faiz oranlarının artış ve azalışları belirtilerek verilebilecek bir karar için sistemin kullanım şekli örnekler ile anlatılmıştır.

Şekil 50: Uygulama Çıkarımlar Arayüzü

#	Çıkarım 1						
#	Mahalle	Oda Sayısı	Isıtma	Kat Sayısı	Daire Sayısı	Asansör	Otopark
#	Adatepe	2+1	Doğalgaz Kombi	5 Katlı	10 Daire	Yok	Yok
#	Haziran (Güncel)	Temmuz (- %5.23)	Ağustos (+ %7.58)	Eylül (+ %14.27)	Ekim (- %2.78)	Kasım (+ %3.17)	Aralık (- %17.21)
Giriş Kat	218.676 TL	207.240 TL	222.948 TL	254.762 TL	247.766 TL	255.531 TL	211.555 TL
1.Kat	218.751 TL	207.311 TL	233.025 TL	254.850 TL	247.766 TL	255.620 TL	211.628 TL
2.Kat	250.568 TL	237.464 TL	255.463 TL	291.917 TL	283.802 TL	292.798 TL	242.408 TL
3.Kat	243.919 TL	231.163 TL	248.685 TL	284.172 TL	276.273 TL	285.030 TL	235.039 TL
4.Kat	252.485 TL	239.281 TL	257.418 TL	294.151 TL	285.974 TL	295.039 TL	244.263 TL
#						1.384.018 TL	1.146.031 TL
Periyod	Temmuz	Ağustos	Eylül	Ekim	Kasım	Aralık	Toplam
1. Sınıf Betonarme inşaat (Ortalama Maliyet)							

Kaynak: Yazar tarafından derlenmiştir.

4.9.2. Sistem Yazılımının Geliştirilmesi

Sistem yazılımı geliştirilmesi esnasında kullanılan veri kaynaklarına hızlı bir şekilde ulaşmak amaçlanmıştır. Bu nedenle uygulamanın arka planında bulunan yazılım makine öğrenmesi tekniklerinde de efektif olarak kullanılabilen bir yazılım dili olması gerekmektedir. Bu sebeple yazılım dili olarak Python tercih edilmiştir. Çok sayıda Python dosyası ve veri setlerinden işlem yapılmasının kolay olması amacı ile Python için geliştirilen web frameworkü Flask tercih edilmiştir. Sistem geliştirilirken Python yazılım dili için birçok kütüphane içeren ve kolay bir kullanım sağlayan Anaconda platformu ve Anaconda ile birlikte gelen Spyder ve Jupyter Notebook kullanılmıştır. Veri görselleştirmek, veri işlemek, matematiksel işlemler yapmak gibi amaçlar için Python kütüphaneleri olan Pandas, NumPy, Math, Seaborn, Matplotlib vb. kütüphanelerden sıkça yararlanılmıştır.

Şekil 51: Flask Frameworkü Örnek Kod Bloğu

```
@app.route('/analytical', methods=['GET'])
def analytics():
    return render_template("analytical.html")

@app.route('/predict', methods=['GET'])
def predict():
    return render_template("predict.html")

@app.route('/outline', methods=['GET'])
def outline():
    return render_template("outline.html")

@app.route('/question', methods=['GET'])
def question():
    return render_template("question.html")

@app.route('/out', methods=['GET'])
def out():
    return render_template("out.html")

@app.route('/result', methods=['POST', 'GET'])
def samplefunction():
    output = {}
    if request.method == 'POST':
```

Kaynak: Yazar tarafından derlenmiştir

SONUÇ VE ÖNERİLER

Uygulama, bilgi sistemleri piramidine göre incelendiğinde, operasyonel açıdan kullanıcı kendisine sağlanan yapılandırılmış verilere hızlı ve etkin bir şekilde ulaşarak kısa dönemli kararlar alabilmektedir. Taktiksel açıdan, zaman serileri ile elde edilen tahminleme raporları ile alt düzey yöneticilere sağlanan yapılandırılmış veri sınırlarını belirleyen kararları verebilmektedir. Ayrıca, sektör trendlerini takip ederek üst düzey yöneticilere etkili ve anlamlı raporlar hazırlanabilir. Stratejik açıdan, sektör içinden ve sektöre etki edebilecek dış kaynaklardan elde edilen veriler ile ileriye yönelik, bölgesel, zamansal ve finansman bağlamında yatırım kararları alabilir.

