

T.C.
BEYKENT ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

**DOĞAL DİL İŞLEME İLE MAKİNELERİN KENDİ
DİLİNİ MODELLEMESİ**

Yüksek Lisans Tezi

Tezi Hazırlayan

Ali DAYAN

İstanbul, 2022

T.C.
BEYKENT ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

**DOĞAL DİL İŞLEME İLE MAKİNELERİN KENDİ
DİLİNİ MODELLEMESİ**

Yüksek Lisans Tezi

Tezi Hazırlayan

Ali DAYAN

1908020004

ORCID: 0000-0001-5924-2299

Danışman

Dr. Öğr. Üyesi Atınc YILMAZ

İstanbul, 2022

YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “**Doğal Dil İşleme İle Makinelerin Kendi Dilini Modellemesi** ” başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere uygun şekilde tarafımdan yazıldığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, yararlandığım eserlerin tamamının kaynaklarda gösterildiğini ve çalışmamın içinde kullanıldıkları her yerde bunlara atıf yapıldığını, patent ve telif haklarını ihlal edici bir davranışımın olmadığını belirtir ve bunu onurumla doğrularım. 04.02.2022

Ali DAYAN

T.C.
BEYKENT ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ MÜDÜRLÜĞÜ
TEZLİ YÜKSEK LİSANS SINAV TUTANAĞI

04/02/2021

Enstitümüz *Bilgisayar Mühendisliği* Anabilim Dalı *Bilgisayar Mühendisliği* Programı yüksek lisans öğrencilerinden 1908020004 numaralı *Ali DAYAN*'ın "*Beykent Üniversitesi Lisansüstü Eğitim – Öğretim Yönetmeliği*"nin ilgili maddesine göre hazırlayarak, Enstitümüze teslim ettiği "*Doğal Dil İşleme İle Makinelerin Kendi Dilini Modellemesi*" konulu tezini, Yönetim Kurulumuzun 28/12/2021 tarih ve 2021/45 sayılı toplantısında seçilen ve On-Line toplanan biz jüri üyeleri huzurunda, Beykent Üniversitesi Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 29. maddesinin 3. fıkrası gereğince 45 dakika süre ile Microsoft Teams programı aracılığıyla on-line olarak aday tarafından savunulmuş ve sonuçta adayın tezi hakkında "*OYBİRLİĞİ*" ile "*KABUL*" kararı verilmiştir.

İşbu tutanak, 2 nüsha olarak hazırlanmış ve Enstitü Müdürlüğü'ne sunulmak üzere tarafımızdan düzenlenmiştir.

DANIŞMAN
Dr. Öğr. Üyesi At*** YI***
(Beykent Üniversitesi)

ÜYE
Dr. Öğr. Üyesi Ed*** ŞA***
(Beykent Üniversitesi)

ÜYE
Dr. Öğr. Üyesi Öm*** ÇE***
(Milli Savunma Üniversitesi)

Adı ve Soyadı : Ali DAYAN
Danışmanı : Dr. Öğr. Üyesi Atıncı YILMAZ
Türü ve Tarihi : Yüksek Lisans Tezi, 2022
Alanı : Bilgisayar Mühendisliği
Anahtar Kelimeler : Doğal Dil İşleme, Doğal Dillerin Modellenmesi, Makinelerin Kendi Doğal Dil Modelini Oluşturması

ÖZ

DOĞAL DİL İŞLEME İLE MAKİNELERİN KENDİ DİLİNİ MODELLEMESİ

Doğadaki tüm canlılar birbirleri ile haberleşmek için çeşitli yöntemler kullanmaktadır. Dillerin oluşumu ve gelişimi hakkında hala yeterli bilgi bulunmamaktadır. Dillerin oluşumu ve doğal dil işleme süreçleri ile ilgili geçmişten günümüze kadar çeşitli çalışmalar yapılmış olup bu çalışmalar farklı metod, farklı analiz ve farklı yaklaşımlar içermektedir.

Makinelerin kendi dillerini üretebilmesi için yaklaşım geliştirilmesi, böylece insansı makinelerin oluşturulma sürecinde bu makinelerin de tıpkı insanların sahip olduğu gibi kendi dilleri ile iletişim kurabilme süreçlerine katkıda bulunulmak, doğal dillerin oluşumu sürecinde de dil bilimcilere yol göstermesi, evrişimli sinir ağlarını kullanarak canlıları ayırt etme ve sınıflandırma yetilerini, yeni sesler türetme yetisini kazanması hedeflenmiştir. Görüntü işleme algoritmalarının ses işlemede kullanılarak daha etkin ses analizlerinin yapılmasının sağlanması ve ses işlemede yeni model geliştirilmesi konularına odaklanılmıştır.

İnsanlar yüzyıllardır dillerin oluşumu ve gelişimi üzerine çalışmalar yapmaktadır. Diller sayesinde insanlar birbirleri ile etkileşimde bulunabilmektedir. Çalışma ile bu yetinin makinelere kazandırılması sağlanması için modeller oluşturulmuştur. Çalışmada var olan doğal dil işleme modelleri ile ses işleme modelleri kullanılmıştır. Öncelikle makineye günümüzde var olan alfabelerden hangisini kullanmak istediği karar verdirilmiş ardından da Yenilemeli Yapay Sinir Ağları, Mel Frekans Cepstral Katsayısı ve Dinamik Zaman Çözgü yaklaşımı

modellerinin birlikte kullanılması ile canlıların seslerine benzer sesler türetilmiş ve canlılar bu sesler ile adlandırılmıştır.

Çalışma sonucunda, makinelerin kendi dillerini üretebilmesi için yaklaşım geliştirilmiştir. Böylece insansı makinelerin oluşturulma sürecinde bu makinelerin de tıpkı insanların sahip olduğu gibi kendi dilleri ile iletişim kurabilme süreçlerine katkıda bulunulmuştur. Doğal dillerin oluşumu sürecinde de dil bilimcilere yol göstermesi hedeflenmiştir.



Name and Surname : Ali DAYAN
Supervisor : Assist. Prof. Dr. Atınç YILMAZ
Degree and Date : Master Thesis, 2022
Major : Computer Engineering
Keywords : Natural Language Processing, Natural Language Modelling,
Modelling Computer's Own Language

ABSTRACT

MODELLING THE MACHINES' OWN LANGUAGE WITH NATURAL LANGUAGE PROCESSING

All living things in nature use various methods to communicate with each other. There is still not enough information about the formation and development of languages. Various studies have been carried out on the formation of languages and natural language processing processes from the past to the present, and these studies include different methods, different analyzes and different approaches.

Developing an approach for machines to produce their own languages, thus contributing to the processes of these machines in the process of creating humanoid machines to communicate with their own language, just like humans have, guiding linguists in the formation of natural languages, and their ability to distinguish and classify living things using convolutional neural networks. It is aimed to gain the ability to generate new sounds. The focus is on providing more effective audio analysis by using image processing algorithms in audio processing and developing a new model in audio processing.

People have been working on the formation and development of languages for centuries. Thanks to languages, people can interact with each other. With the study, models have been created in order to enable this ability to be gained by machines. In the study, existing natural language processing models and sound processing models were used. First, the machine had to decide which alphabet it would like to use today, and then, by using Regenerative Artificial Neural Networks, Mel Frequency Cepstral

Coefficient and Dynamic Time Warp approach models together, sounds similar to the sounds of living things were derived and living things were named with these sounds.

As a result of the study, an approach has been developed so that machines can produce their own languages. Thus, in the process of creating humanoid machines, these machines also contributed to the process of communicating with their own language, just like humans have. It is aimed to guide linguists in the process of formation of natural languages.



İÇİNDEKİLER

Sayfa No.

ÖZ

ABSTRACT

TABLolar LİSTESİ	iii
ŞEKİLLER LİSTESİ	iv
KISALTMALAR	vi
SÖZLÜK.....	vii
GİRİŞ	1

BİRİNCİ BÖLÜM

MATERYALLER VE YÖNTEMLER

1. YAPAY ZEKÂ	9
1.1. Yapay Zekâ, Derin Öğrenme ve Makine Öğrenmesi	11
1.1.1. Yapay Zekâ	12
1.1.2. Makine Öğrenmesi	12
1.1.3. Derin Öğrenme	13
1.2. Yapay Zekâ Uygulamaları	13
1.2.1. Yapay Sinir Ağları	13
1.2.2. Bulanık Mantık.....	20
2. DOĞAL DİL İŞLEME	21
2.1. Doğal Dil İşleme Süreçlerinin Tarihçesi	21
2.2. Doğal Dil İşlemede Klasik Yaklaşım	22
2.2.1. Morfolojik ve Sözcüksel Analiz:.....	23
2.2.2. Sözdizimsel Analiz:.....	23
2.2.3. Anlamsal Analiz:.....	23
2.2.4. Söylem Entegrasyonu:.....	23
2.2.5. Pragmatik Analiz:.....	24
2.3. Yapay Zekâ Ve Doğal Dil İşleme Metodları	24
2.3.1. Anlamsal Ağ.....	24
2.3.2. Ayırıştırma Ağacı	25

3. SES İŞLEME.....	28
3.1. Ses Tanımlama Çalışmalarının Tarihçesi	28
3.2. Ses İşleme Modelleri	29
3.2.1. Gizli Markov Modeli.....	29
3.2.2. Dinamik Zaman Bükme Yaklaşımı	29
3.3. Ses Dosyalarını Tanımlama Ve Sınıflandırma.....	32

İKİNCİ BÖLÜM

UYGULAMA

1. UYGULAMANIN KENDİ DİLBİLGİSİ YAPISINI SEÇMESİ.....	35
2. SES DOSYALARINDA ÖZELLİK ÇIKARIMI	37
3. ÖN İŞLEM.....	44
4. MODELİN OLUŞTURULMASI.....	48
SONUÇ VE ÖNERİLER.....	70
KAYNAKÇA.....	74

TABLÖLAR LİSTESİ

	Sayfa No.
Tablo 1. Ağ Topolojisi Katman Özellikleri ve Değerleri	53
Tablo 2. Nöron Bırakma ile Ağ Topolojisi Katmanı Özellikleri ve Değerleri	62
Tablo 3. Evrişimli Yapay Sinir Ağları Destekli Ağ Topolojisi Katman Özellik ve Değerleri.....	64
Tablo 4. Uzun-Kısa Dönem Hafıza Destekli Ağ Topolojisi Katman Değer ve Özellikleri.....	66



ŞEKİLLER LİSTESİ

Sayfa No.

Şekil 1. Yapay Zekâ, Makine Öğrenmesi ve Derin Öğrenme Arasındaki İlişki	12
Şekil 2. İnsan Beynindeki Nöral Ağ Yapısı	14
Şekil 3. Yapay Sinir Ağlarının Tipik Bir Mimarisi	15
Şekil 4. McCulloch-Pitts Modeli	16
Şekil 5. Rosenblatt α -Perceptron Modeli.....	17
Şekil 6. Adaline Modeli.....	18
Şekil 7. Anlamsal Ağ.....	25
Şekil 8. Ayrıştırma Ağacı	26
Şekil 9. Doğal Dil İşlemenin Temel Yöntemleri.....	26
Şekil 10. Öklid Eşleşmesi.....	31
Şekil 11. Dinamik Zaman Bükme Yaklaşımı Algoritması.....	31
Şekil 12. Otomatik Ses Tanımlama Sistemleri Algoritması Örneği.....	32
Şekil 13. Konuşma Tanımlama Teknikleri.....	33
Şekil 14. Alfabe Seçim Algoritması	36
Şekil 15. Zamana Bağlı Ses Verisi	37
Şekil 16. Ses Dosyası Eşik Noktası Örneği	38
Şekil 17. Seslerin Başlangıç Noktaları	38
Şekil 18. Makine Öğrenmesinde Test Verisinin Eğitimi.....	39
Şekil 19. Veriden İstenilen Çıktının Elde Edilmesi.....	39
Şekil 20. Yapay Sinir Ağı.....	40
Şekil 21. Dalga Formu.....	41
Şekil 22. Frekans/Adım ve Genlik/Ses Şiddeti	42
Şekil 23. Gerçek Dünya Piyano Sesi	43
Şekil 24. Genlik-Zaman Grafiği	43
Şekil 25. Fourier Dönüşümü.....	44
Şekil 26. Fourier Dönüşümü ile Büyüklük-Frekans Grafiği	45
Şekil 27. Kısa Zaman Fourier Dönüşümü	45
Şekil 28. Büyüklük-Frekans Grafiği.....	46
Şekil 29. Frekans, Zaman, Büyüklük Spektrogramı.....	47
Şekil 30. Frekans, Zaman, Desibel Spektrogramı	47

Şekil 31. Mel Frekans Cepstral Katsayısı Oluşturma Algoritması.....	48
Şekil 32. Ses Verileri İçin Geleneksel Makine Öğrenmesi Akışı.....	49
Şekil 33. Mel Frekans Cepstral Katsayıları Grafiği	49
Şekil 34. Mel Frekans Cepstral Katsayıları	50
Şekil 35. Doğrultulmuş Lineer Birim	51
Şekil 36. Doğrultulmuş Lineer Birim Modelin Eğitilme Süreçleri	52
Şekil 37. Erken Durma	54
Şekil 38. Nöron Bırakma	55
Şekil 39. Evrişimli Yapay Sinir Ağları Mimarisi	56
Şekil 40. Tekrarlamalı Yapay Sinir Ağları Mimarisi	57
Şekil 41. Tekrarlamalı Yapay Sinir Ağı Katmanı	57
Şekil 42. Zamanla Geri Yayılım.....	58
Şekil 43. Tekrarlama Katmanı.....	58
Şekil 44. Tekrarlamalı Yapay Sinir Ağları ve Uzun Kısa Dönem Hafıza.....	59
Şekil 45. Uzun Kısa Dönem Hafıza.....	60
Şekil 46. Aşırı Uyumluk Sonrası Eğitim ve Hata Değerlendirmeleri	63
Şekil 47. Evrişimli Yapay Sinir Ağı İle Test ve Hata Değerlendirmesi	65
Şekil 48. Uzun-Kısa Dönem Hafıza Sonrası Test ve Hata Değerlendirmeleri.....	67
Şekil 49. Veri Setine Karşılık Ses Dosyalarının Üretilmesi	68
Şekil 50. Mel Frekans Cepstral Katsayısı İle Ses Karşılaştırması.....	69

KISALTMALAR

- AI** : Yapay Zekâ (Artificial Intelligence)
- CNN** : Evrişimli Sinir Ağları (Convolutional Neural Network)
- DTW** : Dinamik Zaman Çözüğü Yaklaşımı (Dynamic Time Wrapping)
- IDE** : Entegre Geliştirme Ortamı (Integrated Development Enviroment)
- LSTM** : Uzun Kısa Dönem Hafıza (Long-Short Term Memory)
- MFCC** : Mel Frekans Kepstral Katsayısı (Mel Frequency Cepstral Coefficient)
- NLP** : Doğal Dil İşleme (Natural Language Processing)
- NLU** : Doğal Dil Uygulamaları (Natural Language Applications)
- ReLU** : Doğrultulmuş Lineer Birim (Rectified Linear Unit)
- RNN** : Yenilemeli Yapay Sinir Ağları (Recurrent Neural Networks)
- WAV** : Ses dosya biçimi

SÖZLÜK

Doğal Dil İşleme : Doğal Dil İşleme, yaygın olarak NLP olarak bilinen yapay zekâ ve dilbilim alt kategorisidir.

Entegre Geliştirme Ortamı : Entegre geliştirme ortamı yazılımcıların hızlı ve rahat bir şekilde yazılım geliştirebilmesini amaçlayan, geliştirme sürecini organize edebilen birçok araç ile geliştirme sürecinin verimli kullanılmasına katkıda bulunan araçların tamamını içerisinde barındıran bir yazılım türüdür

Yapay Zekâ : Yapay zekâ makineler tarafından kullanılan zekadır.



GİRİŞ

Doğadaki tüm canlılar birbirleri ile haberleşmek için çeşitli yöntemler kullanmaktadır. Haberleşmede kullanılan en etkin yol ise konuşmadır. İnsanlarda konuşma becerisi dil ve yazının gelişmesi ile kazanılmıştır. Dil gelişimi uzun yıllar almakta, yıllar geçtikçe insanların birbirleri ile etkileşimleri sayesinde yeni kelimeler üretilir, insanlardan birbirlerine aktarılır ve bazı kelimeler de yok olur. Eğer dil olmasaydı, kültür, sanat ve bilim de olmazdı.

Doğal dil işleme, doğal dillerde yazılmış metin ve/veya konuşmaların işlenip makineler tarafından anlamlandırılması çalışmalarının bir bütünüdür. NLP çalışmaları, her bir doğal dil farklı özellikler barındırdığı için dillerin özelinde metotlar, analizler ve yaklaşımlar gerektirmektedir.

Makineler doğal dilleri işlerken denetimli, yarı denetimli ve denetimsiz modellemeler kullanılması gerekmektedir.

Doğal dil işleme süreçleri ile ilgili geçmişten günümüze kadar çeşitli çalışmalar yapılmış olup bu çalışmalar farklı metot, farklı analiz ve farklı yaklaşımlar içermektedir. 1950'lerde Alan Turing'in basitçe "Makineler düşünebilir mi?" diye sorduğu sorunun doğal dil işleme çalışmalarını başlattığı kabul edilmektedir. Alan Turing, Almanların Enigma makinesinin şifre algoritmasını çözmeye çalışan matematikçilerin en ünlülerinden biriydi. Şifre çözme amacı ile başlatılan çalışmalar, Turing'in prensiplerini oluşturduğu bilgisayar prototipleri olan bilgisayarlar, Boole cebirine dayanan veri işleme mantığı ile Makine Zekası kavramının oluşmasına sebep olmuştur. Alan Turing 1950 yılında günümüzde "*Turing Testi*"ni anlattığı "*Makine ve Zekâ*" bir adlı makale yayınlamıştır. Turing testi sayesinde makinelerin ne kadar zeki olduğu bulunmaktadır. İnsan karşısındakinin makine olmadığını anlarsa makine testi geçmiş sayılır.

1960'lı yıllarda Terry Winograd tarafından Massachusetts Teknoloji Enstitüsünde (MIT) SHRLU isimli doğal dilleri anlayan bilgisayar programı yazmıştır John Searle ise Searle'nin Çince Odası Tartışması deneyini başlatmıştır. Bu deneyde Çince girdiler manipüle edilerek, Çince bilgisi olamayan Searle tarafından yine Çince

çıktılar üretilir. Searle'nin deneyinin bazı karmaşıklıkları vardır. Searle hala Çince bilmezken Searle, oda ve kurallar bütünü Çinceyi anlayabiliyor ancak Çince bilen birisini bu deney sonucunda geçemezler. Çince konuşulan bir diyalogu taklit edemez.

İlk ses çalışmaları 1881 yılında Alexander Graham Bell tarafından keşfedilen ses kaydeden cihaz ile başlamıştır. 1930'lara gelindiğinde Bell Laboratuvarları konuşma analizi ve sentezi için bir sistem modeli önermiştir. 1952 yılında yine Bell Laboratuvarı tarafından Audrey isminde rakam tanıyabilen sistem geliştirilmiştir. 1960 yılında Gunnar Fant, konuşma üretimi için kaynak-filtre modelini geliştirmiştir. 1970'lerde artık çalışmalara DARPA da dahil olmuştur. 1000 kelimelik konuşma tanıma çalışması yapmıştır. 1990'lardan itibaren de çağrı merkezlerinin yaygınlaşması ile ses tanımlama çalışmaları daha da önem kazanmıştır. Çağrı merkezlerinde her bir çağrıya yanıt verecek yeterli eleman barındırmak firmalar için çok yüksek maliyetler getirmekte iken ses tanımlama yazılımları ücretleri ve yıllık bakım maliyetleri daha düşüktür. Artık çağrı merkezlerini aradığımızda veya internet bankacılığı kullandığımızda sesli yanıt sistemleri sayesinde para transferinden, fatura ödemeye kadar işlerimizin çoğunu çağrı merkezi elemanlarına ihtiyaç duymadan halledebilmekteyiz.

İnsanlar, canlıların dil becerilerini nasıl elde ettiğini ve elde edilen dil becerisinin gelecek teknolojilerde en etkin kullanma yöntemlerini bulmak için yüzyıllardır doğal dil işleme alanında çalışmalar yürütmektedir. Doğal dil işleme çalışmaları, makine öğrenmesi çalışmalarının artması, makine öğrenmesinde, doğal dil işleme yöntemlerinde ve yapay sinir ağlarında farklı yaklaşım ve teknikler geliştirilmesi, entegre geliştirme ortamlarının ve kütüphanelerinin kurulumlarının kolaylaşması, herkes tarafından kolayca erişilebilir olması sebebiyle popülerleşmiştir.

İnsanlar uzun yıllardır bazı sorular sormaktadır:

- İlk insanlar konuşmayı nasıl öğrendi?
- Makineler konuşabilir mi?
- Makineler kendine ait yeni bir dil geliştirebilir mi?
- Makineler canlıları taklit edebilir mi?
- Makineler sesleri ve canlıları sınıflandırabilir mi?

Bu çalışmada, insanların uzun yıllardır sorduğu bu sorulara yanıt aranmıştır. Makinelerin zamanla insanlar gibi kendi doğal dil yetilerine sahip olabileceğine, kendi aralarında bu dille iletişim kurabileceğine, insan dillerinin başka nesne ve canlıların çıkardığı sesleri taklit ederek kendi seslerini ve zamanla dillerini oluşturduğuna ve makinelerin de bu yolları izleyip, diğer diller ile etkileşim kurarak zenginleşebileceği konularına odaklanılmıştır. Özellikle, doğal dil işleme metotları kullanarak makinelerin kendi doğal dillerini oluşturması konusu incelenmiş ve doğal dillerin oluşum ve etkileşim süreçleri incelenerek, ses dosyalarının, görüntü işleme metotları ile işlenmesi ile geliştirilen doğal dil işleme metodu kullanılarak uygulama yapılmıştır.

Buzdolaplarının sütun bittiğini bilmesini ve diğer araçlar ile etkileşime geçmesini veya doğrudan sipariş vermesini, şehirlerdeki park metrelerin boş park yerlerini araçlara bildirmesini ve araçları boş park yerlerine yönlendirmesini sağlayabilen otonom sistemler geliştirilebilmesi için cihazların birbiri ile etkileşiminin kolaylaştırılması gerekiyor. İnsanlardan farklı olarak çok karmaşık düşünceler yerine basit bir kelime veya anahtar sözcük ile rahatlıkla iletişim kurabileceklerdir. Bizim için bu kelime bir şey ifade etmiyor olsa da makineler için anlamı büyük olacaktır. Böylelikle doğada oluşan olayları önceden kendi bilişleri ile analiz edebilecek, buzdolabımızda eksik olan ürünü kendisi sipariş edebilecek, arabamızla ile telefonumuz bizlere ihtiyaç duymadan kendi bilişleri ile iletişim kurabileceklerdir.

Makineler, bilişte sınırları olan insanların ikincil düşüncesine ihtiyaç duymadan, yeni geliştirdikleri dil sayesinde birbirleri ile konuşabileceklerdir.

Makine dilleri konuşmasına ve üretmesine izin verilen makineler, kendileri ile daha iyi iletişim kurmamıza ve hatta onları kontrol etmemize izin verebilir.

Çalışma için içerisinde anlaşılır ve çok çeşitli veri bulundurduğu için Kaggle Audio Cats and Dogs veri seti kullanılması kararlaştırılmıştır. Bu veri seti kedi ve köpek seslerinin çeşitli varyasyonlarını içermektedir.

Çalışmada, DTW, RNN ve MFCC algoritmaları kullanılarak veri setinin özellik çıkarımı ve sınıflandırılması hedeflenmiştir. Böylece makinenin canlıları ve sesleri sınıflandırma özelliğine sahip olması hedeflenmiştir.

Sınıflandırılan sesler için, ilk insanların sesleri taklit ederek yeni kelimeler üretmesi tezlerine dayanarak makinenin sesleri taklit etmesi hedeflenmiş ve başta belirtilen ilk dört sorunun cevabı aranmıştır.

Çalışma sonucunda doğada bulunan canlıların, ilk dillerde çıkardıkları seslere benzer sesler üretilmesi ile seslendirildiği ve diğer toplumlar ve diller ile etkileşime geçildikçe seslerin daha da anlaşılır hale geldiği gösterilmek istenmiş ve bu konuda dil bilime katkı sunulmak istenmiştir.

Bu çalışma sayesinde makinelerin kendi dillerini üretebilmesi için yaklaşım geliştirilmeye çalışılmış. Böylece insansı makinelerin oluşturulma sürecinde bu makinelerin de tıpkı insanların sahip olduğu gibi kendi dilleri ile iletişim kurabilme süreçlerine katkıda bulunulmak istenmiştir. Doğal dillerin oluşumu sürecinde de dil bilimcilere yol göstermesi hedeflenmiştir.

Bilgisayar bilimi alanında insansı makinelerin oluşturulmasında, bu makinelerin insana özgü olan kendi dilini geliştirme ve bu dille iletişim kurmasının sağlanması konularına katkıda bulunulmak istenmiştir. Aynı zamanda ses ve doğal dil işleme konularında yeni bir yaklaşım getirmek hedeflenmiştir. Ses işleme süreçlerinde yeni yaklaşım geliştirilmesi, makinelerin canlıları ayırt etme yetisine sahip olması, makinelerin canlıların seslerini taklit ederek onlardan öğrenebilme yetisine sahip olması, makinelerin kendi geliştirdikleri dil ile daha etkin iletişime geçebilmesini sağlamak, insansı makinelerin üretilmesini sağlama ve karmaşık problemlerin makineler tarafından kendi mukayese yetkinlikleri ile daha kolay çözülmesini sağlamak araştırmanın diğer başlıca hedefleri arasında yer almaktadır.

Çalışmanın ilerleyen bölümlerinde öncelikle literatür taraması yapılarak doğal dil işleme ve ses işleme süreçlerinde yapılmış çalışmalar incelenmiştir. İlerleyen bölümlerde doğal dil işleme süreçleri ve ses işleme süreçleri derinlemesine incelenmiş olup çalışma için ön hazırlık yapılmıştır. Sonraki bölümlerde ise makinenin kendi alfabesini seçmesi ve belirlenen veri setinden özellik çıkarımı yapılarak kedi ve köpek seslerine karşılık ilkel sesler üretilmesi için çalışmalar yapılmıştır.

Geçmiş Çalışmalar

Literatür taraması yapıldığında doğal dil işleme ve ses işleme yöntemleri üzerine birçok çalışma yapıldığı görülmektedir. Çalışmaların temelinde ses dalgalarının çözümlenerek konuşmadan yazıya dönüştürülmesi bulunmaktadır. Konuşmadan yazıya dönüştürülen ses dalgaları işlendikten sonra tekrar yazıdan konuşmaya döndürülmektedir. Çalışmalarda genellikle otomatik yazım hatası tanımlama, denetlenmesi ve düzeltilmesi, ses tanıma, sesli asistan geliştirme, diller arası çeviri gibi biçimsel, sözdizimsel ve anlamsal çözümleme yöntemlerine odaklanılmaktadır.

Muhammed Enes Almahdi, Doğal Dillerin Zaman ve Konuma Bağlı Değişimlerinin Nicel Olarak Ölçülmesi Yüksek Lisans Tezinde yaptığı çalışmada; Gözetimsiz öğrenmeyle paralel korpora kullanmadan kelime ikamesini ve diller arası benzerliği otomatik olarak tespit eden etkin bir sistem oluşturmuştur. Kelime ikamesini keşfetmek ve zamana dayalı kelime vektör uzayları ile zamandan bağımsız evrensel kelime vektör uzayının nasıl hizalanacağını öğretmek için basit ve etkin sözlük tabanlı doğrulama yöntemi ile çekişmeli eğitim prosedürünü kullanmıştır. Ayrıca, tek dilli metinler üzerinde eğitilmiş n-gram modellere dayanarak diller arası benzerliği kestirmiştir.

Ahmet Toprak, Doğal Dil İşleme ile İngilizce Otomatik Sözlük Oluşturma Yüksek Lisans Tezinde yaptığı çalışmada; verilen bir İngilizce dokümanı referans alıp Helmholtz Prensibi ve TF-IDF metriklerini uygulayarak, makale konusuna ait sözlüğün otomatik oluşturulmasını sağlamıştır. WordNet sözlüğünü kullanarak her anlamlı kelimenin oluşmuş sözlük ile benzerliğini hesaplamış. Benzerlik değerleri, belirli bir eşik değerinden yüksek olan anlamlı kelimeleri sözlüğe eklemiştir.

Mehmet Durna, Doğal Dil İşlemesi ve Makine Öğrenmesi ile Türkçe Haberlerde Olay Göstergesi Tespiti Yüksek Lisans Tezinde yaptığı çalışmada; olay göstergesi için kendi oluşturdukları Türkçe veri kümesini nasıl etiketlediklerini ve Türkçe haber metinleri için olay göstergesi tespit ve tür sınıflandırması yapan bir sistem tasarlamışlardır. Veri kaynağı olarak Türkçe haber sitelerinden alınmış haber içerisindeki kelimeleri kullanmışlardır. Farklı makine öğrenmesi algoritmalarını

kullanarak olay göstergesi tespiti, olay göstergesi türü sınıflandırması, realist değeri bulma ve ana olay tespiti için en başarılı modeli bulmaya çalışmıştır.

Berfu Büyüköz, Doğal Dil İşlemede Bağlam Odaklı Derin Öğrenme Yöntemlerinin Transfer Kapasitesinin Metin Sınıflandırması Problemi Üzerinden İncelenmesi Yüksek Lisans Tezinde yaptığı çalışmada; ELMo ve DistilBERT bağlamsal ve doğal dil sistemlerinin farklı girdileri modellemedeki başarısını ölçmek ve doğal dil işlemeyi hedef alan sistemlerin gerçek hayattaki kaynak çeşitliliğine ne kadar adapte olabildiğini görebilmek için, protesto haber metni ve kullanıcı yorumlarının ikili sınıflandırması senaryoları üzerinden kıyaslamıştır.

Deniz Gül Özcan, Doğal Dil İşleme Tekniklerinin İncelenmesi ve Seyahat-Turizm Sesli Asistanı için Türkçe Varlık İsmi Tanıma Aracı Geliştirilmesi Yüksek Lisans Tezinde yaptığı çalışmada; doğal dil işleme teknikleri kullanılarak seyahat sektörü için Türkçe seyahat asistanı geliştirilmesine odaklanılmıştır.

Mihail Duşçu, Makine Öğrenmesi ve Ses İşleme Kullanılarak Sosyal Ağlarda Duygu Analizi Yüksek Lisans Tezinde yaptığı çalışmada; Twitter mesajlarında polarite sınıflandırması için ses verilerine dayalı yeni bir tanım, çıkarma tekniği ve özelliklerin kullanımını geliştirmeye çalışmıştır.

Gökhan Deneç, Canlı Ses İşleme ve Geribesleme Kullanılarak Sempatik Titreşen Teller ile Sound Mass Tarzı Sesler Yaratmak Doktora Tezinde yaptığı çalışmada; birbirinden bağımsız olarak görülen üç müzikal kavramın sentezlenmesi ile temellendirilmiştir. Bu üç müzikal kavram ise “Sentetik Sesler”, “Örneklenmiş Sesler” ve “Ses çıkışının ses girdisini geri beslemesiyle ortaya çıkan seslerin yapıtaşı olan müzikler”.

Bilgi Gökem Yazgaç, Süne ve Kımıl Zararlılarının Ses İşleme Yöntemleri ile Sınıflandırılması ve Bir Gömülü Sistem Gerçeklemesi Yüksek Lisans Tezinde yaptığı çalışmada; tahıl ürünlerinde bulunan süne ve kımıl zararlılarının erken tespiti ve sınıflandırılması amacıyla çeşitli ses analiz ve işleme yöntemlerinden elde edilen verileri kullanan bir sistem modeli önerilmiştir. Doğrusal Öngörülü Kepstral Katsayıları, MFCC ve Çizgi Spektral Frekans yöntemlerinin farklı kombinasyonları kullanılmıştır.

Bahadır Özkan, Doğal Dil İşleme Yöntemleri ile Diyalog Niyet Sınıflandırması Yüksek Lisans Tezinde yaptığı çalışmada; TF-IDF, Doc2Vec ve BERT gibi geleneksel ve gelişmiş doğal dil işleme yöntemlerini kullanarak, sohbet robotu konuşmalarının niyet tahminini geliştirmeye odaklanmıştır.

Hüseyin Emre Altın, Kuantum Yapay Zekâ Algoritmalarının Analizi: Grover Araması ve Doğal Dil İşleme ile Kuantum Haberleşmesi Yüksek Lisans Tezinde yaptığı çalışmada; Kuantum tabanlı Grover arama algoritması, geleneksel doğal dil işleme, kuantum doğal dil işleme ve kriptografik iletişim ile arama ve doğal dil işleme ile ilgili problemleri analiz etmeye odaklanmıştır.

Fatih Demir, Doğal Dil İşleme Kullanarak Türkçe Yazılım Gereksinimlerinden Sınıf Diyagramları Oluşturma Doktora Tezinde yaptığı çalışmada; doğal dil işleme kullanarak yazılım gereksinimlerinden otomatik olarak UML sınıf diyagramı üreten sistem önermiştir.

Mehmet Cengiz, Doğal Dil İşleme Temelli Görsel Senaryo Taslağı Motoru Geliştirilmesi Doktora Tezinde yaptığı çalışmada; görsel senaryo taslağı içerik oluşturucularının geliştirilecek yeni bir araçla bu araçları kullanarak zaman kaybetmeleri yerine yeteneklerini yaratıcılığa yönlendirmelerini hedeflenmiştir. Aracın geliştirme süreci, Doğal Dil İşleme ve çıkarılan nesnelerin çizimi ve eylemleri/hareketleri bir görsel senaryo taslağı üzerinde olan iki ana aşama içermektedir.

Fatma Bozyiğit, Doğal Dil İşleme Kullanımı ile Nesne Tabanlı Analiz ve Kaynak Kod Değerlendirmesi Doktora Tezinde yaptığı çalışmada; doğal dil işleme tekniklerinden faydalanılarak Türkçe gereksinimleri Birleştirilmiş Modelleme Dili sınıf diyagramına dönüştüren bir otomatik kavram tanımlama modeli sunmuştur.

Onur Keklik, Doğal Dil İşleme Tekniklerini Kullanarak Otomatik Soru Üretme Yüksek Lisans Tezinde yaptığı çalışmada; doğal dil işleme ile otomatik soru üretimi için yeni bir kural tabanlı yaklaşım önermektedir.

Değer Ayata, Türkçe ve İngilizce Twitter Duygu Sınıflandırması İçin Makine Öğrenmesi ve Doğal Dil İşleme Tekniklerinin Uygulanması Yüksek Lisans Tezinde yaptığı çalışmada; Twitter verisi için duygu çözümlemesi probleminin çözümünde

makine öğrenmesi ve istatistiksel doğal dil işleme tekniklerinin beraber kullanımının verimliliğini çalışmıştır.

Özge Sevgili Ergüven, Doğal Dil İşlemede Semantik Gösterimlerin Sistematik Değerlendirilmesi Yüksek Lisans Tezinde yaptığı çalışmada; öğrenme tabanlı güncel çalışmalar ve kelime ve onun eşlik ettiği kelimeler arasındaki istatistiklerin sayılmasına dayanan eski sistemler arasındaki ilişkiyi araştırmış ve kelime gösterimleri üzerine odaklanmıştır.

İncelenen bu çalışmalar sonucunda, doğal dil işleme ve ses işleme metotlarının birlikte kullanımı ile doğadaki sesler ve üretilen seslerin RNN, MFCC ve DTW metotlarının kombinasyonları ile karşılaştırması yapılmış ve doğadaki seslere karşılık yeni ilkel sesler ve kelimeler üretilmiştir.

BİRİNCİ BÖLÜM

MATERYALLER VE YÖNTEMLER

1. YAPAY ZEKA

Yapay Zekâ (AI) kavramının literatürde birçok karşılığı bulunmaktadır. Oxford’a göre AI, bir teori ve diller arası çeviri, karar verme, ses tanıma, görsel tahmin gibi normalde insan zekâsı gerektiren görevleri gerçekleştirebilen bilgisayar sistemlerinin geliştirilmesidir (Saygılı, 2020). 1956 yılında Minsky, McCarthy, Rochester ve Shannon önderliğinde düzenlenen Dartmouth Konferansında, McCarthy tarafından önerilen Yapay Zekâ tanımı ilk defa kullanılmıştır (Yılmaz,2015; Saygılı, 2020).

1950’den önce zekâ üzerine araştırmalar ve elektronik gereçler üzerine uygulamaları olsa da sayısal zekânın varlığından net olarak söz edemeyiz. İlk olarak 1950 yılında Alan Turing “*Makine ve Zekâ*” adlı makalesinde “Makineler Düşünebilir mi?” sorusu ile makinelerin zekâyâ sahip olabilecek şekilde programlanabileceğini gündeme getirmiştir. Bu makalede makinelerin zeki olup olmadığının anlaşılabilmesi için “*Turing Testi*” adında bir yöntem de paylaşmıştır (Doğan, 2018).

Stuart J. Russell ve Peter Norvig, 2009 yılında yayınladıkları “Yapay Zekâ: Bir Modern Yaklaşım” adlı kitapta AI tanımını “İnsan gibi düşünen sistemler”, “İnsan gibi davranan sistemler”, “Rasyonel düşünen sistemler” ve “Rasyonel davranan sistemler” olmak üzere 4 başlığa ayırmışlardır (Acarer, 2021).

a) İnsan Gibi Davranma: Turing Test Yaklaşımı

Turing testi, Alan Turing tarafından zekânın tatmin edici operasyonel tanımı sağlamak için tasarlanmıştır. Bilgisayarların sahip olması gereken yetenekler şunlardır:

- Başarılı bir şekilde İngilizce iletişim kurabilmek için **doğal dil işleme (natural language processing)**,

- Bildiklerini ve duyduklarını saklamak için **bilgi temsili (knowledge representation)**,
- Yeni sonuçlar üretmek ve sorulara cevap verebilmek için saklanan bilgileri kullanan **otomatik akıl yürütme**,
- Yeni durumlara uyum sağlamak ve desenlere tahmin yürütebilmek için **makine öğrenmesi**.

Zekâ için fiziksel temas gerekmediği için Turing testi fiziksel temastan uzak durur ancak tamamen Turing testi video sinyallerini de içerir bu yüzden bilgisayarların şu iki disipline de ihtiyacı vardır:

- Nesneleri algılamak için **bilgisayar görüşü**,
- Nesneleri manipüle etmek ve tanımlamak için **robotik**.

Test uygulamasında, sorgulayıcı kişi sorularını hem bir insana hem de yapay zekâ ile programlanmış bir makineye, onları görmeden istediği tüm soruları yöneltir ve hangisinin makine hangisinin de insan olduğunu ayırt etmeye çalışır. Makine ve insan da kendilerinin insan olduğunu ispatlamaya çalışır. Sorular ve cevaplar klavye ile yazılır ve bir ekran vasıtası ile gösterilir (Yılmaz, Sinirsel Bulanık Mantık Modeliyle Kanseri Risk Analizi, 2015).

a) İnsan Gibi Düşünme: Bilişsel Modelleme Yaklaşımı

İnsan gibi düşünebilen sistemler oluşturabilmek için öncelikle insan düşünmesinin çözülmesi gerekmektedir. İnsan düşünmesini çözmek için 3 yöntem vardır: iç gözlem, psikolojik deneyler veya beyin görüntülemesi. Başarılı olduğu takdirde bilgisayar program teorisini açıklamak mümkün olabilecektir.

b) Rasyonel Düşünme: Düşünme Kuralları Yaklaşımı

Rasyonel sistemler mantıksal yaklaşımlara dayanmaktadır. Bu yöntem belli kurallar altında sonuca ulaşmak için çıkarımlar yapılmasını hedeflemektedir. Örneğin; Sokrates bir insandır, tüm insanlar ölümlüdür, o zaman Sokrates bir ölümlüdür.

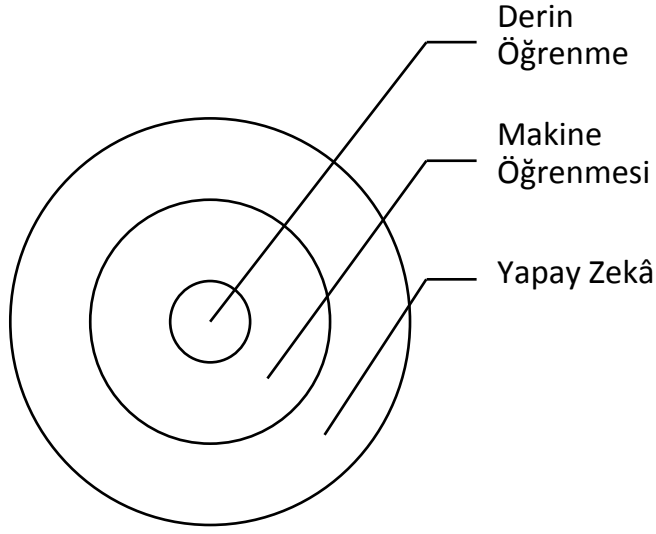
c) Rasyonel Davranma: Rasyonel Etken Yaklaşımı

Bir temsilci (İngilizce: agent, Latince: agere, yapmak anlamına gelmektedir) eylemleri olan bir şeydir. Tüm programlar bir şeyler yaparken bilgisayar temsilcisinin daha fazla şey yapması beklenir. Rasyonel temsilci de en iyi sonuca veya kesinlik olmadığı durumlarda beklenen en iyi sonuca ulaşmaya çalışır. Doğru çıkarımlar yapmak için bazen rasyonel davranmak gerekse de doğru çıkarımlar her zaman rasyonel değildir. Ancak hala bir şeylerin yapılması, tamamlanması beklenir. Örneğin sıcak sobaya tekrar kömür atmak üzerine düşünülecek bir olay değil bir reflektir. Bundan dolayı rasyonel davranma her zaman tek başına yeterli olmayabilir (Stuart, 2009).

Massachusetts Teknoloji Enstitüsü (MIT) Bilgisayar Bilimleri laboratuvar yöneticisi Edward Fredkin, yapay zekâ için “Tarihte üç büyük olay vardır. Bunlardan ilki kâinatın oluşumudur. İkincisi yaşamın başlangıcının olmasıdır. Üçüncüsü de yapay zekânın ortaya çıkışıdır.” demiştir (Pirim, 2006).

1.1. Yapay Zekâ, Derin Öğrenme ve Makine Öğrenmesi

Yapay zekâ, derin öğrenme ve makine öğrenmesi kavramları birbiri ile sıklıkla karıştırılan kavramlardır. *Şekil 1*'de Derin Öğrenme, Makine Öğrenmesi ve Yapay Zekânın arasındaki ilişki gösterilmiştir. Burada görüleceği üzere Derin Öğrenme ve Makine Öğrenmesi, Yapay Zekânın birer dalıdır.



Şekil 1. Yapay Zekâ, Makine Öğrenmesi ve Derin Öğrenme Arasındaki İlişki

1.1.1. Yapay Zekâ

Yapay Zekâ, Makinelerin veya makine kontrollü bir nesnenin bazı faaliyetlerini insan, maymun, yunus, arı gibi zeki canlılara benzer şekilde yerine getirme kabiliyetine sahip olmasıdır.

Yapay zekâ çalışmalarında genellikle insanın düşünme biçimi veya arıların sürü halinde hareket edebilme yetenekleri taklit edilmeye çalışır. Yapay zekâ çalışmaları ile hedeflenen öğrenebilen ve insan zekasından bağımsız gelişebilen yapay zekâ metodolojilerini geliştirmektir.

1.1.2. Makine Öğrenmesi

Makine Öğrenmesi, makinelerin veriler aracılığı ile öğrenmesini ve görevleri belirgin talimatlar olmadan yapma kabiliyetini kazanmasına olanak sağlar. Makine öğrenmeli çoğu sistem ham veriden öğrenme yeteneğine sahip değildir. Verinin özelliklere dönüştürülmesi gerekmektedir. Bu süreç özellikleştirme olarak adlandırılır ve veriyi alakasız içerikten arındırarak sistemin anlayabileceği ve kullanabileceği yapıya dönüştürür.

Makine Öğrenmesinde veri tek başına yeterli değildir. Yazılımın diğer alanlarındaki çöp toplama işlemi makine öğrenmesinde daha da değerlidir. İşe yaramayan petabyte boyutunda veri ile uğraşılmasının önüne geçilmesi gerekmektedir (Lanzetta, 2018).