Makine öğrenmesi ile ilan veri setine uygulanan algoritmalarda Random Forest algoritması %86 ile en başarılı sonucu vermiştir.

İş zekası çözümlerine gerek kamu gerekse özel kuruluşlar bazında ilginin gün geçtikçe daha çok arttığı günümüzde, iş zekasının farklı yönetim seviyelerinde karar destek amaçlı kullanımının işletmelere rekabet avantajı sağladığı görülmektedir. İş zekası yazılımlarının insandan kaynaklı hatalara yer vermeyen, yüksek performanslı yapısı işletmelerin bu tip sistemleri belirli operasyonlarına entegre etmesiyle sonuçlanmaktadır. Özellikle emlak sektöründe bu sistemlerin sıklıkla kullanılmaya başlanması, iş zekası uygulamalarına yeni bir soluk getirmektedir.

Bu anlamda iş zekası sistemlerinde makine öğrenmesi tekniklerinin kullanımı işletmelerin operasyonlarını gerçekleştirme şekillerinde de değişim yaratmaktadır. Bu yüzden işletmeler, hali hazırda kullandıkları sistemleri yeni bir teknolojik yapıya büründürmektedir.

Bu çalışma yönetim bilişim sistemleri kapsamında değerlendirildiğinde, farklı yönetim seviyelerine anlık raporlar sunması ve geleceğe yönelik tahminlemeleri içinde barındırması sebebiyle işletmeler için büyük önem taşımaktadır. Detaylı bir yaklaşımla incelendiğinde uygulamanın, Yönetim Bilişim Sistemleri alanına katkıları aşağıdaki gibi listelenebilir:

- Uygulama, bilgi sistemleri piramidine göre incelendiğinde, operasyonel açıdan kullanıcı kendisine sağlanan yapılandırılmış verilere hızlı ve etkin bir şekilde ulaşarak kısa dönemli kararlar alabilmektedir.

- Taktiksel açıdan, zaman serileri ile elde edilen tahminleme raporları ile alt düzey yöneticilere sağlanan yapılandırılmış veri sınırlarını belirleyen kararları verebilmektedir. Ayrıca, sektör trendlerini takip ederek üst düzey yöneticilere etkili ve anlamlı raporlar hazırlanabilir.
- Stratejik açıdan, sektör içinden ve sektöre etki edebilecek dış kaynaklardan elde edilen veriler ile ileriye yönelik, bölgesel, zamansal ve finansman bağlamında yatırım kararları alabilir.

Geliştirilen sistem şu anda prototip aşamasındadır. Ayrıca tez kapsamında geliştirilen uygulamanın YBS alanına katkısının süreklilik arz edebilmesi için, gelecek çalışmalarda, bu alanla ilgili yapılabilecek geliştirmeler şu şekilde listelenebilir.

- Yapay sinir ağı yöntemi kullanılarak tahminlemelerin yapılması,
- Farklı özellikler (özellikle konum verisi) kullanılarak modellerin oluşturulması,
- Daha fazla veri kaynağı elde edilerek daha farklı çapraz sorguların gerçekleştirilmesi,
- Çalışmanın farklı sektörlerde uygulanabilirliğinin değerlendirilmesi,

Sonuç olarak, geliştirilen uygulamayla, emlak sektöründe iş zekası amaçlı kullanılması bağlamında, farklı araştırmacılar için bir kaynak niteliğindedir. Gelecek bu konuyla ilgili çalışmaların artacağı beklenmektedir.

KAYNAKÇA

Abar, H. ve Karaaslan, A. (2013). Konut Talep Edenlerin Özellikleriyle Talep Edilen Konutun Özellikleri Arasındaki İlişkinin Çoklu Uyum Analizi Yöntemi ile İncelenmesi: Atatürk Üniversitesi Personeli Örneği. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 27(3): 323-339

Agrawal, D. (2008). The Reality Of Real-Time Business Intelligence. *International Workshop on Business Intelligence for the Real-Time Enterprise*. Springer, Berlin, Heidelberg, <http://birte08.stanford.edu/ppts/08-agrawal.pdf> Erişim Tarihi: 17.06.2019

Ahn, J.J., Byun, H.W., Oh, K.J., Kim, T.Y. (2012). Using Ridge Regression With Genetic Algorithm To Enhance Real Estate Appraisal Forecasting. *Expert Systems with Applications*. 39(9): 8369-8379.