1.1.3. Derin Öğrenme

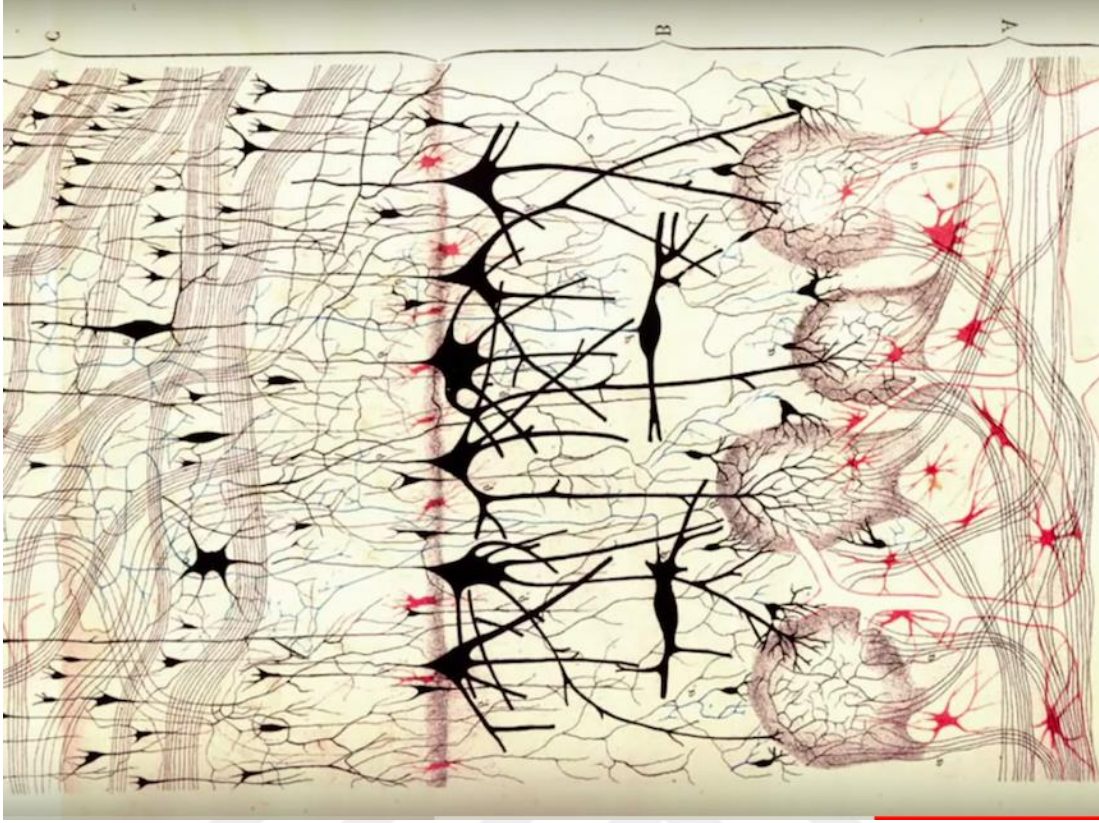
Derin Öğrenme sayesinde Makine öğrenmesinde, Geri Besleme, Stokastik Gradyan İnişi, Lineer Olmayan Aktivasyon Fonksiyonları ve bunların kombinasyonları kullanılarak daha karmaşık lineer olmayan fonksiyonlara göre daha az kaynakla derin ağların eğitilmesine olanak sağlanır.

1.2. Yapay Zekâ Uygulamaları

1.2.1. Yapay Sinir Ağları

İnsanlar tarafından kolaylıkla gerçekleştirilen, desen tanıma ve zekâ gerektiren işleri otomatikleştirmek oldukça zordur. İnsanlar çevrelerindeki çeşitli nesneleri tanıyıp, çeşitli bilgiler çerçevesinde büyük anlamlar çıkarabilmektedir. Makinelerin ise fizik sınırları içerisinde insanların bunları nasıl ağıladığını çözüp simüle etmesi gerekmektedir. Bu da sinir ağlarının araştırılması gerekliliğini doğurmuştur (Maind, 2014).

Binlerce nöronun birbirine bağlanması ile oluşan sinir ağları, insan sinir sisteminin bir parçasını oluşturmaktadır. ***Şekil 2***'de insan beynindeki nöral ağ yapısı gösterilmiştir.

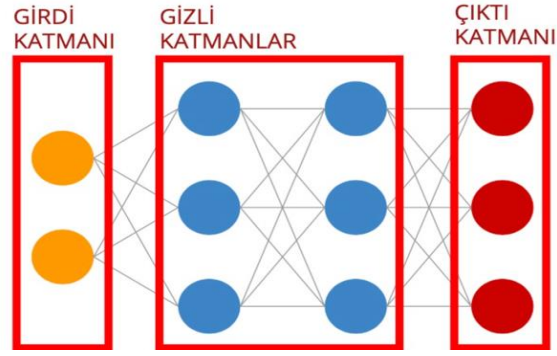


Şekil 2. İnsan Beynindeki Nöral Ağ Yapısı

Son zamanlarda yapay sinir ağları (artificial neural networks), birçok disiplinde sınıflandırma, kümeleme, örüntü tanıma ve tahmin için oldukça popüler ve yararlı bir model haline gelmiştir. Yapay sinir ağları, bir tür makine öğrenmesi (machine learning) modelidir. Geleneksel regresyon ve istatistiksel modellerle de nispeten rekabetçi bir hale gelmiştir (Dave & Dutta, 2014).

Yapay sinir ağların avantajı, büyük girdili karmaşık doğal sistemlerden daha kesin ve uygulaması kolay modeller yaratılabilmesidir (Jahnavi, 2017).

Genel olarak yapay sinir ağları insan beyninin bir kopyasıdır. **Şekil 3**'te yapay sinir ağlarının tipik bir mimarisi gösterilmiştir.



Şekil 3. Yapay Sinir Ağlarının Tipik Bir Mimarisi

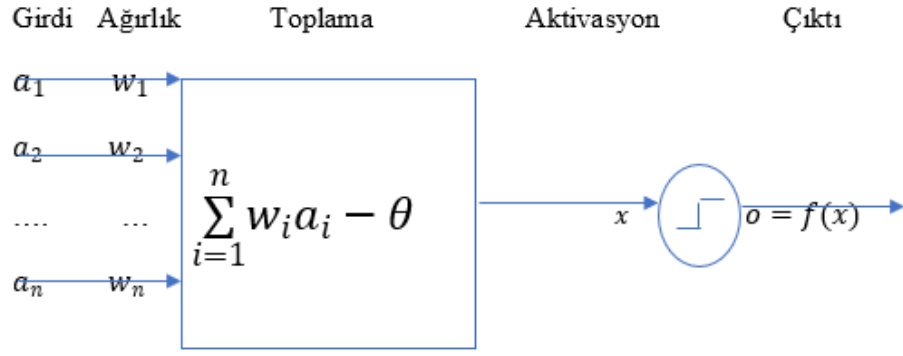
Yapay sinir ağı oluşturulması için tipik bir insanda bulunduğu gibi nöron yapısı oluşturulması gerekmektedir. Bu yüzden yapay sinir ağlarında nöron modelleri geliştirilmiştir.

Nöron Modelleri

Nöron modelleri oluşturulurken genellikle üç klasik model tercih edilmektedir. Bu modeller McCulloch-Pitts Modeli, Perceptron Modeli ve Adaline Modelidir.

McCulloch-Pitts Modeli

McCulloch-Pitts Modelinde, aktivasyon değeri (x), girdilerin (a), ağırlıkları (w) ile çarpımlarının toplamından öngörü (θ) değerinin çıkarılmasından meydana gelmektedir. *Şekil 4*'te McCulloch Pitts Modeli gösterilmiştir. *Şekil 4*'te verilen Toplama Denklemi, *Denklem 1*'de gösterilmiştir.



Şekil 4. McCulloch-Pitts Modeli

Denklem 1'de her bir ağırlık değeri (w) ile girdi değerinin (a) çarpımlarının öngörü değerinden (θ) çıkarılmasının birbiri ile toplanması ile oluşan McCulloch-Pitts Modelinde kullanılan denklemi verilmiştir.

$$\sum_{i=1}^n w_i a_i - \theta$$

Denklem 1. McCulloch-Pitts Modeli Toplama Denklemi

Denklem 2'de McCulloch-Pitts Modelinin çıktı denklemi verilmiştir. **Denklem 2**'de görüldüğü gibi ikili düzende çıktı sonucu 0'dan büyük ise 1 0'dan küçük veya eşit ise de 0 değeri alabilir. Çıktı sinyali (o), aktivasyon değerinin (x), lineer olamayan bir ikili (binary) fonksiyonudur ($f(x)$).

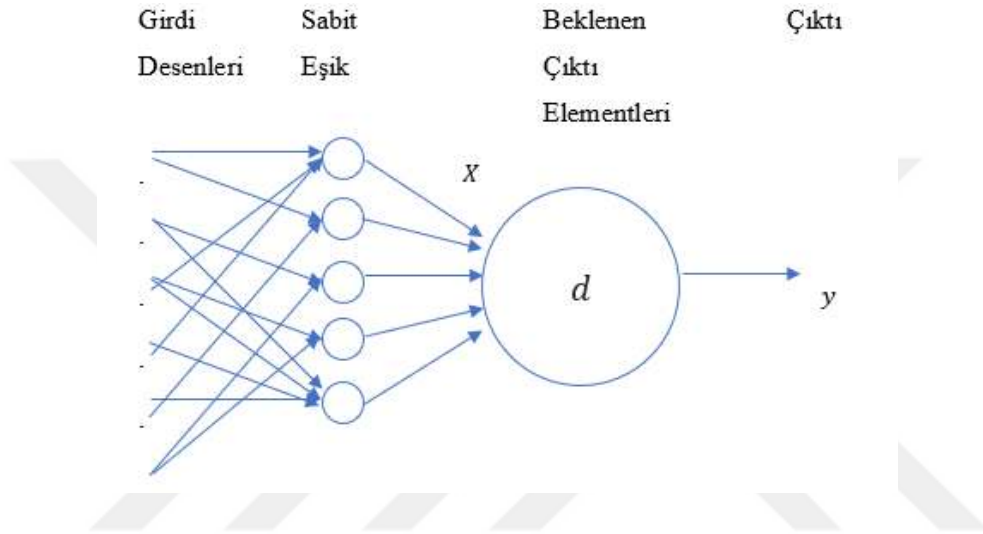
$$f(x) = \begin{cases} 1 & \text{eğer } x > 0 \\ 0 & \text{eğer } x \leq 0 \end{cases}$$

Denklem 2. İkilik sistemde çıktı

Bu modelde ağırlıklar sabittir. Bu sebeple makinenin öğrenme yeteneği bulunmamaktadır (Yegnanarayana, 2009).

Rosenblatt α -Perceptron Modeli

Şekil 5'te gösterilen Rosenblatt modelinde rastgele verilmiş sabit ağırlıklar ile çarpılan girdi desenleri rastgele bağlantılar ile ikinci katmanı beslemektedir. İkinci katman tek doğrusal eşik elemanından meydana gelmektedir. Eşik elemanlarından elde edilen çıktı beklenen çıktı ile karşılaştırılır ve bir niceleyici hata $\tilde{\varepsilon}$ elde edilir.



Şekil 5. Rosenblatt α -Perceptron Modeli

Perceptron modelinin uygulanması, niceleyici hatanın $\tilde{\varepsilon}$ kullanımı ile olur.

Niceleyici hata $\tilde{\varepsilon}$, beklenen çıktı (d) ile çıktısının (y) farkının bulunması ile **Denklem 3**'te verildiği şekilde hesaplanır.

$$\tilde{\varepsilon}_k \triangleq d_k - y_k$$

Denklem 3. Perceptron Modeli niceleyici hata denklemi

$\tilde{\varepsilon}_k = 0$ gibi y_k 'nin doğru olduğu durumlarda ağırlıklar güncellenmez. Eğer y_k ve d_k birbirinden farklı ise ağırlıklar $\tilde{\varepsilon}_k$ ile güncellenir. Eğer $\tilde{\varepsilon}_k > 0$ ise eklenir. $\tilde{\varepsilon}_k < 0$ ise çıkartılır.

Denklem 4'te niceleyici hata ($\tilde{\varepsilon}$), ağırlık ($\tilde{W}$), düzeltme katsayısı ($\tilde{\mu}$) ve girdi ($\tilde{X}$) olmak üzere Perceptron Modelinin ağırlık hesaplama denklemi verilmiştir.

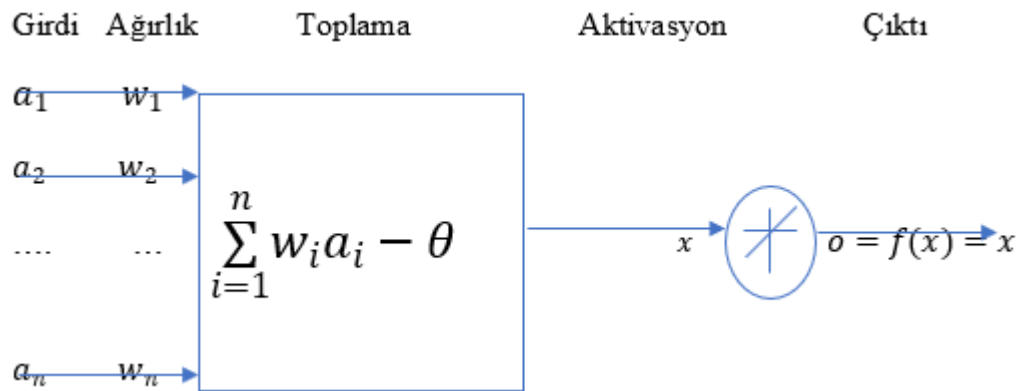
$$\tilde{W}_{k+1} = \tilde{W}_k + 2\tilde{\mu} \frac{\tilde{\varepsilon}_k}{2} \tilde{X}_k$$

Denklem 4. Perceptron Modeli ağırlık hesaplama denklemi

Rosenblatt, **Denklem 4**'te verilen μ değerini 1 olarak almıştır. Rosenblatt'ın modeli sadece ikili sayı sistemi ile cevap beklenen sistemler için kullanılabilir (WIDROW & LEHR, 1995).

Adaline Modeli

ADaptive LINear Element (Uyarlamalı Doğrusal Element) kelimelerinden uyarlanan Adaline, Widrow tarafından geliştirilen bir hesaplama modelidir. Perceptron modeline benzemekle beraber ana farkı doğrusal fonksiyon kullanmasıdır.



Şekil 6. Adaline Modeli

Şekil 6'da Adaline Modeli gösterilmiştir.

Adaline modelinde analog aktivasyon değeri (x), beklenen çıktı (o) ile karşılaştırılır.

Denklem 5'te ağırlık değeri (w), girdi değeri (a) ve öngörü değeri (θ) olmak üzere Adaline Modelinde kullanılan toplama denklemi verilmiştir.

$$\sum_{i=1}^n w_i a_i - \theta$$

Denklem 5. Adaline Modeli Toplama Denklemi

Denklem 6'da beklenen çıktı (o) ve toplama denklemi sonucu (x) olmak üzere bir hata (ε) elde edilir.

$$\varepsilon = o - x$$

Denklem 6. Adaline Modeli hata denklemi

Denklem 7'de katsayı (μ), hata (ε), girdi değeri (a) olmak üzere ağırlık hesaplama denklemi verilmiştir.

$$\Delta w_i = \mu \varepsilon a_i$$

Denklem 7. Adaline Modeli ağırlık hesaplama denklemi

Ağırlıkların her biri öğrenme katsayısı μ ve hata ε ile girdinin a_i çarpılması ile güncellenir. Ağırlık güncellemesi ortalama kare hatasını μ^2 azaltır. Bu da Least Mean Square (LMS- En Küçük Ortalama Kare) hata öğrenme kuralıdır. Bu kural, ağırlık uzaydaki hata yüzeyinin negatif gradyanı kullanılarak elde edildiği için gradyan iniş algoritması olarak da bilinmektedir (Hekimoğlu, 2010).

1.2.2. Bulanık Mantık

İnsanlar gündelik yaşamlarında sorunlar ile karşılaşırken bu sorunları geçmişte edindikleri bilgi ve tecrübelerinden yararlanarak çözmektedir. Sorunların bir kısmına kolayca çözümler üretebilirken bir kısmı belirsizlikler içermekte ve çözümü ise nispeten zor ve görecelik içermektedir (Yılmaz, Sinirsel Bulanık Mantık Modeliyle Kanser Risk Analizi, 2015).

Gerçek dünyada, nesne sınıfları genellikle belirli bir kritere sahip değildir. Örneğin köpekler ve atlar hayvan sınıfına aitken, taşlar ve bitkiler kesinlikle hayvan değildir. Ancak bakteriler ve deniz yıldızları gibi nesneler hayvan sınıfına göre belirsiz durumdadırlar. 1'den çok büyük tüm reel sayılar sınıfı, güzel kadınlar sınıfı veya uzun erkekler sınıfı matematiksel anlamda alışılmış kümeler veya sınıflar oluşturmamaktadır. Yine de insan düşünmesinde sınıflar kesin olarak önemli bir rol oynamaktadır (Lanzetta, 2018).

Bulanık mantık, klasik mantıktaki doğru ve yanlış (0-1) değerleri arasındaki kısmen ifadelerini de kapsayacak şekilde ikili mantık kümesinin genişletilmiş bir üst kümesidir (Erdal, 2008).

2. DOĞAL DİL İŞLEME

Doğal dil işleme (**natural language processing**), yapay zekâ ile programlanmış yazılımlar sayesinde, konuşma ve yazma gibi dile özgü işlevleri analiz ederek bilgiye dönüştürülmesini amaçlayan, Dil Bilimi ile yapay zekâyı birleştiren Bilgisayar Bilimleri dalıdır (Yılmaz, 2020). Bu sayede makineler ile makine dilinde konuşulabilmektedir. Tezin de amacı, makineler ile makine dilinde konuşmak yerine, makinelerin, insanların geliştirdiği biçimde kendilerine özgü doğal dil geliştirmeleri ve böylece insan dilinin de oluşumuna ışık tutulmasıdır.

Doğal dil işlemede, işleme kelimesi anlama kelimesi yerine kasıtlı olarak kullanılmıştır. Yapay zekanın ilk yıllarında doğal dil işleme yerine doğal dil anlama ifadesi kullanılırken, bugün doğal dil işlemenin asıl amacının doğru doğal dil anlama olduğu kanısına varılmıştır (Liddy, 2001). Tam bir doğal dil anlama sistemi aşağıdaki yeterliliklere sahip olmalıdır.

1. Girdiyi başka kelimeler ile ifade edebilme
2. Metni başka bir dile çevirebilme
3. Metin içeriğindeki sorulara yanıt verebilme
4. Metinden çıkarımlar yapabilme

Doğal dil işleme süreçleri ilk üç maddede bayağı yol kat etmişken hala kendi başına bir metinden çıkarımlar yapabilme yeteceğine sahip değildir.

Bu tez ve gelecekteki çalışmalar sayesinde kendi dilini üretebilecek hale gelebilecek bir makine, metinden çıkarımlar yapma yetisine de sahip olabilecektir.

2.1. Doğal Dil İşleme Süreçlerinin Tarihçesi

1950’lerde Alan Turing’in basitçe “Makineler düşünebilir mi?” diye sorduğu sorunun doğal dil işleme çalışmalarını başlattığı kabul edilmektedir. Alan Turing 1950 yılında günümüzde “*Turing Testi*”ni anlattığı “*Makine ve Zekâ*” bir adlı makale

yayınlamıştır. Turing testi sayesinde makinelerin ne kadar zeki olduğu bulunmaktadır. İnsan karşısındakinin makine olmadığını anlarsa makine testi geçmiş sayılır.

1960'lı yıllarda Terry Winograd tarafından Massachusetts Teknoloji Enstitüsünde (MIT) SHRLU isimli doğal dilleri anlayan bilgisayar programı yazmıştır (Abhimanyu Chopra, 2013). John Searle ise Searle'nin Çince Odası Tartışması deneyini başlatmıştır. Bu deneyde Çince girdiler manipüle edilerek, Çince bilgisi olmayan Searle tarafından yine Çince çıktılar üretilir. Searle'nin deneyinin bazı karmaşıklıkları vardır. Searle hala Çince bilmezken Searle, oda ve kurallar bütünü Çinceyi anlayabiliyor ancak Çince bilen birisini bu deney sonucunda geçemezler. Çince konuşulan bir diyalogu taklit edemez.

2.2. Doğal Dil İşlemede Klasik Yaklaşım

Doğal dil işleme, temelde verilerin büyüklüğü, kuralsız ve anlaşılmasız konuşma, kuralsız ve bozuk yazılar, sözcük niteliklerindeki belirsizlikler, çok uzun kelime bulunan dillerde kelimeleri dilimleme ve kelimelerin gündelik dilde dahi kullanım anlamlarının çokça değişmesi gibi insan dillerindeki belirsizlikler nedeni ile çok zor olsa da insan dilindeki süreçleri inceleyip modelleme çalışır. Modelleme çalışmalarında, doğal dil işleme süreçleri çeşitli görev ve parçalara bölünür ve ardından ayrı ayrı tekrar birleştirilirler. Basit görüş, bir metnin tümcelerinin önce sözdizimleri açısından analiz edilmesidir; bu, semantik veya gerçek anlam açısından bir analize daha uygun bir düzen ve yapı sağlar ve bunu, sözcenin veya metnin bağlam içindeki anlamının belirlendiği bir pragmatik analiz aşaması takip eder (Indurkha & Damerau, 2010). Klasik yaklaşımda doğal dil işleme süreçleri beş aşamada incelenebilir:

1. Morfolojik ve Sözcüksel Analiz
2. Sözdizimsel Analiz
3. Anlamsal Analiz
4. Söylem Entegrasyonu

2.2.1. Morfolojik ve Sözcüksel Analiz

Kelimeler doğal dillerin yapı taşlarıdır. Morfoloji kelimelerin yapısını analiz eder, tanımlar ve açıklar. Sözcüksel analiz ile metinler paragraflara, kelimeler de cümlelere bölünür. Tüm bunlar sesbilim ile mümkün kılınmaktadır. Sesbilim birtakım bileşenlerden meydana gelmektedir. Bu bileşenlerin kullanılmayıp sesbilimin dil bilim ile birleşmesine morfofonoloji veya morfoloji denir.

2.2.2. Sözdizimsel Analiz

Bir cümledeki kelimelerin analizini yaparak cümlenin dilbilgisi yapısını tasvir eder. Kelimelerin birbirleri ile olan ilişkileri incelenir. Kelimelerin bir cümle oluşturmak için nasıl sıralanması gerektiği belirlenir. Türkçe’de devrik cümleler olduğu için bunun tespiti daha zorken İngilizce veya Almanca gibi dilbilgisi yapısında katı kurallar olan dillerde bunların tespiti kolaydır. Örneğin “Das Mädchen das geht das zur Schule (Okula giden kız)” ve “Das Mädchen zur Schule geht (Kız okula gidiyor)” gibi cümleler Almanca sözdizimsel analizden geçmeyecektir.

2.2.3. Anlamsal Analiz

Cümle yapısı incelenir ve sözlük anlamı ile gerçek anlamları ayrıştırılır. Dillerde kelimeler sözlük anlamlarından farklı kullanılır ve bunlar anlamsal analizciler ile ayırt edilir. Analizciler anlamsal yapılar üretirler. Kelimeler ve bu yapılar arasında eşleşme oluştururlar. Bu makinelerin çalışma mantığına aykırı bir süreçtir. Örneğin “ışksız lamba odayı aydınlattı” böyle bir cümle anlamsal analizci tarafından reddedilecektir. Işıksız lamba olmayacağı gibi ışksız oda da aydınlatılamayacaktır.

2.2.4. Söylem Entegrasyonu

Cümleler tek başına anlamlı olmayabilirler. Bu yüzden cümlelerin anlamları değerlendirilirken paragraf içerisindeki önceki cümleler de analiz edilir. Örneğin bir cümlede geçen “onun” kelimesi önceki cümlelerde muhtemelen bir canlıyı işaret

ediyordur. Makineler de anlamlandırmak için önceki cümleleri de analiz etmek ve bir bütün halinde değerlendirmek durumundadır.

2.2.5. Pragmatik Analiz

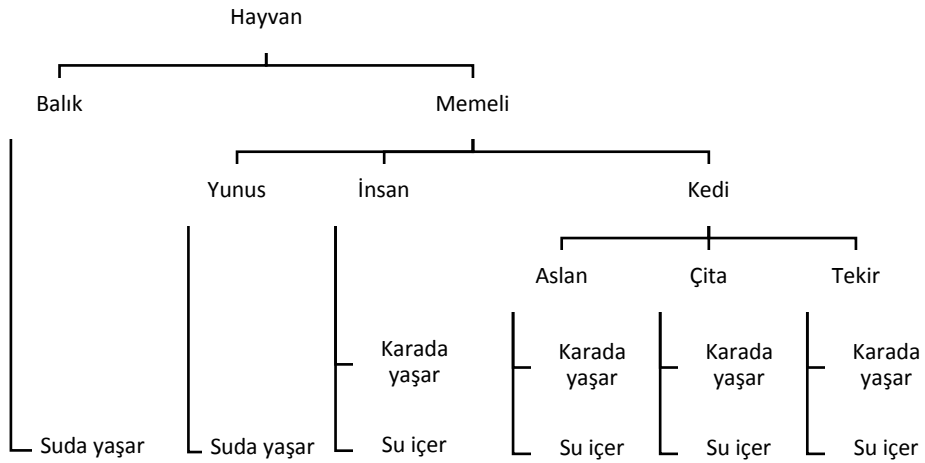
Pragmatik analiz ile cümlede gerçekte ifade edilmek anlam çıkartılır. Örneğin “O mu büyük?” cümlesi mecaz anlam içererek bahsettiği nesnenin/canlının aslında boyunun büyük olmadığından bahsediyor olabileceği gibi nesnelerden hangisinin daha büyük boyutlu veya canlılardan hangisinin yaşının daha büyük olduğunu da soruyor olabilir.

2.3. Yapay Zekâ Ve Doğal Dil İşleme Metodları

Yapay zekâ ve doğal dil işleme metotları bir bütünlük içermektedir. Yapay zekada bilgi temsili ile kelimeler birbirleri ile ilişkilendirilip gruplandırılır. Bunun için de iki temel yöntem kullanılmaktadır. Bunlar nesneler arasındaki ilişkilere odaklanan “*Anlamsal Ağ*” ve tüm alternatiflere sahip “*Ayrıştırma Ağacı*”dır.

2.3.1. Anlamsal Ağ

Anlamsal analiz, bir cümleyi bir tür anlam temsiline, örneğin mantıksal bir ifadeye eşlemeyi içerir (Brill & Mooney, 1997). Anlamsal bir ağaç, her biri belirli bir cümle sözcük birimini temsil eden temel anlamsal düğümlerden oluşturulur (Novichkova, Egorov, & Daraselia, 2003). *Şekil 7*’de örnek bir anlamsal ağ verilmiştir.

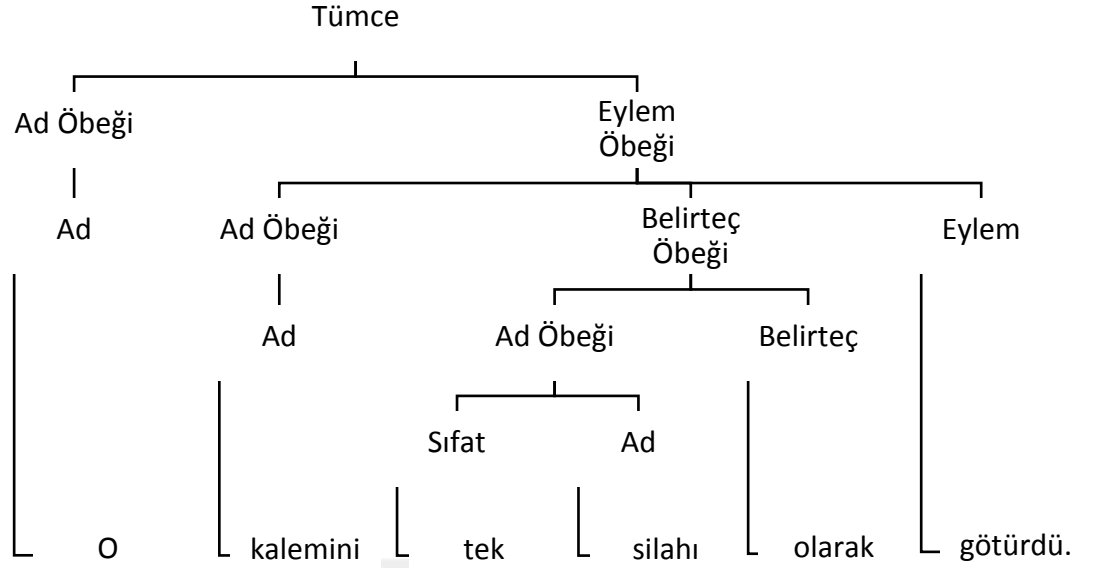


Şekil 7. Anlamsal Ağ

Anlamsal ağlar nesneleri gruplar ve nesneler arasında ilişkiler belirler. Bu ağdan makineler insan karada yaşayan, su içen memeli bir hayvandır. Balık ise suda yaşayan hayvandır gibi çıkarımlar yapabilirler.

2.3.2. Ayırıştırma Ağacı

Cümleleri, oluşturan parçaların yapısını ortaya çıkarmaya yardımcı olması için ağaç veya grafik olarak temsil etmek uygundur. Örneğin, ‘O kalemini tek silahı olarak götürdü’ cümlesi, **Şekil 8**’de gösterilen bir ayırıştırma ağacı olarak temsil edilebilir (Chowdhary, 2020).

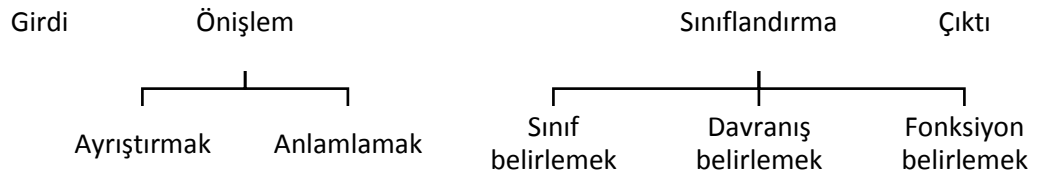


Şekil 8. Ayırıştırma Ağacı

Ayırıştırma ağaçları sayesinde cümleler içerisinde bulunan kelimelerin dilbilgisi yapısının doğru sıralamasını göstermek için kullanılır.

2.4.Doğal Dil İşleme Süreçlerindeki Yöntem

Doğal dil işlemedeki amaç makinelerin insanların ana dillerini anlamasını ve onlara yanıtlar vermesini sağlamaktır. Bunun için bilgisayarın anadilin dil yapısını ve o dildeki kelimeleri bilmesi gerekmektedir. Kullanıcıdan bir girdi geldikten sonra belirlenen algoritmalar sayesinde girdi anlamlı hale getirilir ve bir çıktı oluşturulur.



Şekil 9. Doğal Dil İşlemenin Temel Yöntemleri

Kullanıcıdan alınan girdi önişlemlerden geçirildikten sonra sınıflandırılır ve sınıflandırma sonucunda bir çıktı üretilir. **Şekil 9**'da doğal dil işlemenin temel yöntemleri verilmiştir. Doğal dil işlemenin temel yöntemleri şu şekildedir (Kothari, 2012).

Girdi: Kullanıcının makineye verdiği ses, yazı gibi metaryellerdir.

Önişlem: Önişlemde girdi cümlelere daha sonra da kelimelere ayrıştırılır. Gereksiz kelimeler ayrıştırılır. Örneğin konuşma esnasında duraklama sesleri.

Sınıflandırma: Önişlemden geçen kelimelerin anlamsal ağırları çıkarılır, sınıfları, davranışları ve fonksiyonları belirlenir. Eğer girdiye uygun sınıf, davranış veya fonksiyon bulunamazsa benzer anlamlı kelimeler ile birlikte değerlendirilir.

Çıktı: Girdi sınıflandırıldıktan sonra girdinin isteğine göre bir yanıt oluşturulur ve kullanıcıya sunulur.

3. SES İŞLEME

Son yıllarda makine algısında istikrarlı bir ilerleme akışı sağlandı. Makineler, zaman zaman insan yeteneklerini bile aşan, giderek daha karmaşık görevleri üstlenmektedir. Ses işleme araştırmaları, çoğunlukla konuşma ve müzik işlemeye odaklanmaktadır (Piczak, 2015).

3.1. Ses Tanımlama Çalışmalarının Tarihçesi

Konuşma canlıların birbirleri ile iletişimin en temel ögesidir. Makineler ile insanlar arasında etkileşim kurulma isteği dolayısı ile canlıların konuşma özelliğinin makinelere kazandırılması için yüzyıllardır bilim insanları çalışmalar yürütmekte. Bu çalışmalar ile insan seslerinin tanınması ve makinelerin insanları seslerinden ayırt etmesi ve/veya insanların verdiği komutlara cevap vermesini sağlamak istenmektedir.

İlk ses çalışmaları 1881 yılında Alexander Graham Bell tarafından keşfedilen ses kaydeden cihaz ile başlamıştır (Searle, 1980). 1930'lara gelindiğinde Bell Laboratuvarları konuşma analizi ve sentezi için bir sistem modeli önermiştir. 1952 yılında yine Bell Laboratuvarı tarafından Audrey isminde rakam tanıyabilen sistem geliştirilmiştir. 1960 yılında Gunnar Fant, konuşma üretimi için kaynak-filtre modelini geliştirmiştir. 1970'lerde artık çalışmalara DARPA da dahil olmuştur. 1000 kelimelik konuşma tanıma çalışması yapmıştır. 1990'lardan itibaren de çağrı merkezlerinin yaygınlaşması ile ses tanımlama çalışmaları daha da önem kazanmıştır. Çağrı merkezlerinde her bir çağrıya yanıt verecek yeterli eleman barındırmak firmalar için çok yüksek maliyetler getirmekte iken ses tanımlama yazılımları ücretleri ve yıllık bakım maliyetleri daha düşüktür. Artık çağrı merkezlerini aradığımızda veya internet bankacılığı kullandığımızda sesli yanıt sistemleri sayesinde para transferinden, fatura ödemeye kadar işlerimizin çoğunu çağrı merkezi elemanlarına ihtiyaç duymadan halledebilmekteyiz.

Günümüzde sesli yanıt sistemleri her ne kadar gelişmiş de olsa hala çağrı merkezlerinde çalışan kişilere ihtiyacımız var. Araştırmalar çok fazla ilerlemiş ticari yansımalarını görüyor olmamıza rağmen makineler diğer insanlar ile bir insan gibi iletişim kuramamaktadır.

3.2. Ses İşleme Modelleri

Ses işleme ve konuşma işlemede, akustik modelleme ve doğal dil işleme yöntemleri ana aktörleri oluşturmaktadır. En yaygın kullanılan modeller ise Gizli Markov Modeli ve Dinamik Zaman Bükme Yaklaşımıdır.

3.2.1. Gizli Markov Modeli

Gizli Markov Modeli, el yazısı, konuşma tanımlama gibi uygulamaları ile moleküler biyoloji hesaplamalarının, skotastik süreç ve dizilerinin modellenmesinde tercih edilen model haline gelmiştir. Gizli Markov Modeli, aynı zamanda doğal dil işleme süreçlerinde de kullanılmaktadır (Fine, Singer, & Tishby, 1998).

Markov süreci 1906 yılında Rus matematikçi Andrei Markov tarafından geliştirilen, Markov özelliği olarak da bilinen zamana göre indeksli ve mevcut durumunun koşullu olasılık dağılımının önceki duruma bağlı olduğu bir stokastik süreç türüdür.

Denklem 8'de Markov süreçleri (X) ve Markov süreçlerinin o andaki durumları (j, i, x) olmak üzere Markov özelliğinin matematiksel tanımı yapılmıştır.

$X(t), t \geq 0$ olmak üzere $X(t)$ Markov süreci olsun. Bu durumda Markov özelliğini matematiksel tanımı aşağıda verildiği gibidir:

$s, t \geq 0$ ve $0 \leq u < s$ olmak üzere;

$$P(X_{t+s} = j | X_s = i | X_u = x_u, 0 \leq u < s) = P(X_{t+s} = j | X_s = i)$$

Denklem 8. Markov özelliğini matematiksel tanımı

Markov Süreçleri, parametre uzayı ve durum uzayının ayrık veya sürekli olması durumlarına göre değişiklik gösterir. Eğer ikisi de ayrık ise ayrık-zaman Markov zincirleri olarak tanımlanırlar.

Denklem 9'da Markov süreçleri (X) ve Markov süreçlerinin o andaki durumları (j, i) olmak üzere Ayırık-zaman Markov özelliğinin matematiksel tanımı yapılmıştır.

$X(t), t \geq 0$ olmak üzere $X(t)$ Markov süreci olsun. Bu durumda ayırık-zaman Markov sürecinin matematiksel tanımı aşağıda verildiği gibidir:

i ve j, t süreçlerin durumları ve $t \geq 0$ olmak üzere;

$$P(X_t = j | X_0 = i_0, X_1 = i_1, \dots, X_{t-1} = i) = P(X_t = j | X_s = i)$$

Denklem 9. Ayırık-Zaman Markov sürecinin matematiksel tanımı

Markov süreçleri (X) ve Markov süreçlerinin o andaki durumları (j, i) olmak üzere Markov sürecinin matematiksel denklemi **Denklem 10**'da görülmektedir (Danışman, 2021).

p_{ij}, i durumundan j durumuna bir-adım geçiş olasılığı ve $\Omega > 0$ durum uzayı için;

$t \geq 0$ ve $j, t \in \Omega$ olmak üzere;

$$p_{ij} = P(X_{t+1} = j | X_t = i)$$

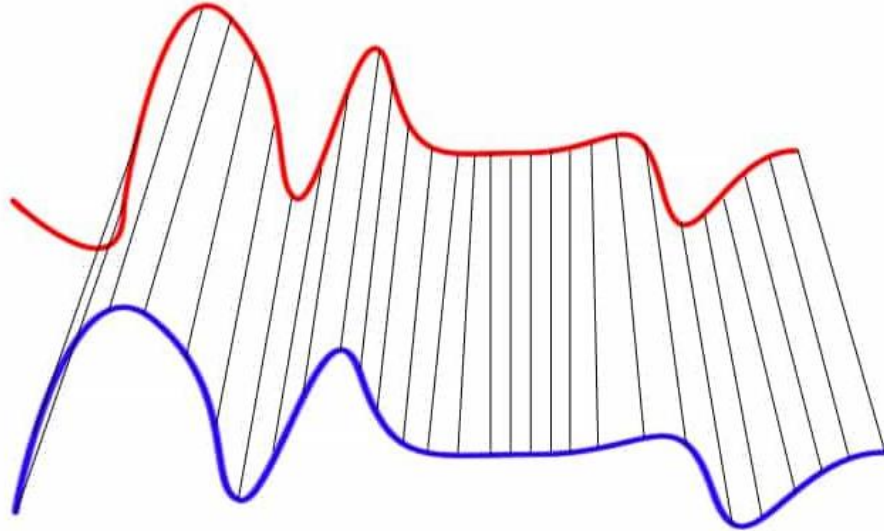
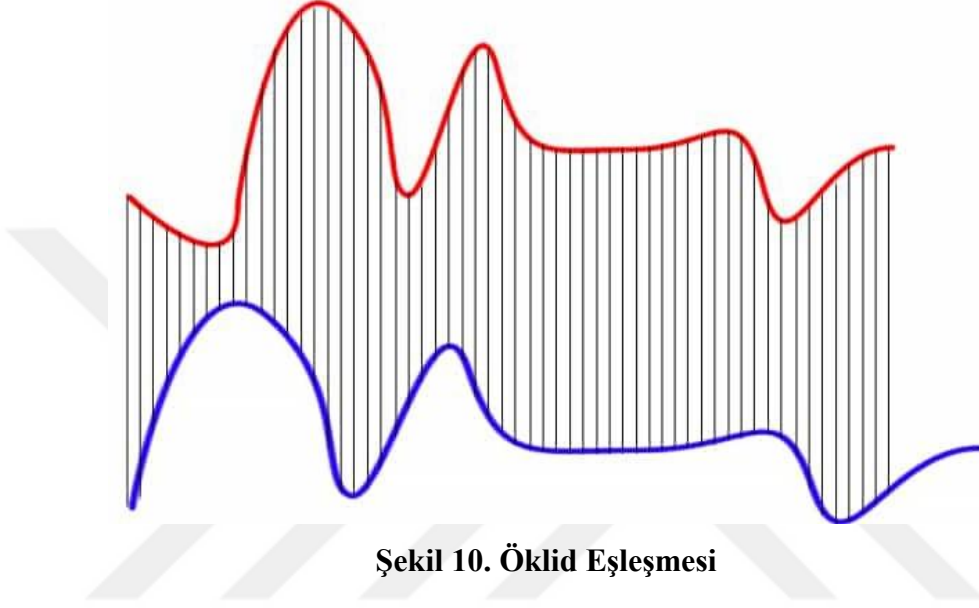
Denklem 10. Markov Süreci

3.2.2. Dinamik Zaman Bükme Yaklaşımı

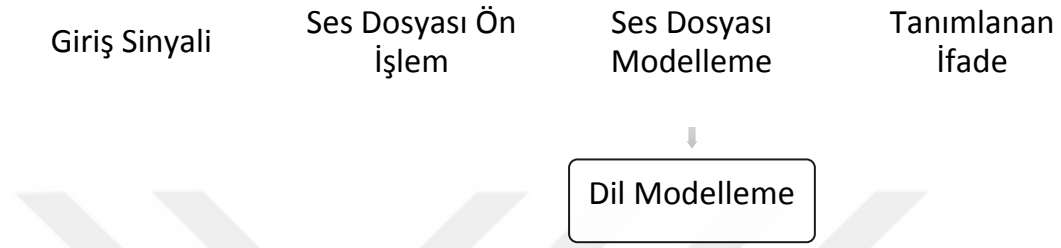
DTW algoritması, farklı fazlara sahip benzer şekilleri tespit etmek için zaman serilerinin “elastik” dönüşümüne izin vererek zaman içindeki kayma ve bozulma etkilerini en aza indiren zaman serisi benzerlik ölçüsü olarak son derece verimli olmasıyla popülerliğini kazanmıştır (Senin, 2008). İki ses dosyası arasındaki en uygun çözümleme yöntemini bulmak için kullanılan bir algoritmadır. İki ses dosyası arasındaki çözümleme yolu, iki deseninin birbirinden uzaklığını vermektedir. Desenler birbirlerine ne kadar yakın olursa desenler birbirine o kadar çok benziyordur.

Şekil 10'da iki ses dosyasının Öklid Eşlemesi kullanılarak karşılaştırılmak istendiği durumda karşılaşılan grafik gösterilmiştir.

Şekil 11'de iki ses dosyasının Dinamik Zaman Bükme Yaklaşımı kullanılarak karşılaştırılmak istendiği durumda karşılaşılan grafik gösterilmiştir.



İki ses dosyasına ait desenler Öklid Eşleşmesi ile karşılaştırıldığında makineler bu iki ses dosyasının birbirlerinden farklı olduğunu algılayabilirler ancak aynı ses dosyalarını Dinamik Zaman Bükme Yaklaşımı Algoritması ile karşılaştırdığımızda ses dosyalarının desenleri arasındaki mesafenin aslında çok fazla olmadığı ve bu iki ses dosyasının benzer sesler olduğu makine tarafından anlaşılabilir.



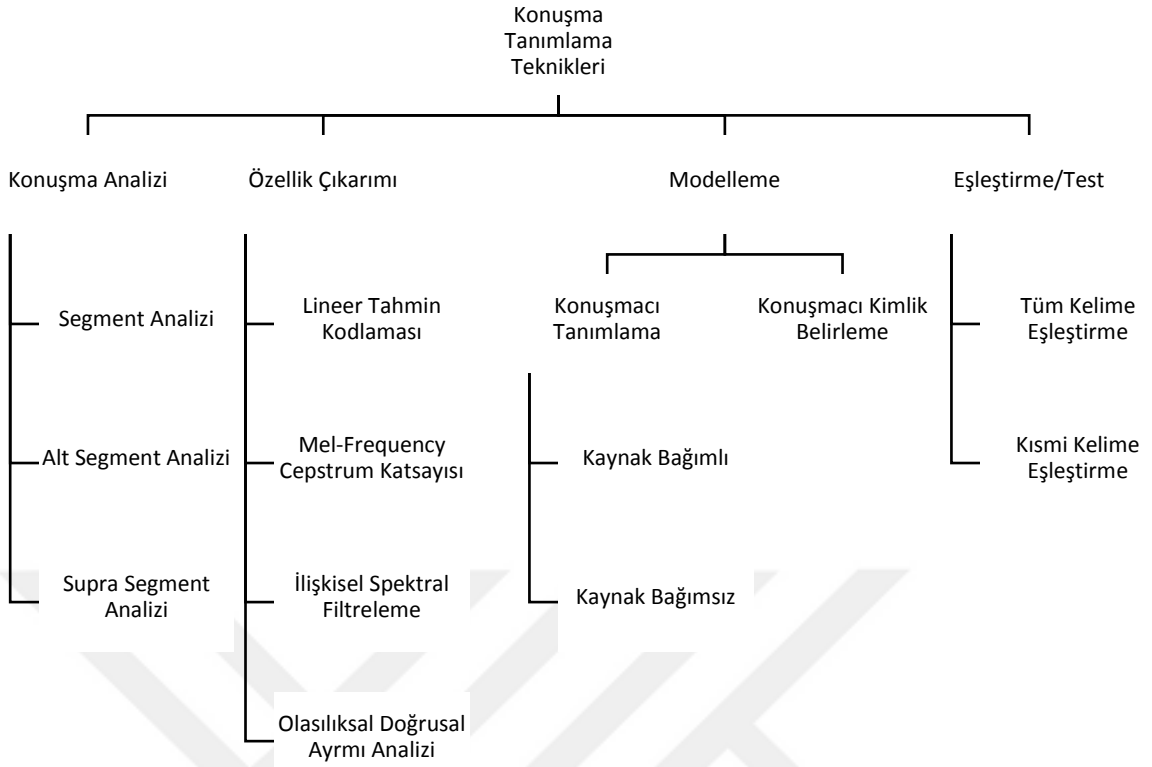
Şekil 12. Otomatik Ses Tanımlama Sistemleri Algoritması Örneği

Şekil 12'de otomatik ses tanımlama sistemlerinde yaygın olarak kullanılan bir algoritma örneği verilmiştir.