Albayrak, A.S. (2005). Çoklu Doğrusal Bağlantı Halinde Enküçük Kareler Tekniğinin Alternatif Yanlı Tahmin Teknikleri Ve Bir Uygulama. *ZKÜ Sosyal Bilimler Dergisi* 1(1): 105-126

Alpaydın, E. (2004). *Introduction To Machine Learning*. United States of America: MIT Press.

Alpaydın, E. (2010). *Introduction to Machine Learning*. Massachusetts Institute of Technology, United States of America: MIT Press

Baars, H., Kemper, H. G., Siegel, M. (2008). Combining RFID Technology And Business Intelligence For Supply Chain Optimization Scenarios For Retail Logistics. *In Hawaii International Conference on System Sciences, Proceedings of the 41st Annual* (ss. 73-73), IEEE, Waikoloa, HI, USA. 7-10 Ocak 2008

Baldominos, Alejandro & Blanco, Iván & José Moreno, Antonio & Iturrarte, Rubén & Bernárdez, Óscar & Afonso, Carlos. (2018). Identifying Real Estate Opportunities Using Machine Learning. *Applied Sciences*. 8(11):2321

Bayraktar, E. (2007). *Bir İnsanlık Hakkı Konut: TOKİ'nin Planlı Kentleşme Ve Konut Üretim Seferberliği*. İstanbul: Boyut Yayıncılık.

Linoff, G.S. ve Michael, J.A. (2004). *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. Indiana: Wiley Publishing Inc

Bishop, C. M. (2006). *Pattern Recognition And Machine Learning*. Singapore: Springer.

Blumenfeld, H. (1944). A Neglected Factor in Estimating Housing Demand. *The Journal of Land & Public Utility Economics*, 20(3), 264-270.

Brohman, M.K. vd. (2000). The Business Intelligence Value Chain: Data-Driven Decision Support In A Data Warehouse Environment: An Exploratory Study. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. doi: 10.1109/HICSS.2000.926905.

Bulut, H., Öner, Y., İslamoğlu, E. (2015). The Investigation Of The Factors Affecting On The Prices Of Real Estates In Samsun Via Hedonic Price Model. *Alphanumeric Journal*, 3(2): 121–130

Burges, C. (1998). A Tutorial On Support Vector Machines For Pattern Recognition. *Data Min Knowl Discov* 2(2): 121–167

Chan, L. K., Sim, Y. W., & Yeoh, W. (2011). A SOA-Driven Business Intelligence Architecture. *Communications of the IBIMA*, 2011(216423): 1-8.

Chee, T., Chan, L., Chuah, M., Tan, C., Wong, S., & Yeoh, W. (2009). Business Intelligence Systems: State-Of-The-Art Review And Contemporary Applications. *Symposium on Progress in Information & Communication Technology* (s.s 96-101)

Cristianini N, Shawe-Taylor J. (2000). *An Introduction To Support Vector Machines And Other Kernel-Based Learning Methods*. Cambridge: Cambridge University Press

Cui, G., Wong, M. L. ve Lui, H. K. (2006). Machine Learning for Direct Marketing Response Models: Bayesian Networks with Evolutionary Programming. *Management Science*. 52(4): 597-612.

Daşkiran F. (2015), Denizli Kentinde Konut Talebine Etki Eden Faktörlerin Hedonik Fiyatlandırma Modeli İle Tahmin Edilmesi, *Uluslararası Sosyal Araştırmalar Dergisi*. 8(37): 850-857

Del Giudice, V., De Paola, P., Forte, F. (2017). Using Genetic Algorithms for Real Estate Appraisals. *Buildings*. 7, 31.