3.3. Ses Dosyalarını Tanımlama Ve Sınıflandırma

Konuşma/ses tanımlama teknikleri dört aşamadan oluşmaktadır (Almahdi, 2020). Bu aşamalar, Konuşma Analizi, Özellik Çıkarımı, Modelleme ve Eşleştirme/Test olarak sınıflandırılmaktadır.

Şekil 13'te konuşma tanımlama teknikleri verilmiştir. Konuşma tanımlama tekniklerinde sırası ile konuşma analizi, özellik çıkarımı, modelleme ve eşleştirme/test yapılır.



Şekil 13. Konuşma Tanımlama Teknikleri

Konuşma Analizi: Konuşma analizinin ilk aşamasıdır. Segment Analizi, Alt Segment Analizi, Supra Segment Analizi olmak üzere üç farklı teknik kullanılır.

1. Segment Analizi ses dosyasının 10 ila 30 ms aralıklara bölünerek analiz edilmesidir.
2. Alt Segment Analizi ses dosyasının 3 ila 5 ms boyutlarına indirgenerek analiz edilmesidir.
3. Supra Segment Analizi konuşmacının davranış ve karakteristik analizi için de tüm ses dosyasının analiz edilmesidir.

Özellik Çıkarımı: Sistemin kalbi olarak nitelendirilen özellik çıkarımında ana hedef konuşmadaki alakalı özellikleri bulmak ve alakasız olan tüm özellikleri çıkarmaktır. Ses sinyalinin gücüne zarar vermeden giriş sinyalinin vektör boyutu sıkıştırılır. Özellik çıkarımı için kullanılan tekniklerden bazıları şu şekildedir (Almahdi, 2020):

1. Lineer Tahmin Kodlaması: Hedef sinyalin önceki örnekleri toplanarak n. örneği tahminlenir. Otomatik regresyon tabanlı konuşma özellikleri sağlar.
2. Mel-Frequency Cepstrum Katsayısı: Amaç insan kulağı davranışını kopyalamaktır. Konuşma tanımlama görevleri için kullanılır. İnsan işitme sistemini taklit eder.
3. İlişkisel Spektral Filtreleme: Filtreleme sayesinde ses gürültüden arındırılır. Buradaki gürültü ses yüksekliği olmayıp cızırtı ve çevreden gelen istenmeyen seslerdir.
4. Olasılıksal Doğrusal Ayrım Analizi: Daha önceleri yüz tanımlama sistemleri için kullanılan analiz tekniği ses tanımlama teknolojilerinde de kullanılmaktadır.

Modelleme: Amaç çıkartılan özellikler ile konuşmacı modelinin oluşturulmasıdır. Konuşmacının tanımlanması için kullanılır. Akustik-Fonetik Yaklaşım, Desen Tanımlama Yaklaşımı, Dinamik Zaman Çözümleri Yaklaşımı ve Yapay Zeka Yaklaşımı modellemede kullanılmaktadır.

Eşleştirme/Test: Ses tanımlama motoru kısmi kelime tanımlama ve tüm kelime tanımlama tekniklerini kullanarak girdilerin daha önce tanımlanmış kelimeler ile eşleştirir.

1. Tüm kelime tanımlama tekniği kullanıldığı zaman daha önceden kaydedilen ses dosyası ile girdinin karşılaştırılmasıdır. Büyük veri tabanları ve sunucu boyutların gerekir.
2. Kısmi kelime tanımlama tekniği ise daha çok işlem ve daha küçük sunucu boyutu ve veri tabanı gerektirir.

İKİNCİ BÖLÜM

UYGULAMA

1. UYGULAMANIN KENDİ DİLBİLGİSİ YAPISINI SEÇMESİ

Tüm doğal diller ilk oluştuğu zamandan itibaren sürekli değişim halindedir. Ancak tüm dillerin bir dilbilgisi yapısı vardır. Dilbilgisi yapısı, bir dilin ses, biçim ve cümle yapısının kuralları bütünüdür. Makinelerin kendi dillerini oluşturmaları için kendi dilbilgisi yapılarına sahip olmaları gerekmektedir. Doğal diller biçim ve köken olmak üzere iki ayrı grupta değerlendirilebilir (Durna, 2019).

Biçim bakımından doğal diller üçe ayrılmaktadır:

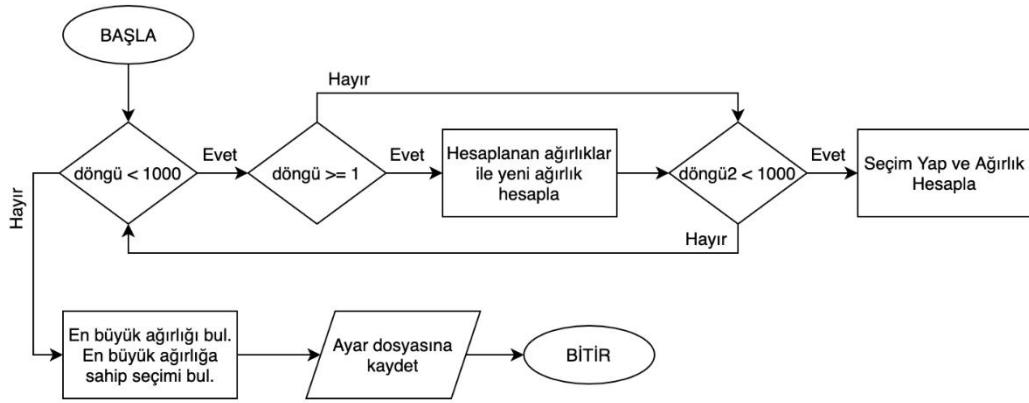
Tek Heceli Diller: Bu tür dillerde hiçbir sözcük herhangi bir ek almazken tamamı tek hecelidir. Örnek olarak Çince, Vietnam, Bask, Himalaya ve Afrika dilleri verilebilir.

Eklemlili Diller: Bu tür diller kök sözcükler içermektedir. Yeni kelimeler türetilmesi veya çekimleneceği zaman kök sözcük genellikle bozulmadan sözcüğe eklemeler yapılır. Bu dillere örnek olarak Türkçe, Macarca ve Moğolca verilebilir.

Bükümlü Diller: Kök bükümlü ve gövde bükümlü olarak ikiye ayrılan bükümlü dillerde çekimlemeler ve yeni kelime üretimlerinde kök sözcük değişikliğe uğrar. Arapça kök bükümlü dillerden iken İngilizce, Almanca ve Fransızca ise gövde bükümlü dillerdir.

Makinenin kendi diline sahip olabilmesi için dilbilgisi kurallarına ve alfabeye de ihtiyaç duymaktadır.

Dilbilgisi dil geliştikçe eklenen kurallar ile zenginleşip şekillenen kurallar bütünüdür. Makinenin çıkardığı sesleri yazı diline dökebilmesi için öncelikle alfabe seçimine odaklanıldı. Bunun için ağırlıklı seçme algoritması geliştirilmiştir.



Şekil 14. Alfabe Seçim Algoritması

Şekil 14'te uygulamadan kullanılan alfabe seçim algoritması gösterilmiştir. Ağırlıklı seçme algoritması iki döngüden oluşmaktadır. Eğitim esnasında Latin, Kiril, Yunan, Ermeni, Gürcü, Hangil, Hanzi, Kana, Arap, İbrani, Kuzey Hint, Güney Hint, Etiyopya, Tana, Kanada Hece alfabeleri veya kendi alfabesi arasında seçim yapması istendi.

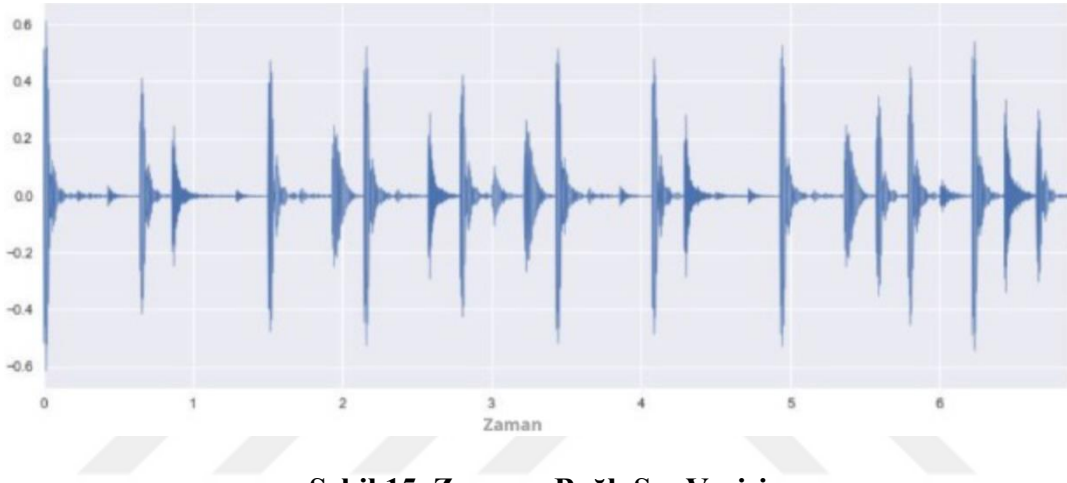
Eğitim süreci binlik bölümlere ayrıldı ve ilk binlik bölümde kendi alfabesini seçme ağırlığı iki diğer ağırlıklar ise bir olarak verildi. Her bölüm sonunda ağırlık güncellemesi yapıldı. İlk bölümün sonunda Hanzi alfabesinin ağırlığı 1 olarak kalırken diğer alfabeler için makine yeni ağırlık değerleri belirledi. Her binlik döngünün sonunda hesaplanan ağırlıklar ilk ağırlıklara eklendi ve yeni bir binlik döngü başlatıldı. Toplamda bin kez, binerlik döngü ile karar süreci tamamlandı.

Ağırlıklı seçme algoritması sonucunda Latin alfabesinin seçildiği görülmüştür. Eğitim süreci günlük dosyasına kaydedildi. Eğitim sonucunda 424 Megabayt boyutunda veri elde edildi. Günlük dosyası incelendiğinde makinenin başlarda Hanzi, Kanada Hece, Ermeni alfabesi gibi farklı alfabeleri seçip her seferinde yeni ağırlıklar hesapladığı görülmektedir. Makinenin yönelimi daha sonra Latin alfabesine doğru olmuştur.

Eğitimin sonucunda Latin Alfabesi 0.17028569202215496 gibi yüksek bir ağırlık alarak makine tarafından seçilmiştir. Seçilen alfabe dilin tüm özelliklerini içerisinde barındıracak olan ayar dosyasına kaydedilmiştir.

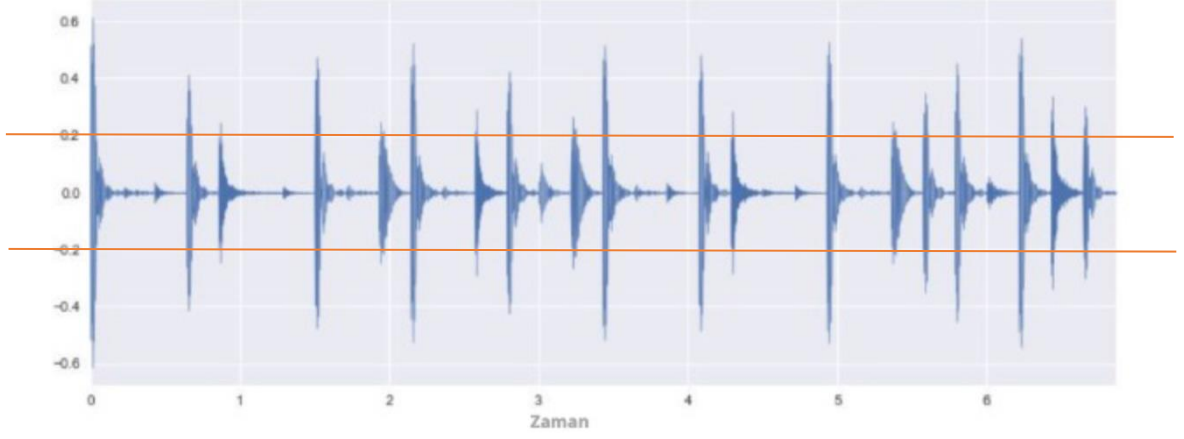
2. SES DOSYALARINDA ÖZELLİK ÇIKARIMI

Ses dosyalarında özellik çıkarımı ile sınıflandırma işlemi yapabilmek mümkün kılınmaktadır. Uzman sistemlerde ses işleme yöntemlerinde makinelere belli kurallar verilir ve istenilen çıktı alınır. Makine öğrenimli sistemlerde ise makineye istenilen çıktılar verilir ve kuralları öğrenmesi beklenir. Kuralları öğrenen sisteme yeni veriler girdi olarak verildiğinde artık bize istenilen çıktıyı kendisi üretebilir.



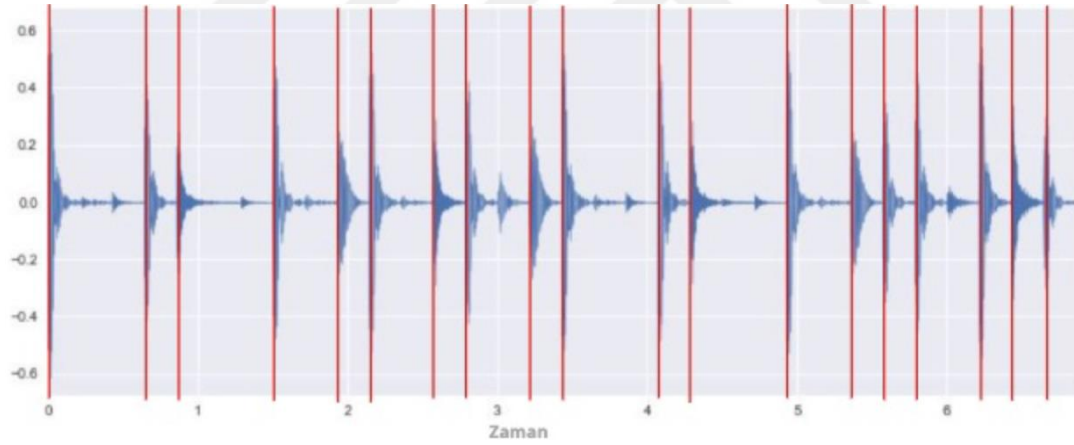
Şekil 15. Zamana Bağlı Ses Verisi

Şekil 15'te zamana bağlı ses verisi gösterilmiştir. Gösterilen veride seslerin başlangıç noktalarının hesaplanması istenildiğinde uzmanlaşmış sistemlerde “belli bir şiddetin yüksekliğinde ise ses başlamıştır” gibi kurallar verilerek seslerin başladığı noktalar belirlenebilir.



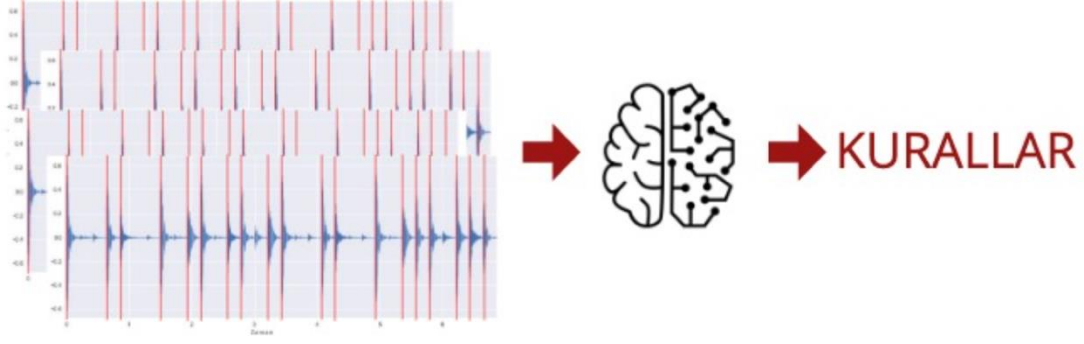
Şekil 16. Ses Dosyası Eşik Noktası Örneği

Şekil 16'da ses dosyası eşik noktası örneği gösterilmiştir. Ses dosyası için 0.2 ve -0.2 ses şiddeti eşik noktası olarak belirlenmiştir. Makineye 0.2'nin üzerinde olanlar ile -0.2'nin altında olanların seslerin başlangıç noktası olduğu söylenmiştir.



Şekil 17. Seslerin Başlangıç Noktaları

Şekil 17'de seslerin başlangıç noktaları gösterilmiştir. Verilen kural sonucunda uzman sistemlerde makinenin seslerin başlangıç noktalarını bulabildiği görülmektedir. Makine öğrenimli sistemlerde ise kuralı makinelerin oluşturması beklenmektedir.



Şekil 18. Makine Öğrenmesinde Test Verisinin Eğitimi

Şekil 18'de makine öğrenmesinde test verisinin eğitimi gösterilmiştir. Makinelere öncelikle test verileri verilir ve makinelerin bu eğitim verilerinden kurallar çıkarması beklenir.



Şekil 19. Veriden İstenilen Çıktının Elde Edilmesi

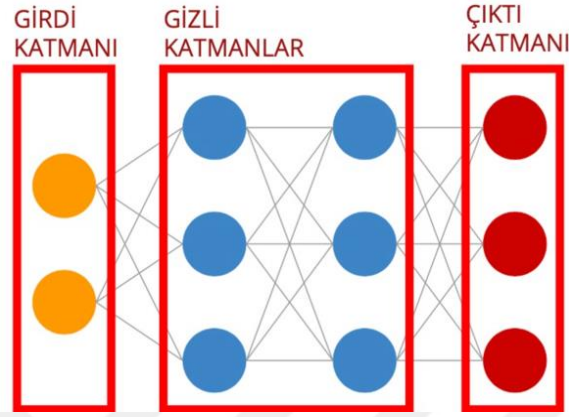
Şekil 19'da veriden istenilen çıktının elde edilme süreci gösterilmiştir. Daha önceden test verileri ile kuralları öğretilen makine öğrenimi destekli sisteme verilen girdinin istenilen çıktı olarak sistemin işlemesi görülmektedir.

Makine öğreniminde çeşitli öğrenme metotları kullanılmaktadır. Bunlardan en yaygın olarak kullanılanları denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenmedir.

Denetimli öğrenmede veriler etiketler ile işaretlenmiş şekilde makineye verilir.

Denetimsiz öğrenmede veriler etiketler ile işaretlenmeden makineye verilir. Denetimli öğrenmenin aksine kedi ve köpek seslerinin aynı klasörlerde ayırt edici bir etiket olmadan makineye eğitim verisi olarak verilmesi gerekmektedir.

Pekiştirmeli öğrenmede eğitim verileri eğitildikten sonra hatalar hesaplanıp yeni ağırlıklar ile aynı veriler ile sistem tekrar eğitilir. Sistemde belirlenen sayıda eğitim gerçekleştirilir.



Şekil 20. Yapay Sinir Ağı

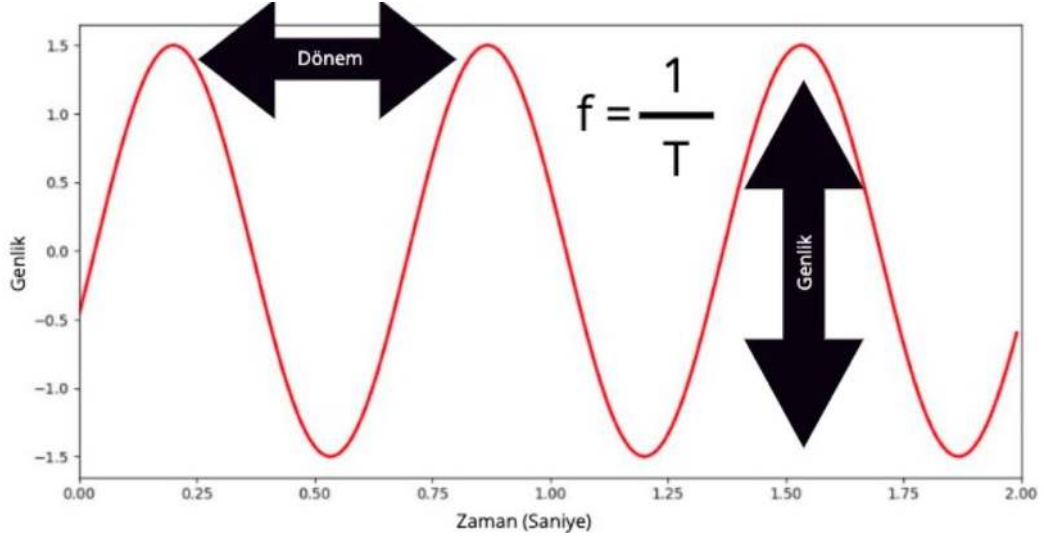
*Şekil 20'*de yapay sinir ağının katmanları gösterilmiştir.

Makine öğrenimi algoritmalarında yapay sinir ağları oluşturulur. Sinir ağlarındaki nöronların ağırlıkları değiştirilerek makinenin eğitimi gerçekleştirilir. Yapay sinir ağları girdi, gizli ve çıktı katmanları olmak üzere üç grupta incelenmektedir.

Derin Öğrenme: Makine öğrenmesinin alt katmanı olan derin öğrenmede birden fazla gizli nöron katmanı bulunmaktadır. Derin öğrenme algoritmaları ses işleme uygulamalarında kullanılmaktadır.

Derin öğrenme metodolojisini kullanarak ses dosyalarının özelliklerini çıkarmadan önce ses dosyalarını anlamamız gerekmektedir.

*Şekil 21'*de ses dosyalarının zaman(saniye) – genlik türünde dalga formu grafiği gösterilmiştir.