Deveci Kocakoç, İ., Erdem, S. (2010). Business Intelligence Applications in Retail Business: OLAP, Data Mining & Reporting Services. *Journal of Information & Knowledge Management*. 9(2): 171-181.

Durkaya, M. (2002). *Türkiye’de Konut Piyasasının Talep Yönlü Analizi*. (Yayınlanmamış Doktora Tezi). Trabzon: Karadeniz Teknik Üniversitesi, Sosyal Bilimler Enstitüsü.

Elaoud, A., Hassen, H.B., Salah, N.B. vd. (2017). Modeling of Soil Penetration Resistance Using Multiple Linear Regression (MLR). *Arabian Journal of Geosciences*. 10: 442.

Martins-Filho, C. ve Bin, O. (2005). Estimation of Hedonic Price Functions via Additive non- Parametric Regression. *Empirical Economics*. 30: 93–114.

Gerlein, E. A., McGinnity, M., Belatreche, A. ve Coleman, S. (2016). Evaluating Machine Learning Classification for Financial Trading: An Empirical Approach. *Expert Systems with Application*. 54: 193-207.

Gibson, M., Arnott, D., Jagielska, I., & Melbourne, A. (2004). Evaluating the intangible benefits of business intelligence: Review & research agenda. *In Proceedings of the 2004 IFIP International Conference on Decision Support Systems (DSS2004): Decision Support in an Uncertain and Complex World* (ss. 295-305). Prato, Italy.

Gollapudi, S. ve Laxmikanth, V. (2016). *Practical Machine Learning*. Birmingham, UK: Packt Publishing.

Gunn, S. (1998). Support Vector Machines For Classification And Regression. University of Southampton, ISIS Technical Report

Han, J. ve Kamber, M. (2001). *Data Mining: Concepts And Techniques*, San Francisco: Morgan Kauffman

Hardle, W. (1989). *Applied Nonparametric Regression*. Cambridge: Cambridge University

Haroon, D., (2017) . *Python Machine Learning Case Studies*. Berkeley:Apress.

Ishaya, T. ve Folarin, M. (2012). A Service Oriented Approach to Business Intelligence in Telecoms Industry. *Telematics and Informatics*, 29(3): 273-285.

İçli, G. (2008). *Konut Pazarlaması*. İstanbul: Beta Basım Yayım Dağıtım.

İlhan, B. (2008). *Türk İnşaat Sektörünün Girdi-Çıktı Analizi Ve İthalata Bağımlılığı*. (Yayınlanmamış Yüksek Lisans Tezi). İstanbul: İstanbul Teknik Üniversitesi.

İmir, E. (1986). *Çoklu Bağıntılı Doğrusal Modellerde Ridge Regresyon Yöntemiyle Parametre Kestirimi*. Eskişehir: Anadolu Üniversitesi Basımevi.

İpek, O. (2001). Ridge Regresyon Üzerine Bir Çalışma. *VI. Ulusal Ekonometri ve İstatistik Sempozyumu*. Çukurova Üniversitesi, Adana, 20-22 Eylül 2001.

Kartal, E. (2015). Sınıflandırmaya Dayalı Makine Öğrenmesi Teknikleri ve Kardiyolojik Risk Değerlendirmesine İlişkin Bir Uygulama. (Yayımlanmamış doktora tezi). İstanbul: İstanbul Üniversitesi Fen Bilimleri Enstitüsü.

Taylor, S.J. ve Letham, B. (2017). Forecasting at Scale. *The American Statistician*.

JamesG., Witten D., Hastie T., vd. (2013). Linear Regression. *An Introduction to Statistical Learning* (ss.59-126) New York: Springer

Kalelkar, M., Churi, P., ve Kalelkar, D. (2014). Implementation of Model-View-Controller Architecture Pattern for Business Intelligence Architecture. *International Journal of Computer Applications*, 102(12): 16-21.

Karagöz, Y., Eş, A. ve Yavuz, S. (2015). Nominal İlişki Ölçüleriyle Konut Satın Almaya Etki Eden Faktörlerin İncelenmesi: Bolu İli Örneği. *C. Ü. İktisadi ve İdari Bilimler Fakültesi Dergisi*, 16(1): 247-260.