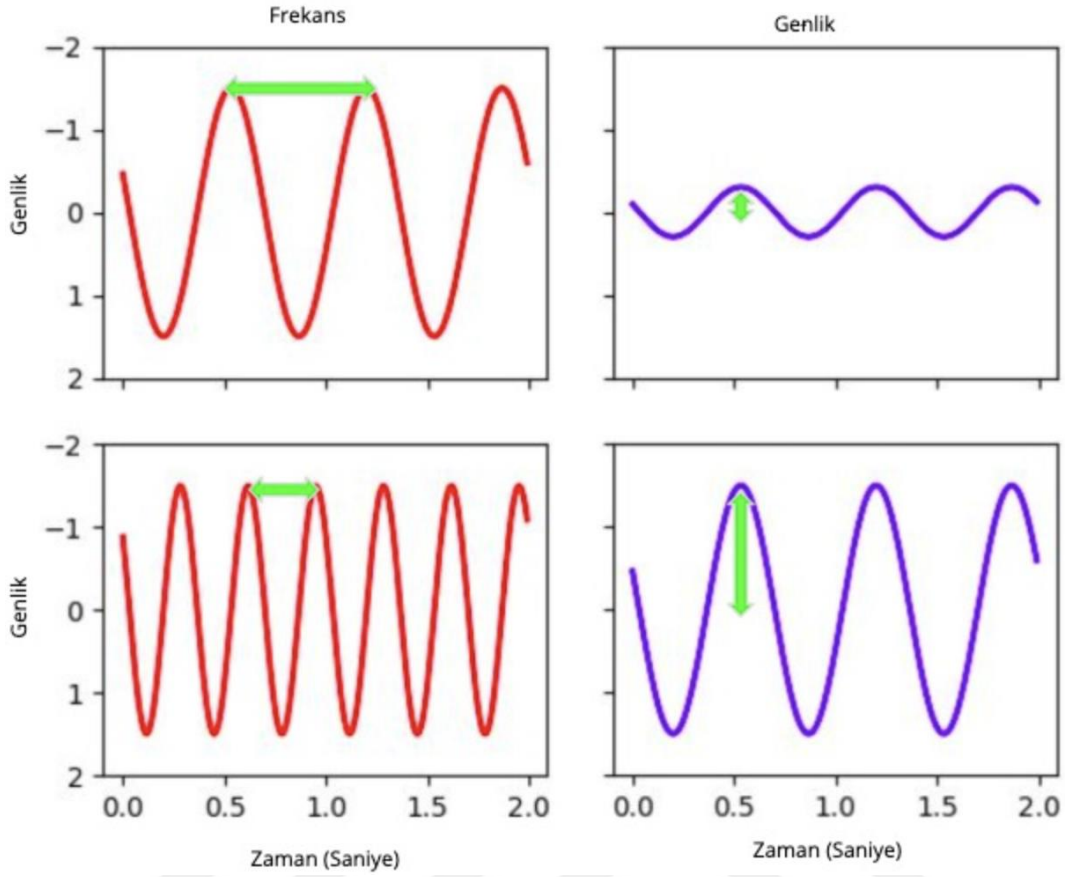
Şekil 21. Dalga Formu

Ses dosyaları dalga formundan meydana gelmektedir. Ses dosyasındaki iki tepe noktası arasındaki zamana dönem adı verilmektedir. Tepe noktalarının şiddetleri ise genlik olarak adlandırılmaktadır. t anındaki genliğin hesaplanması için aşağıdaki formül kullanılmaktadır.

Denklem 11'de faz ofseti (φ) ve frekans (f) olmak üzere t anındaki genliğin ($y(t)$) hesaplama denklemi verilmiştir.

$$y(t) = A \sin(2\pi f t + \varphi)$$

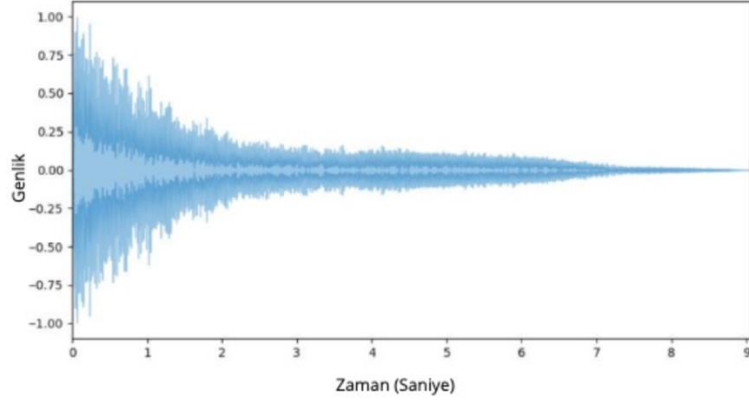
Denklem 11. t anındaki genliğin hesaplanma denklemi



Şekil 22. Frekans/Adım ve Genlik/Ses Şiddeti

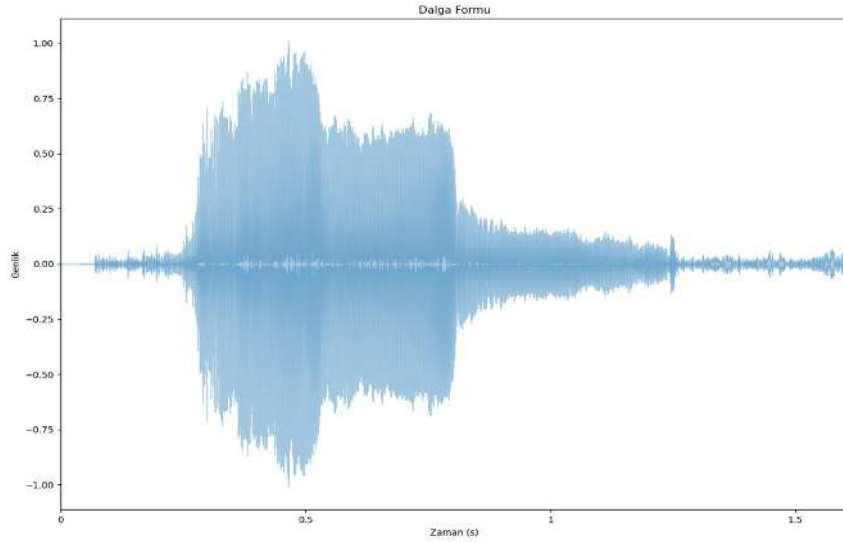
Şekil 22'de ses dosyalarının frekans – adım ve genlik – ses şiddeti türünde dalga formu grafiği gösterilmiştir.

Daha uzun frekansa sahip sesler daha uzun adımları, daha fazla genliğe sahip sesler de daha şiddetli anları ifade etmektedir. Ses şiddeti azaldıkça genlik, konuşma hızlandıkça frekans azalır.



Şekil 23. Gerçek Dünya Piyano Sesi

Şekil 23'te piyano sesinin Zaman-Genlik grafiği gösterilmiştir.



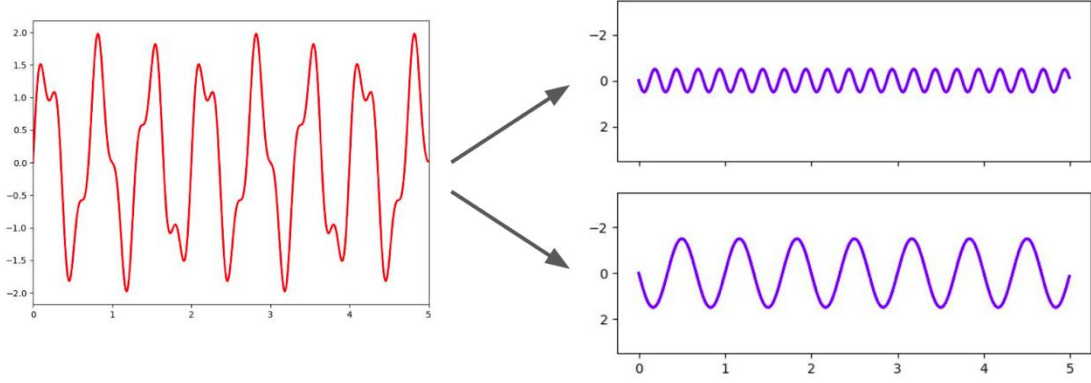
Şekil 24. Genlik-Zaman Grafiği

Şekil 24'te kedi sesinin genlik-zaman grafiği gösterilmiştir.

Ses dosyalarında özellik çıkarımı için Python programlama dilinin Librosa kütüphanesi kullanılmıştır. Veri seti olarak kullandığım Kaggle Audio Cats and Dogs veri setindeki kedi ve köpek seslerinin 22050 örnekleme ve 0.4 alfa değeri ile Librosa kütüphanesinin Waveplot metodu kullanılarak dalga formları hesaplanmıştır. Veri setinde bulunan cat_4.wav dosyası Matplotlib kütüphanesi oluşturulan ve *Şekil 24*'te gösterilen genlik-zaman grafiği incelendiğinde 500 milisaniyeye yaklaşıldığında kedi sesinin tepe noktasına gelmiş daha sonra da sesin azalmaya başladığı görülmektedir.

3. ÖN İŞLEM

Ses verisinin Zaman-Genlik grafiği bize zamana bağlı ses şiddeti hakkında bilgi verse de seslerin sınıflandırılabilmesi için yeterli değildir. Bir sonraki adım olarak ses verisinin “Fourier Dönüşümü” ile şiddet spektrumunun çıkarılması gerekmektedir.



Şekil 25. Fourier Dönüşümü

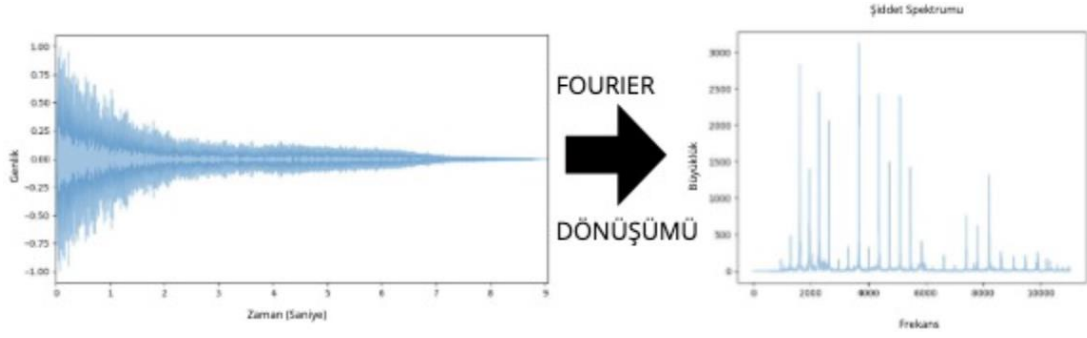
Şekil 25’te Fourier dönüşümü grafiği gösterilmiştir.

Fourier dönüşümü ile karmaşık periyodik ses dosyaları farklı frekanslardaki salınımlı sinüs dalgalarının toplamaları biçiminde ayrıştırılabilmektedir.

Denklem 12’de faz ofseti (φ), frekans (f) ve t anındaki genlik ($y(t)$) olmak üzere Fourier dönüşüm denklemi verilmiştir.

$$s = A_1 \sin(2\pi f_1 t + \varphi_1) + A_2 \sin(2\pi f_2 t + \varphi_2)$$

Denklem 12. Fourier Dönüşümü Denklemi

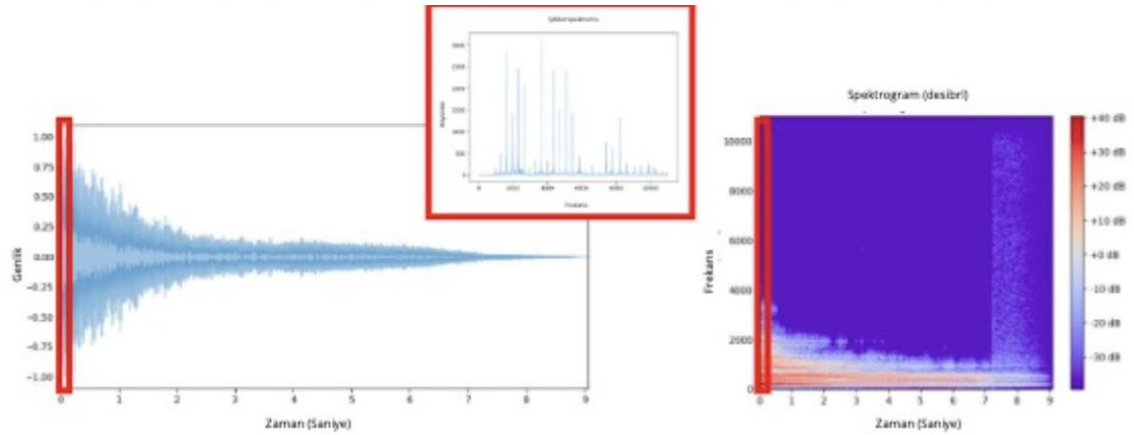


Şekil 26. Fourier Dönüşümü ile Büyüklik-Frekans Grafiği

Şekil 26'da Fourier dönüşümü ile oluşturulan büyüklik-frekans grafiği gösterilmiştir.

Fourier Dönüşümü ile veri zaman tabanından frekans tabanına evrilmiştir ve zaman bilgisi artık mevcut değildir. Zaman bilgisini de elde edebilmek için Kısa Zaman Fourier dönüşümü kullanılması gerekmektedir.

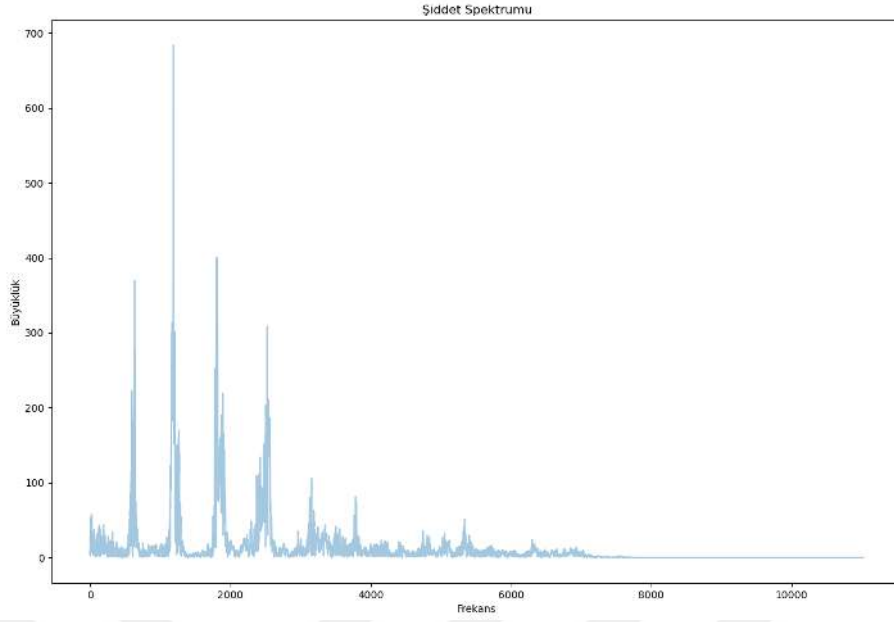
Kısa Zaman Fourier Dönüşümü farklı aralıklardaki birden fazla Fourier Dönüşümünü hesaplanabiliyor böylelikle zaman bilgisine erişebiliyoruz. Bize zaman ve frekansın bir fonksiyonu olarak büyüklik spektrogramını vermektedir.



Şekil 27. Kısa Zaman Fourier Dönüşümü

Şekil 27'de Kısa Zaman Fourier dönüşümü gösterilmiştir.

Veri setindeki kedi ve köpek seslerinin dalga formundan Fourier Dönüşümü ile büyüklik-frekans grafiği oluşturulmuştur. Fourier Dönüşümü ile karmaşık periyodik ses dosyaları farklı frekanslardaki salınımlı sinüs dalgalarının toplamları biçiminde ayrıştırılmıştır.

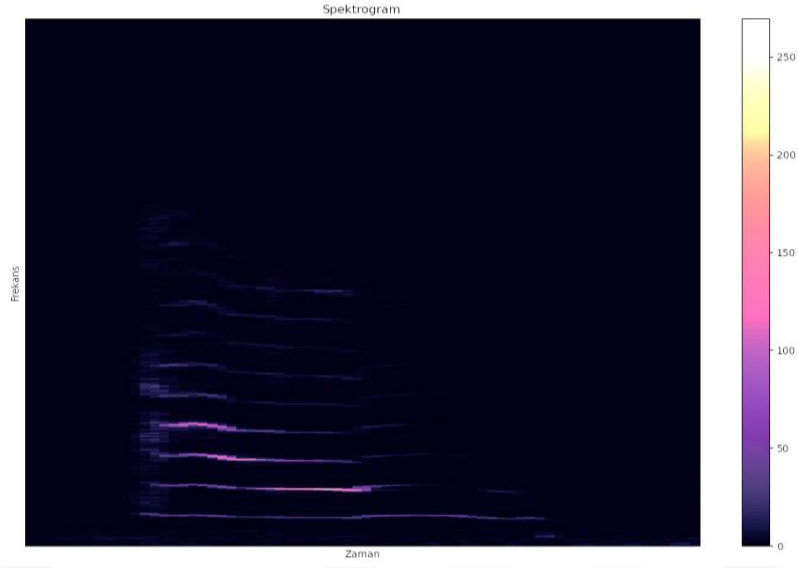


Şekil 28. Büyüklük-Frekans Grafiği

Şekil 28'de büyüklük-frekans grafiği gösterilmiştir.

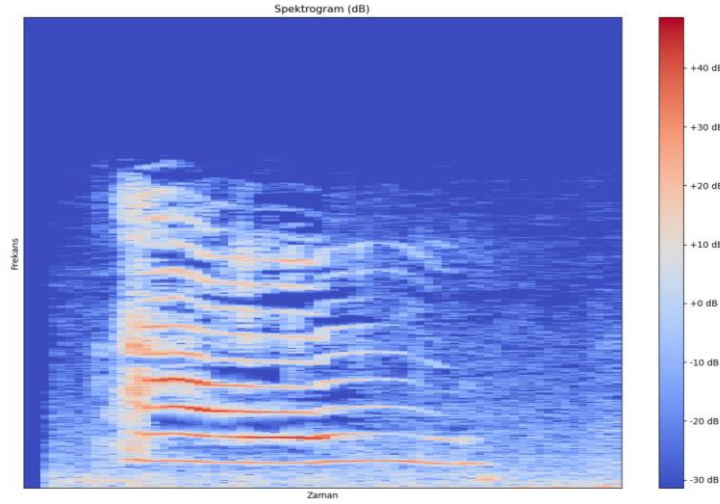
Bir önceki bölümde Genlik-Zaman grafiği çıkarımı yaptığımız veri, Numpy kütüphanesi kullanılarak dalga formundan Fourier Dönüşümü ile büyüklük-frekans grafiği oluşturulmuştur. Fourier Dönüşümü ile karmaşık periyodik ses dosyaları farklı frekanslardaki salınımlı sinüs dalgalarının toplamaları biçiminde ayrıştırılmıştır.

Veride Fourier Dönüşümü yaptıktan sonra Kısa Zaman Fourier Dönüşümü yapılmıştır. Kısa Zaman Fourier Dönüşümü ile farklı aralıklardaki birden fazla Fourier Dönüşümünü hesaplanabiliyor böylelikle zaman bilgisine erişebiliyoruz. Bize zaman ve frekansın bir fonksiyonu olarak büyüklük spektrogramını vermektedir.



Şekil 29. Frekans, Zaman, Büyüklük Spektrogramı

Şekil 29'da frekans, zaman, büyüklük spektrogramı gösterilmiştir. *Şekil 29*'da gösterilen spektrogram verisine logaritma uygulanarak genlikler desibel değerlerine çevrilerek *Şekil 30*'da bulunan frekans ve zamanın bir fonksiyonu olarak desibel spektrogramı elde edilmiştir.



Şekil 30. Frekans, Zaman, Desibel Spektrogramı

Şekil 30'da frekans, zaman, desibel spektrogramı gösterilmiştir.

4. MODELİN OLUŞTURULMASI

Özellikleri çıkarılan veri setleri model oluşturulup ilk kelimeler üretilmeden önce ön işlemden geçirilmelidir. Kısa zaman Fourier dönüşümü yapılan veri setlerinin ilk olarak Mel Frekans Cepstral Katsayısının hesaplanması yapılmıştır. Mel Frekans Cepstral Katsayısı uzun yıllardır konuşma tanımlama için belirgin özellik olmuştur (Büyüköz, 2016). Mel Frekans Cepstral Katsayıları kısa dönem spektral tabanlı özelliklerdir. Mel Frekans Cepstral Katsayıları müzik ve ses dosyalarının modellenmesinde kullanılır.



Şekil 31. Mel Frekans Cepstral Katsayısı Oluşturma Algoritması

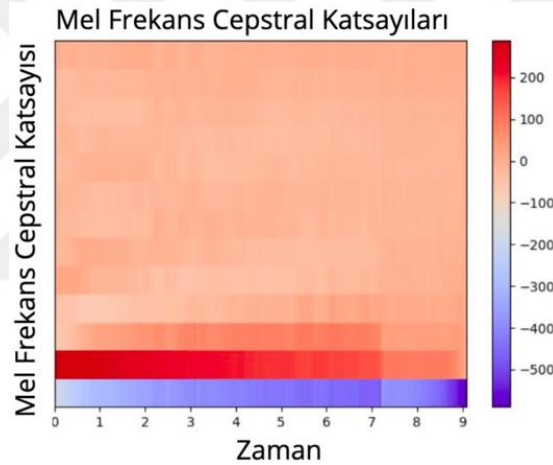
Şekil 31'de Mel Frekans Cepstral Katsayısı oluşturma algoritması gösterilmiştir.

Ses verileri için geleneksel makine öğrenmesi algoritma süreçlerinde Fourier Dönüşümlerinin ardından genlik zarfı, spektral ağırlık merkezi, 0 geçiş oranı ve spektral akış hesaplamaları yapıp makine öğrenmesi algoritmaları ile ses verileri işlenir.



Şekil 32. Ses Verileri İçin Geleneksel Makine Öğrenmesi Akışı

*Şekil 32'*de ses verileri için geleneksel makine öğrenmesi akışı gösterilmiştir.



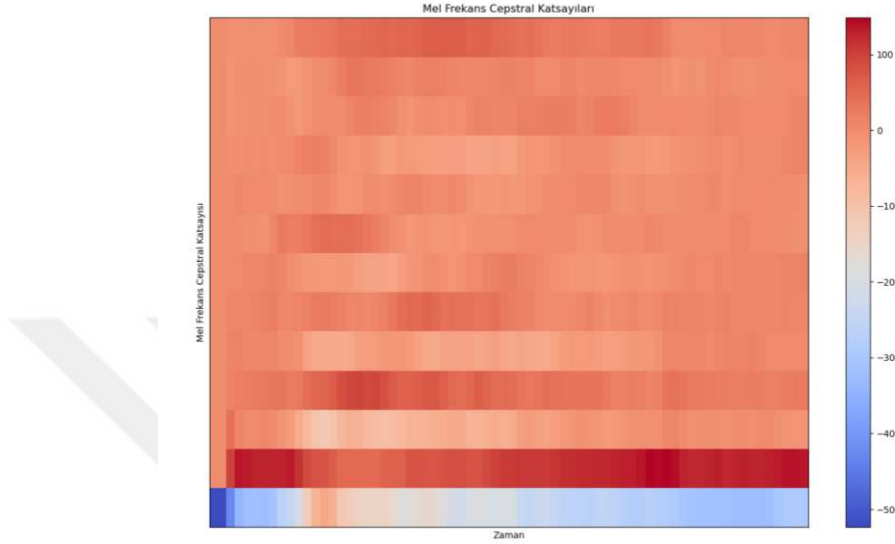
Şekil 33. Mel Frekans Cepstral Katsayıları Grafiği

*Şekil 33'*te Mel Frekans Cepstral Katsayıları grafiği gösterilmiştir.

*Şekil 32'*de verilen algoritma sonucunda veri setlerine ait zamana bağlı Mel Frekans Cepstral Katsayıları hesaplanabilir. Bunun sonucunda da *Şekil 33'*te gösterilen grafik elde edilir. Mel Cepstral Katsayıları hesaplanan veri setine yapay sinir ağları eklenerek ses dosyalarının sınıflandırılması sağlanabilmektedir.

Tek veri için modellemesi yaptığımız ve Mel Cepstral Katsayısını hesapladığımız algoritmayı yapay sinir ağları kullanarak tüm veri setimize uygulamamız gerekmektedir.

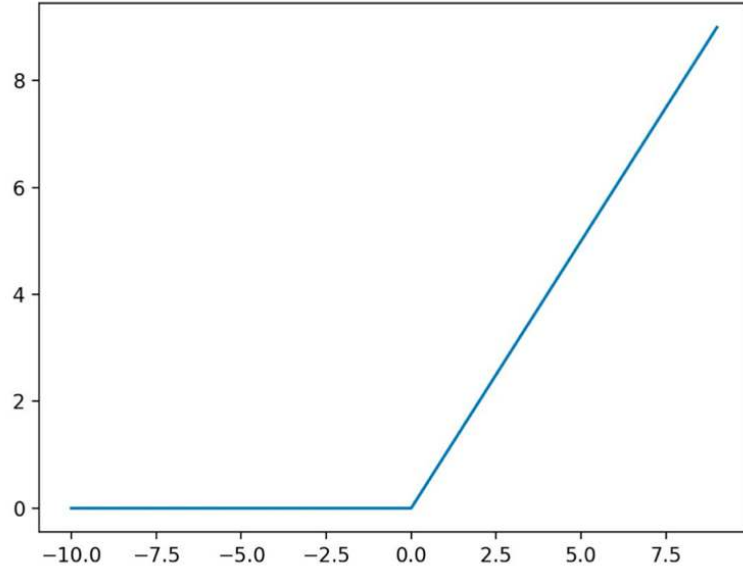
*Şekil 30'*da spektrumu gösterilen ses yapılarak sesin verisinin timbral/dokusal yönlerinin yakalanması, frekans alan özelliğinin ve yaklaşık insan işitme sisteminin elde edilmesi için her karede Mel Frekans Cepstral Katsayısı hesaplaması yapılmıştır.



Şekil 34. Mel Frekans Cepstral Katsayıları

*Şekil 34'*te Mel Frekans Cepstral Katsayıları gösterilmiştir.

Veri setine Doğrultulmuş Lineer Birim uygulanmıştır. Doğrultulmuş Lineer Birim derin öğrenme sistemlerindeki nöral birimler için temel modeldir (Özcan, 2020). Doğrultulmuş Lineer Birim kullanılarak daha iyi bir yakınsama ve degradenin kaybolma olasılığının azaltılması hedeflenmiştir.



Şekil 35. Doğrultulmuş Lineer Birim

Şekil 35'te Doğrultulmuş Lineer Birim grafiği gösterilmiştir. Grafiği verilen Doğrultulmuş Lineer Birim için **Denklem 13**'te çıktı (h) olmak üzere Doğrultulmuş Doğrusal Birim ($ReLU(h)$) denklemi gösterilmiştir.

$$ReLU(h) = \begin{cases} 0, & h < 0 \\ h, & h \geq 0 \end{cases}$$

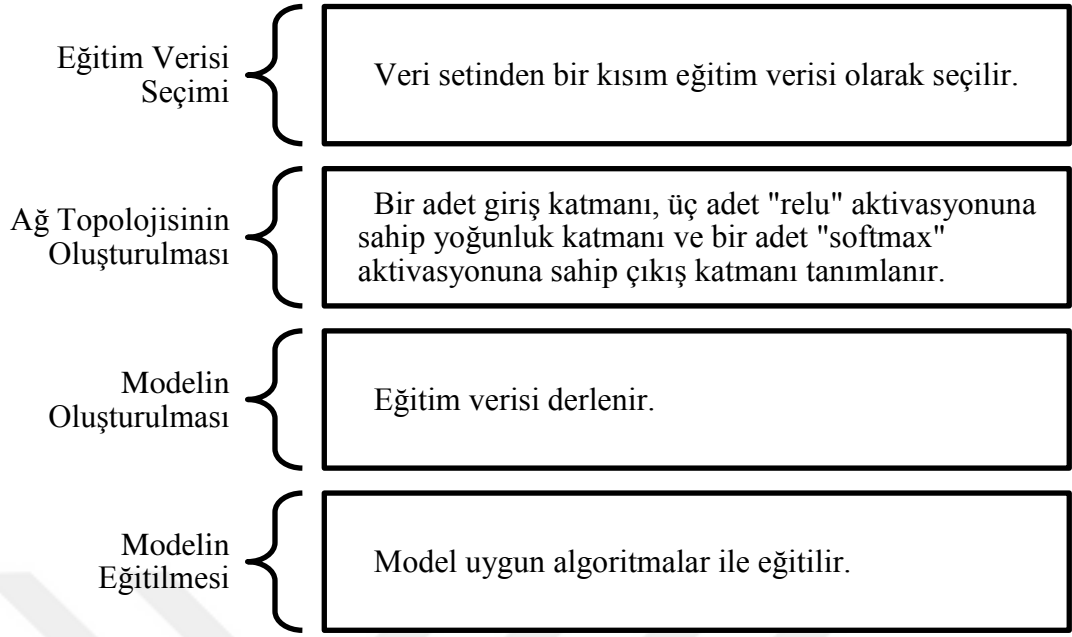
Denklem 13. Doğrultulmuş Lineer Birim Denklemi

Doğrultulmuş Lineer Birimin, stokastik, tam harmanlama, mini harmanlama olmak üzere üç farklı harmanlama biçimi vardır.

Stokastik: Bir örnek için gradyan hesaplar. Hızlıdır ancak kusurludur.

Tam Harmanlama: Tüm veri seti için gradyan hesaplar. Yavaş, bellek bağımlı ve kusursuzdur.

Mini Harmanlama: 16-128 örnekleme ile tüm veri setinin alt kümesi için gradyan hesaplar.



Şekil 36. Doğrultulmuş Lineer Birim Modelin Eğitilme Süreçleri

Şekil 36'da Doğrultulmuş Lineer Birim modelin eğitilme süreçleri gösterilmiştir.

Mel Frekans Cepstral Katsayısı hesaplanan verilere Doğrultulmuş Lineer Birim uygulanmıştır. *Tablo 1*'de Ağ topolojisi oluşturulurken kullanılan katman özellikleri ve değerleri gösterilmiştir.

Tablo 1. Ağ Topolojisi Katman Özellikleri ve Değerleri

Katman (tip)	Çıktı, Şekil	Değer #
Düzleştirilmiş (Flatten)	(Yok, 169)	0
Yoğunluk (Dense)	(Yok, 512)	87040
Yoğunluk_1 (Dense)	(Yok, 256)	131328
Yoğunluk_2 (Dense)	(Yok, 64)	16448
Yoğunluk_3 (Dense)	(Yok, 10)	650
=====		
Toplam değer: 235466		
Eğitilebilir değer: 235466		
Eğitilemez değer: 0		

Veriler, verilen değerler ile elli kez eğitilmiştir. Eğitimde ilk adım 1 saniye 26 milisaniye sürerken diğer adımlar ortalama 3 milisaniyede tamamlanmıştır. Eğitim toplamda 1 saniye 421 milisaniye Eğitimin başında 0,2380 olan doğruluk oranı eğitimin sonunda 0,9819 ve yine eğitimin başında 52,8339 olan kayıp oranı eğitimin sonunda 0,0572 olmuştur.

Veriler eğitilirken birden çok yoğunluk katmanı kullanılmasının sebebi modeli daha derin hale getirmek ve Doğrultulmuş Lineer Birimini bir Derin Öğrenme tekniği olarak daha doğru değerler ile eğitmek. Genellikle yaklaşımların çok çeşitli zorlu tahmin problemlerindeki başarıları, sinir ağlarının derinliği ile ilişkilendirilir. Ancak katmanlar derinleştikçe eğitilecek değer sayısı da azalmaktadır.

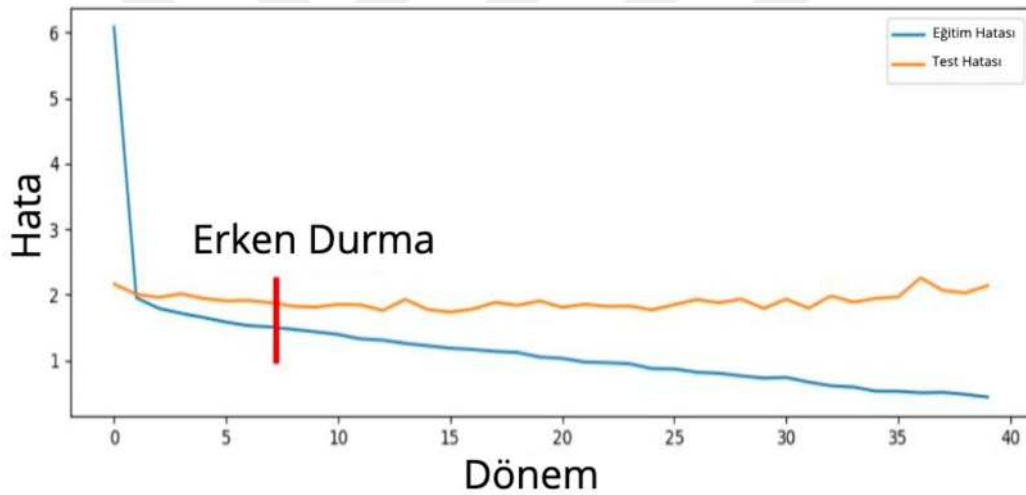
Doğrultulmuş Lineer Birim modelinin eğitilmesi için Şekil 25'te belirtilen algoritma kullanılmıştır. Doğrultulmuş Lineer Birim modeli eğitildikten sonra yapay

sinir ağındaki aşırı uyum gösterme sorununun çözülmesi gerekmektedir. Bunun için daha basit mimari, veri büyütme, erken durma, bırakma ve düzenlilik kurallarının izlenmesi gerekmektedir.

Daha basit mimari: Katmanların kaldırılması ve nöronların sayısının azaltılması gerekmektedir. Evrensel herhangi bir kuralı yoktur.

Veri büyütme: Eğitim verilerinin sayısının yapay olarak artırılması ve ses dosyalarına süre uzatma ve arka plan gürültüsü ekleme gibi dönüşümlerin uygulanması gerekmektedir.

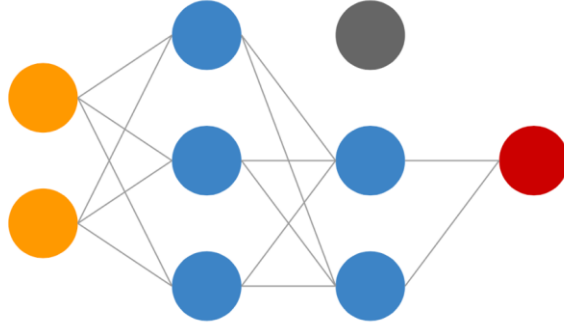
Erken durma: Eğer eğitim verisi hatası ile test hatası arasındaki fark artıyor ise farkın artmasına izin verilmeden eğitimin yarıda kesilmesi gerekmektedir.



Şekil 37. Erken Durma

*Şekil 37'*de dönem – hata bazlı erken durma grafiği gösterilmiştir.

Bırakma: Eğitim esnasında rastgele olarak nöronlar serbest bırakılır ve yapay sinir ağının sağlamlığı artırılır.



Şekil 38. Nöron Bırakma

Şekil 38'de nöron bırakma örneği gösterilmiştir.

Düzenlilik: Hata fonksiyonuna ve büyük ağırlıklara cezalandırma uygulayıp, L1 ve L2 Düzenlilik fonksiyonları uygulanmalıdır.

L1 Düzenlilik Fonksiyonu: Ağırlık değerlerinin, mutlak değerlerini küçültür. Aykırı değerlere karşı dirençlidir ve basit model oluşturur.

Denklem 14'te mutlak ağırlık ($|W|$), katsayı (λ), beklenen çıktı (p) ve çıktı (y) olmak üzere L1 Düzenlilik Denklemi verilmiştir.

$$E(p, y) = \frac{1}{2}(p - y)^2 + \lambda \sum |W_i|$$

Denklem 14. L1 Düzenlilik Denklemi

L2 Düzenlilik Fonksiyonu: Ağırlık değerlerinin karesini küçültür. Aykırı değerlere karşı dirençli değildir ve karmaşık modelleri öğrenir.

Denklem 15'te ağırlık (W), katsayı (λ), beklenen çıktı (p) ve çıktı (y) olmak üzere L2 Düzenlilik Denklemi verilmiştir.

$$E(p, y) = \frac{1}{2}(p - y)^2 + \lambda \sum W_i^2$$

Denklem 15. L2 Düzenlilik Denklemi

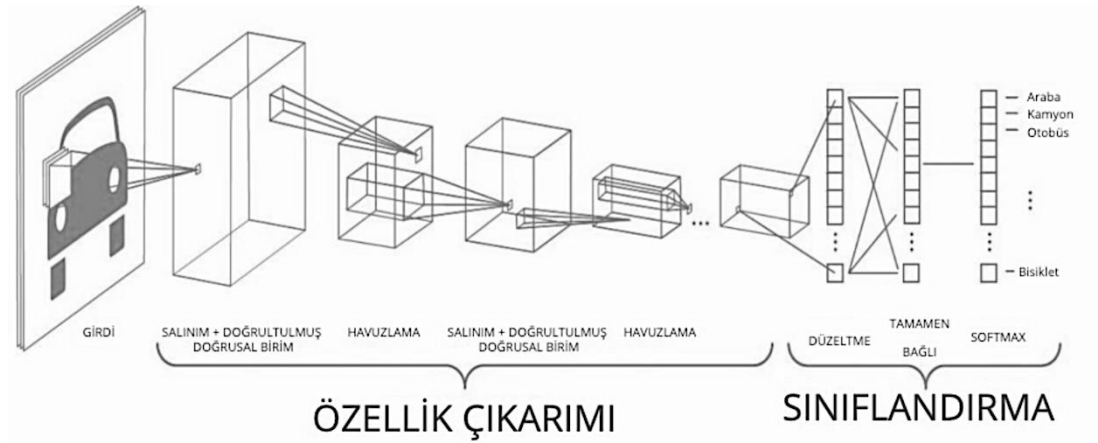
Tez yazım aşamasında ses dosyalarının işlenmesi ve sınıflandırılmasında Evrişimli Yapay Sinir Ağları kullanılması tercih edilmiştir. Evrişimli Yapay Sinir Ağları, temelde görüntü işlemede kullanılmaktadır. Çok katmanlı algılayıcılara nazaran daha iyi performans göstermekte olup yoğunluk katmanlarına göre de daha az parametre almaktadır.

Denklem 16'da i anındaki görüntü ($görüntü_i$) ve belirlenmiş matris (K) olmak üzere Görüntünün Matrise Dönüşüm Denklemi verilmiştir

$$\sum_{i=1}^P görüntü_i \cdot K_i$$

Denklem 16. Görüntünün Matrise Dönüşüm Denklemi

Matrise çevrilmiş görüntü, belirlenmiş başka bir matris ile çarpılır ve görüntü üzerinde derinlik, kirlilik arındırması, kenar düzeltme ve özellik çıkarımı yapılabilmektedir.



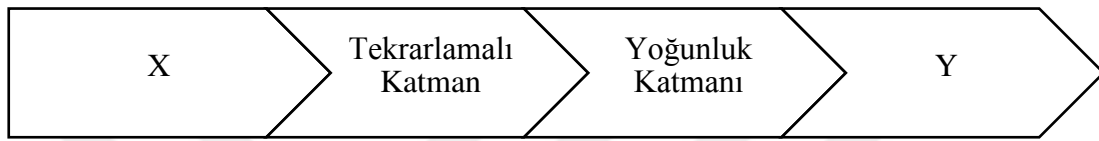
Şekil 39. Evrişimli Yapay Sinir Ağları Mimarisi

Şekil 39'da Evrişimli Yapay Sinir Ağları Mimarisi gösterilmiştir.

Evrişimli Yapay Sinir Ağları, temelde görüntü işlemede kullanılırken bu çalışmada ses işlemede kullanılacaktır. Ses dosyalarını işlerken Mel Frekans Cepstral

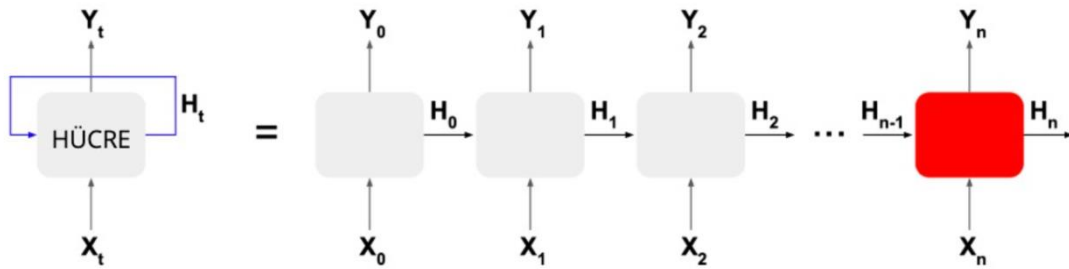
Katsayısı/Spektrum görüntü, zaman ve frekans görüntü matrisinde bulunan x ve y ve genlik ise piksel değeri olarak kabul edilecektir.

Evrişimli Yapay Sinir Ağları uygulanan modelimize Tekrarlamalı Yapay Sinir Ağları uygulanmıştır. Tekrarlamalı Yapay Sinir Ağlarında, sıra ve değişken uzunluğu önemlidir. Ardışık veriler için kullanılır. Her öge bağlam içinde işlenir. Ses ve müzik uygulamaları için ideal bir çözümdür. Veri noktaları tek tek girilip sonraki adımlar tahminlenir. Tahminlemeler önceki veri noktasına bağlıdır.



Şekil 40. Tekrarlamalı Yapay Sinir Ağları Mimarisi

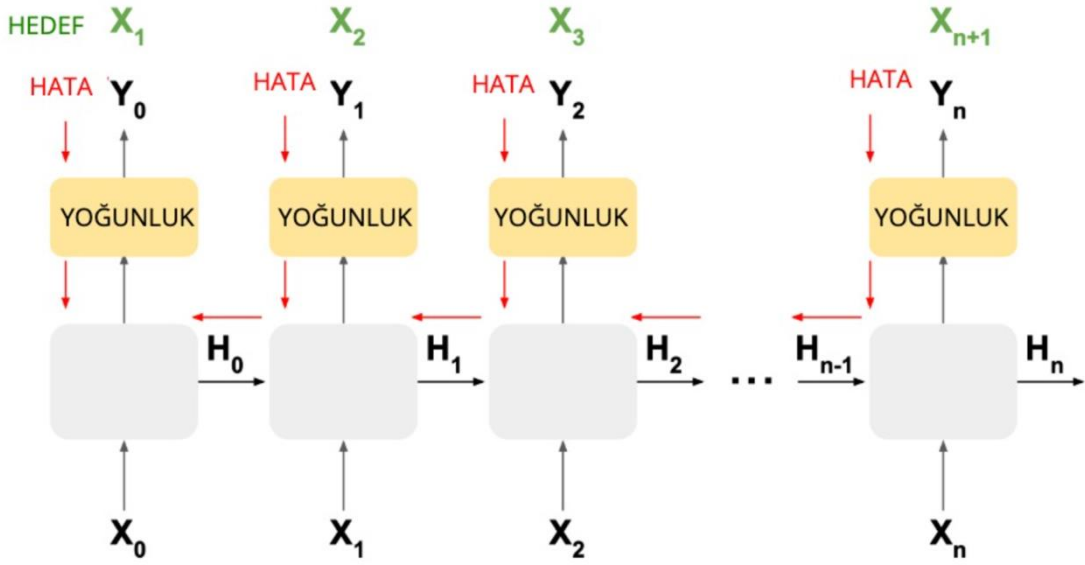
Şekil 40'da Tekrarlamalı Yapay Sinir Ağları Mimarisi gösterilmiştir.



Şekil 41. Tekrarlamalı Yapay Sinir Ağı Katmanı

Şekil 41'de Tekrarlamalı Yapay Sinir Ağı Katmanı gösterilmiştir.

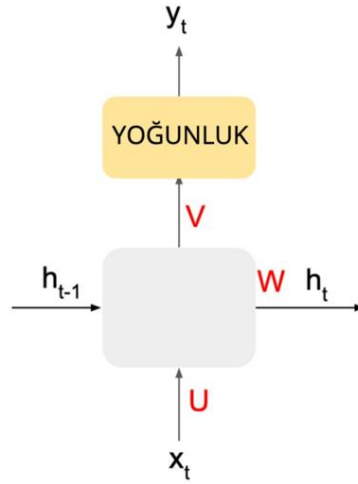
Temel bir Tekrarlamalı Yapay Sinir Ağında yoğunluk katmanı, girdi olarak alan vektörü ve girdi verisi ile Tekrarlamalı Yapay Sinir Ağlarının zorlukları ve Doğrultulmuş Lineer Birimin hata almaması için tanh kullanılmaktadır.



Şekil 42. Zamanla Geri Yayılım

Şekil 42'de Zamanla Geri Yayılım gösterilmiştir.

Tekrarlamalı Yapay Sinir Ağlarında hatalar yoğunluk hesabı yapıldıktan sonra hedefe ulaşana kadar tekrar hesaplatılır.



Şekil 43. Tekrarlama Katmanı

Şekil 43'te tekrarlama katmanını gösterilmiştir. Katman hesaplanırken aşağıdaki formülden yararlanılır.

Denklem 17'de ağırlık (W), çıktı (y) ve girdi (x) olmak üzere Tekrarlama Katmanı (h_t) Denklemi verilmiştir.