Keleş, R. (1983). *100 soruda Türkiye’de şehirleşme, konut ve gecekond* (3. Baskı). İstanbul: Gerçek Yayınevi.

Köklü, K. (2018) , İş Analizi, İş Analistliği ve İş Zekası , Lectio Socialis July 2018, Volume 2, Issue 2, 121-142

Korhonen, K.T., ve Kangas, A. (1997). Application of nearest- neighbour regression for generalizingsample tree information. *Scandinavian Journal of Forest Research*. (12): 97-101

Lantz, B. (2013). *Machine Learning With R*. Birmingham: Packt Publishing.

Lee, T. H. (1963). Demand for Housing: A Cross-Section Analysis. *The Review of Economics and Statistics*, 45(2): 190-196.

Li, W., Meng, W., Zhao, Y., Xu, S. (2009). Study on the Risk Prediction of Real Estate Investment Whole Process Based on Support Vector Machine. *2009 Second International Workshop on Knowledge Discovery and Data Mining*. DOI: 10.1109/WKDD.2009.40.

Lin, Y. H., Tsai, K. M., Shiang, W. J., Kuo, T. C., & Tsai, C. H. (2009). Research On Using ANP To Establish A Performance Assessment Model For Business Intelligence Systems. *Expert Systems with Applications*, 36(2): 4135-4146.

Loshin, D. (2003). Getting Onboard with Emerging IT, Business Intelligence: The Savvy Manager's Guide, Morgan Kaufmann,

Martin, P. (1966). Aggregate Housing Demand: Test Model, Southern California. *Land Economics*, 42(4): 503-513.

Maslow, A.H. (1970). Motivation And Personality. New York: Harper and Row Publishers Inc.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). Foundations Of Machine Learning. Cambridge, London: The MIT Press.

Mu, J., Zhang, A., Wu, F. (2014). Housing Value Forecasting Based on Machine Learning Methods. *Abstract and Applied Analysis*. 2014(4):1-7

Niculescu-Mizil, A., Caruana, R. (2005). Predicting Good Probabilities With Supervised Learning. *Proceedings Of The 22nd International Conference On Machine Learning* (ss. 625-632). 07–11 August 2005, Bonn.

Ong, I. L., Siew, P. H., & Wong, S. F. (2011). A Five-Layered Business Intelligence Architecture. *Communications of the IBIMA*. (2011): 1-11.

Özkurt, H. (2007). Türkiye Ekonomisinde Konut Sektörü: Gelişimi Ve Alternatif Finansman Modelleri. *İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (1): 159-173.

Öztürk, N. ve Fitöz E. (2009). Türkiye’de Konut Piyasasının Belirleyicileri: Ampirik Bir Uygulama. *ZKÜ Sosyal Bilimler Dergisi*, 5(10): 21-46.

Padgett JE, DesRoches R. (2008). Three- Dimensional Nonlinear Seismic Performance Evaluation Of Retrofit Measures For Typical Steel Girder Bridges. *Eng Struct*. 2008(30): 1869- 1878.

Pantic, M., Rothkrantz, L.J.M. (2004). Facial Action Recognition for Facial Expression Analysis From Static Face Images. *IEEE Transactions on Systems, Man and Cybernetics, Part B Cybernetics*. 34(3): 1449–1461.

Park, B., Bae, J. (2015). Using Machine Learning Algorithms For Housing Price Prediction: The case of Fairfax County, Virginia Housing Data. *Expert Systems with Applications* 42(6): 2928-2934.

Pashardes, P. Savva, C. (2009) Factors Affecting House Prices in Cyprus: 1988-2008. Economics Research Centre. 3(1): 3-25

Pazarçeviren, S.Y., Zor, Ü., Gürbüz, F. (2015). İş Zekâsı: Kavramsal Çerçeve, Bileşenler ve İşleyiş. *Siyaset, Ekonomi ve Yönetim Araştırmaları Dergisi*. 3(1): 75–91.

Rafiei, M.H., Adeli, H. (2016). A Novel Machine Learning Model for Estimation of Sale Prices of Real Estate Units. *Journal of Construction Engineering and Management* 142(2): 04015066

Raschka, S., Julian, D., Hearty, J. (2016). *Python: Deeper Insights Into Machine Learning*. Birmingham, UK: Packt Publishing.

Sharer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of datawarehousing*, 5(4), 13-22.

Şeker, S.E. (2016). İş analitiği (Business Analytics). *YBS Ansiklopedi*. 3(4), 3-9

Şeker, S.E. (2013). *İş Zekası ve Veri Madenciliği WEKA ile*, İstanbul:Cinius Yayınları

Sharer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of datawarehousing*, 5(4), 13-22.

Shwartz, S. S. ve David, S. B. (2014). *Understanding Machine Learning From Theory To Algorithms*. New York: Cambridge University Press.

Simon, P. (2013). *Too Big to Ignore: The Business Case for Big Data*. North Carolina, USA: Wiley.

Smola, A.J. ve Schölkopf, B. (1998). *Learning with kernels*. Berlin, Germany: Department of Computer Science: Technical University

Tecim, V. (2015). “Sistem Geliştirme Yaşam Döngüsü”. <http://debis.deu.edu.tr/userweb/vahap.tecim/dosyalar/sgyd.pdf> (05.05.2018).

Tibshirani, R. (1996). Regression Shrinkage And Selection Via The Lasso. *J R Stat Soc B Methodol*. 58(1): 267- 288.

Vapnik, V. (1995). *The Nature Of Statistical Learning Theory*. New York: Springer.

Viswanathan, V. (2016). *R: Recipes for Analysis, Visualization and Machine Learning*. Birmingham, UK: Packt Publishing.

Wadhvani, R., Shukla, S. (2019). Analysis Of Parametric And Non-Parametric Regression Techniques To Model The Wind Turbine Power Curve. *Wind Engineering*. 43(3): 225-232.

Yayar R. Ve Gül D. (2014), Mersin Kent Merkezinde Konut Piyasası Fiyatlarının Hedonik Tahmini, *Anadolu Üniversitesi Sosyal Bilimler Dergisi*. 14(3): 87-99

Zocca, V., Spacagna, G., Slater, D. ve Roelants, P. (2017). *Python Deep Learning*. Birmingham: Packt Publishing.

Zou, H., Hastie, T. (2005). Regularization And Variable Selection Via The Elastic Net. *J R Stat Soc Series B Stat Methodology*. 67(2): 301- 320.

EKLER

Ek 1: Makine öğrenmesi için kullanılan model

```
trainALL=[]
valALL=[]
true_=[]
pred_=[]
kf = KFold(n_splits=5,shuffle=True,random_state=60)
k=1
for train_index, val_index in kf.split(x):
    print("FOLD :",k)
    x_train, x_val = x.loc[train_index,:], x.loc[val_index,:]
    y_train, y_val = y[train_index], y[val_index]

    reg=GradientBoostingRegressor(random_state=60,n_estimators=100,loss="ls",max_depth=10,min_samples_split=2,learning_rate=0.1)
    reg.fit(x_train,y_train)

    predicted_train=reg.predict(x_train).clip(min=0,max=None)
    predicted_val=reg.predict(x_val).clip(min=0,max=None)
    pred_.extend(predicted_val)
    true_.extend(y_val)

    train_MAE=mean_absolute_error(y_train,predicted_train)
    val_MAE=mean_absolute_error(y_val,predicted_val)
    train_MSE=mean_squared_error(y_train,predicted_train)
    val_MSE=mean_squared_error(y_val,predicted_val)
    train_R2=r2_score(y_train,predicted_train)
    val_R2=r2_score(y_val,predicted_val)

    print("MAE Score Train: ",train_MAE)
    print("MAE Score Val : ",val_MAE)
    print("*****")

    trainALL.append([train_MAE,train_MSE,train_R2])
    valALL.append([val_MAE,val_MSE,val_R2])
    k+=1

trainALL=np.array(trainALL)
valALL=np.array(valALL)
true_=np.array(true_)
pred_=np.array(pred_)

print("Train Results For GradientBoosting"+'-'*20)
print("Mean MAE Score: ",trainALL[:,0].mean())
print("Mean MSE Score: ",trainALL[:,1].mean())
print("Mean R2 Score: ",trainALL[:,2].mean())
```

Kaynak: Yazar tarafından derlenmiştir.