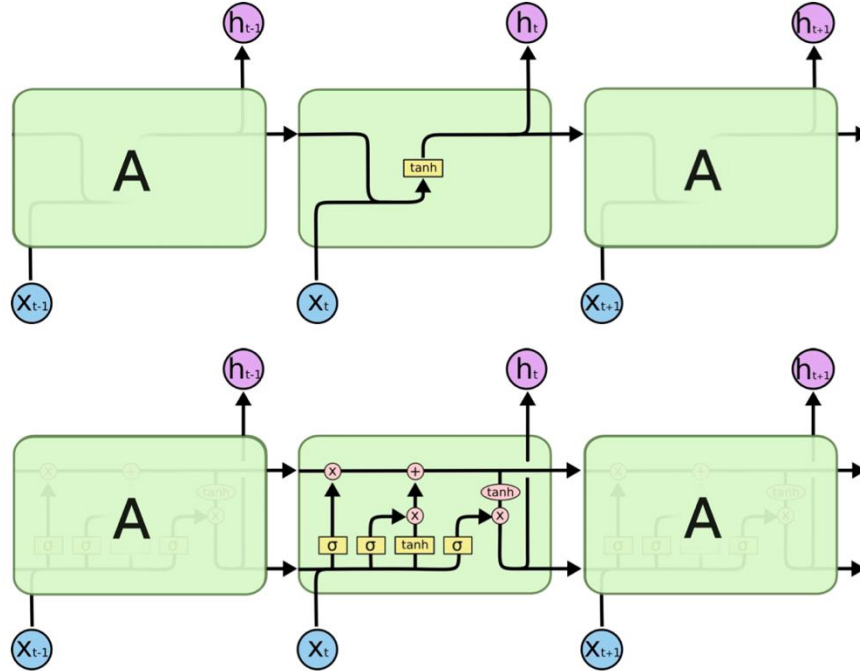
$$h_t = f(Ux_t + Wh_{t-1})$$

$$y_t = \text{softmax}(Vh_t)$$

Denklem 17. Tekrarlama Katmanı Denklemi

Tekrarlamalı Yapay Sinir Ağlarında, uzun dönem hafıza bulunmamaktadır, uzak geçmişin bilgisine sahiptir değildir ve uzun bağımlılıklara sahip desenlerden öğrenemez. Bu yüzden modelimize son olarak Uzun Kısa Dönem Hafıza eklememiz gerekmektedir.

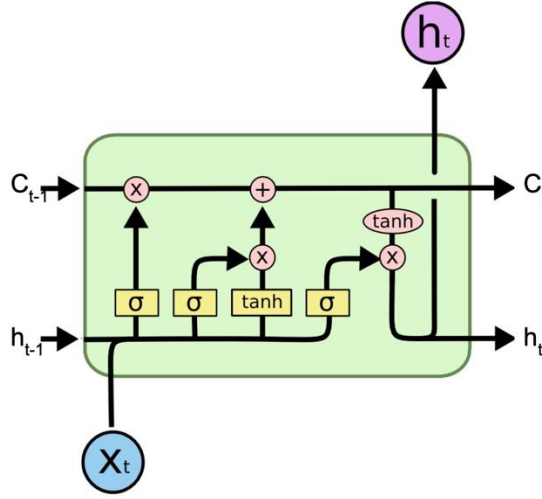
Uzun-Kısa Dönem Hafıza, Tekrarlamalı Yapay Sinir Ağlarının özel bir türüdür. Uzun dönem desenleri öğrenebilir ve 100 adımlı desenleri tespit edebilir.



Şekil 44. Tekrarlamalı Yapay Sinir Ağları ve Uzun Kısa Dönem Hafıza

Şekil 44'te Tekrarlamalı Yapay Sinir Ağları ve Uzun Kısa Dönem Hafıza mimarisi gösterilmiştir.

Bir Uzun Kısa Dönem Hafıza, temel bir Tekrarlamalı Yapay Sinir Ağını barındırır, Unutma Kapısına, Girdi Kapısına ve Çıktı Kapısına sahiptir. Kapılar filtre görevi görmektedir.



Şekil 45. Uzun Kısa Dönem Hafıza

Şekil 45'te Uzun Kısa Dönem Hafıza verilmiştir. Uzun Kısa Dönem Hafızada belirtilen öğeler aşağıda verilmiştir:

$\tanh$ = Temel Tekrarlamalı Yapay Sinir Ağı

h_t = Çıktı

X_t = Girdi

C_t = Hücre durumu

$h_{t-1} - h_t$ = Kısa dönem hafıza

$C_{t-1} - C_t$ = Hücre durumu. Hücre durumu hattında unutulacak ifadeler x ile hatırlanacak ifadeler ise + ile ifade edilir. Hücre durumları iki kez güncellenir.

Denklem 18'de çıktı (h), girdi (X) ve hücre durumu (C) olmak üzere Unutulacak Hücre Denklemi verilmiştir.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

$$C_t^f = C_{t-1} * f_t$$

Denklem 18. Unutulacak Hücre Denklemi

Denklem 19'da çıktı (h), girdi (X) ve hücre durumu (C) olmak üzere Girdi Katmanı Denklemi verilmiştir.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$C_t' = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

$$C_t^i = C_t' * i_t$$

Denklem 19. Girdi Katmanı Denklemi

Denklem 20'de Hücre Durumu (C) Denklemi verilmiştir.

$$C_t = C_t^f + C_t^i$$

Denklem 20. Hücre Durumu Denklemi

Denklem 21'de çıktı (h), girdi (X) ve hücre durumu (C) olmak üzere Çıktı Katmanı Denklemi verilmiştir.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

Denklem 21. Çıktı Katmanı Denklemi

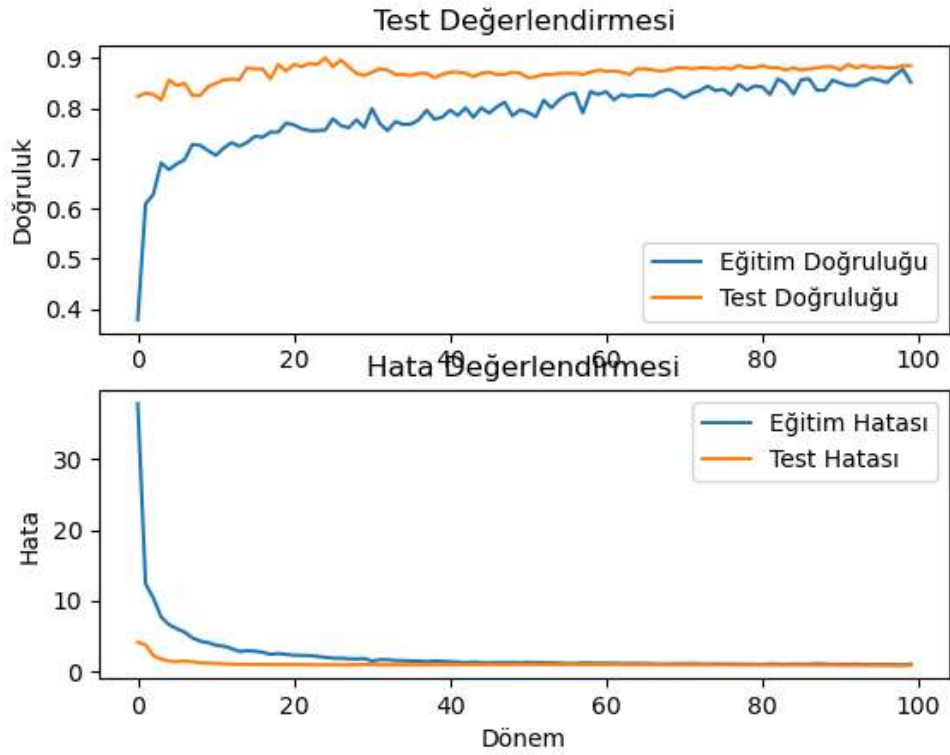
Yapay sinir ağları uyguladığımız veri setlerimizdeki aşırı uyum gösterme sorununu çözmek için yine Doğrultulmuş Lineer Birim kullanılmıştır bu sefer veriler eğitilirken her bir yoğunluktan sonra nöron bırakma uygulanmıştır. Nöron bırakma uygulaması sırasında kullanılan ağ topolojisi katmanında verilen değerlerin program çıktısı **Tablo 2'**de gösterilmiştir.

Tablo 2. Nöron Bırakma ile Ağ Topolojisi Katmanı Özellikleri ve Değerleri

Katman (tip)	Çıktı, Şekil	Değer #
Düzleştirilmiş (Flatten)	(Yok, 169)	0
Yoğunluk (Dense)	(Yok, 512)	87040
Bırakma (Dropout)	(Yok, 512)	0
Yoğunluk_1 (Dense)	(Yok, 256)	131328
Bırakma_1 (Dropout)	(Yok, 256)	0
Yoğunluk_2 (Dense)	(Yok, 64)	16448
Bırakma_2 (Dropout)	(Yok, 64)	0
Yoğunluk_3 (Dense)	(Yok, 10)	650
Toplam değer: 235466		
Eğitilebilir değer: 235466		
Eğitilemez değer: 0		

Tablo 2'de Nöron Bırakma ile Ağ Topolojisi Katmanı Özellikleri ve Değerleri verilmiştir.

Verilen, verilen değerler ile yüz kez eğitilmiştir. Eğitim başlarken doğruluk oranı 0,2206 iken eğitim sonrasında 0,8507 ve kayıp oranı 60,2963 iken eğitim sonrasında 1,0290 olmuştur.



Şekil 46. Aşırı Uyumluk Sonrası Eğitim ve Hata Değerlendirmeleri

Şekil 46'da Aşırı Uyumluk Sonrası Eğitim ve Hata Değerlendirmeleri grafikleri gösterilmiştir. Grafiklerde de görüldüğü gibi eğitimin başlarında eğitim doğruluğu ile test doğruluğu arasında çok fark varken eğitim devam ettikçe eğitim verisinin doğruluğu test verisinin doğruluğuna yaklaşmıştır. Aynı şekilde eğitim başlarında çok yüksek eğitim hata oranı olduğu gözlemlenmiştir, eğitimin sonuna doğru eğitim hatası minimuma indirilmiştir. Ancak yapı hala kararsızdır.

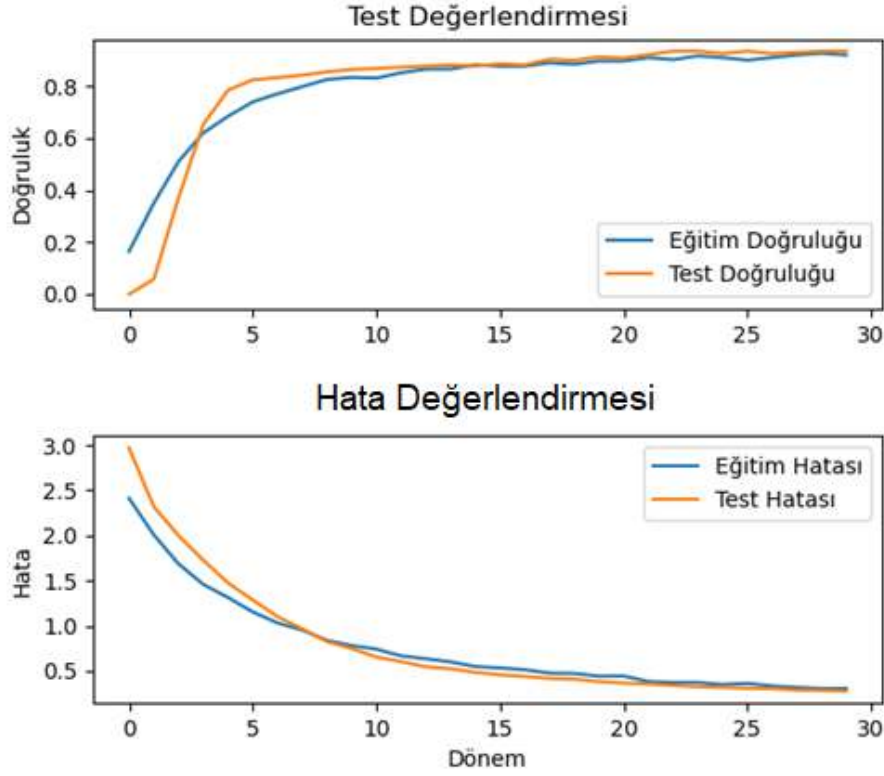
Modelimizi daha kararlı bir yapıya taşımak için Doğrultulmuş Lineer Birimi, Evrişimli Sinir Ağları ile desteklememiz gerekmektedir. Modelimizde kullanılan ağ topolojisi katmanında kullanılan değerlerin program çıktısı **Tablo 3**'te gösterilmiştir.

Tablo 3. Evrişimli Yapay Sinir Ağları Destekli Ağ Topolojisi Katman Özellik ve Değerleri

Katman (tip)	Çıktı, Şekil	Değer #
Dönüşüm_2D (Conv2D)	(Yok, 11, 11, 32)	320
Maks_Havuzlama2D (MaxPooling ...)	(Yok, 6, 6, 32)	0
İş_Parçacığı_Normalleştirme (Batch ...)	(Yok, 6, 6, 32)	128
Dönüşüm_2D_1 (Conv2D)	(Yok, 4, 4, 32)	9248
Maks_Havuzlama2D_1 (MaxPooling ...)	(Yok, 2, 2, 32)	0
İş_Parçacığı_Normalleştirme_1 (Batch ...)	(Yok, 1, 1, 32)	128
Dönüşüm_2D_2 (Conv2D)	(Yok, 1, 1, 32)	4128
Maks_Havuzlama2D_2 (MaxPooling ...)	(Yok, 1, 1, 32)	0
İş_Parçacığı_Normalleştirme_2 (Batch ...)	(Yok, 1, 1, 32)	128
Düzleştirilmiş (Flatten)	(Yok, 32)	0
Yoğunluk (Dense)	(Yok, 64)	16448
Bırakma (Dropout)	(Yok, 64)	0
Yoğunluk_1 (Dense)	(Yok, 10)	650
=====		
Toplam değer: 16842		
Eğitilebilir değer: 16650		
Eğitilemez değer: 192		

Tablo 3'te Evrişimli Yapay Sinir Ağları Destekli Ağ Topolojisi Katman Özellik ve Değerleri verilmiştir.

Evrişimli Sinir Ağları uygulanan modelimiz 30 kez eğitilmiştir. Eğitim öncesinde doğruluk oranı 0,1165 iken eğitim sonrasında doğruluk oranı 0,8939 ve eğitim öncesinde kayıp oranı 2,5465 iken eğitim sonrasında 0,2994 olmuştur.



Şekil 47. Evrişimli Yapay Sinir Ağı İle Test ve Hata Değerlendirmesi

Şekil 47'de Evrişimli Yapay Sinir Ağı İle Test ve Hata Değerlendirmesi grafikleri gösterilmiştir.

Evrişimli Yapay Sinir Ağları uygulanan modelimiz sonunda *Şekil 47*'de görüldüğü gibi 10. dönemden sonra test ve eğitim doğruluğu birbirine paralel ve yakın değerlerde seyretmektedir. Eğitim ve test hataları da daha küçük ve yakın değerlerde logaritmik olarak azalmaktadır.

Evrişimli Yapay Sinir Ağları, hataları kısa dönem hafızada tutmamaktadır. Bu yüzden Uzun-Kısa Dönem Hafıza kullanım gereksinimi doğmaktadır. LSTM ile eğitim esnasında *Tablo 4*'te gösterilen ağ topolojisi katman özellik ve değerleri ile değerlendirme yapılmıştır.

Tablo 4. Uzun-Kısa Dönem Hafıza Destekli Ağ Topolojisi Katman Değer ve Özellikleri

Katman (tip)	Çıktı, Şekil	Değer #
Uzun-Kısa_Dönem_Hafıza (LSTM)	(Yok, 13, 64)	19968
Uzun-Kısa_Dönem_Hafıza_1 (LSTM)	(Yok, 64)	33024
Yoğunluk (Dense)	(Yok, 64)	4160
Bırakma (Dropout)	(Yok, 64)	0
Yoğunluk_1 (Dense)	(Yok, 10)	650
=====		
Toplam değer: 57802		
Eğitilebilir değer: 57802		
Eğitilemez değer: 0		

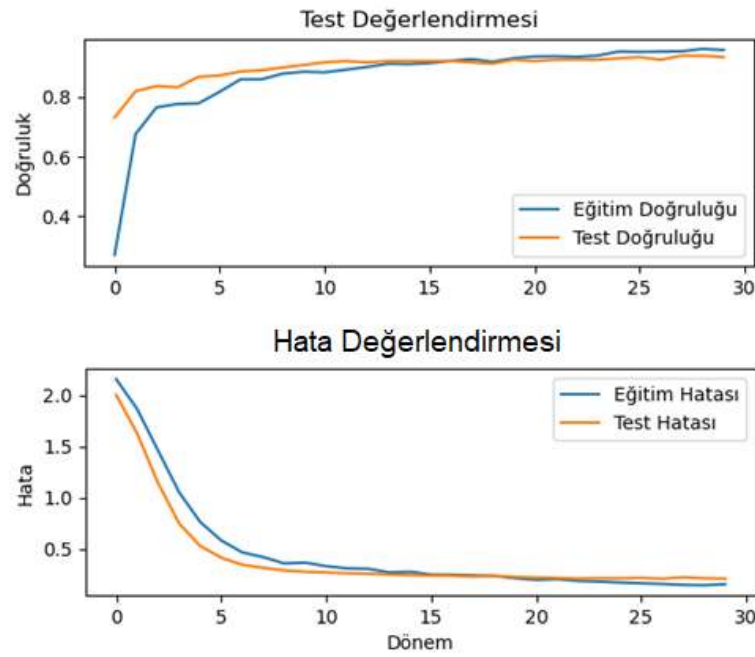
Tablo 4'te Uzun-Kısa Dönem Hafıza Destekli Ağ Topolojisi Katman Değer ve Özellikleri verilmiştir.

Veriler 30 kez eğitilmiştir. Eğitim başlarken seçilen veri setlerinin doğruluk oranları 0,1891 iken eğitim sonrasında doğruluk oranları 0,9151 ve kayıp oranları 2,2099 iken eğitim sonrasında kayıp oranları 0,2358 olmuştur.

Uzun-Kısa Dönem Hafızanın birden fazla katmanda kullanılmasının ana nedeni, daha fazla model karmaşıklığına izin vermektir. Basit bir ileri beslemeli ağ topolojisinde, girdi verilerin hiyerarşik bir özellik temsilini oluşturmak için birden fazla katman kullanılır ve bunlar makine öğrenimi görevlerinde kullanılır. Bu durum Uzun-Kısa Dönem Hafızada da geçerlidir. Girdi hali hazırda bir Uzun-Kısa Dönem Hafıza katmanının sonucu ise mevcut Uzun-Kısa Dönem Hafıza, mevcut girdinin daha karmaşık bir özellik temsilini oluşturabilir.

Uzun-Kısa Dönem Hafıza sonrasında İleri Besleme katmanı da kullanılabilir. Modelimizde yeni bir Uzun-Kısa Dönem Hafıza kullanılmasını sebebi, ileri besleme katmanının, önceki zaman adımından geri bildirim almaması ve bu nedenle girdiyi yorumlayamamasıdır. Bunun yerine yeni bir Uzun-Kısa Dönem Hafıza kullanıldığı takdirde her katmanda daha karmaşık girdi modelleri tanımlanabilir.

Katmanların sayısı arttıkça zamana bağlı kayıp miktarı da artacaktır, bu yüzden Uzun-Kısa Dönem Hafıza katmanı iki adet ile sınırlandırılmıştır.

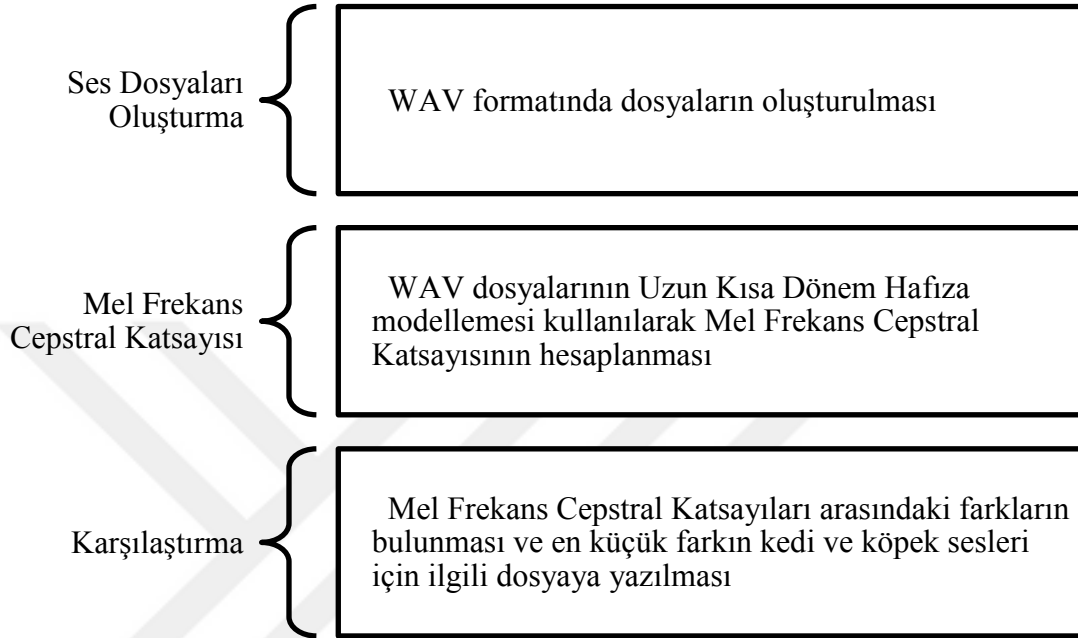


Şekil 48. Uzun-Kısa Dönem Hafıza Sonrası Test ve Hata Değerlendirmeleri

Şekil 48'de Uzun-Kısa Dönem Hafıza Sonrası Test ve Hata Değerlendirmeleri grafikleri gösterilmiştir.

Tezimde ilk kelimelerin oluşumunda çevre seslerden etkilendiği savunulmuştur. Bu bağlamda daha önce Latin alfabesi seçimi yapıldığı için kullandığım makinenin performansına uygun ve konuşma dili İngilizce olduğu için İngiliz alfabesinde bulunan A-Z arası harfler a, aa, aaa, aaaa şeklinde $P(32,1)$, $P(32,2)$, $P(32,3)$ ve $P(32,4)$ permütasyonları ile WAV formatında dosyalar oluşturulup bu dosyaların da sınıflama modellenmesinde olduğu gibi Mel Frekans Cepstral Katsayısı hesaplaması yapılması gerekmektedir.

Mel Frekans Cepstral Katsayılarının görüntü işleme modelli ses dosyası sınıflama modellemesi ile karşılaştırılarak en yakın ses dosyasının seçimini yapılması gerekmektedir.



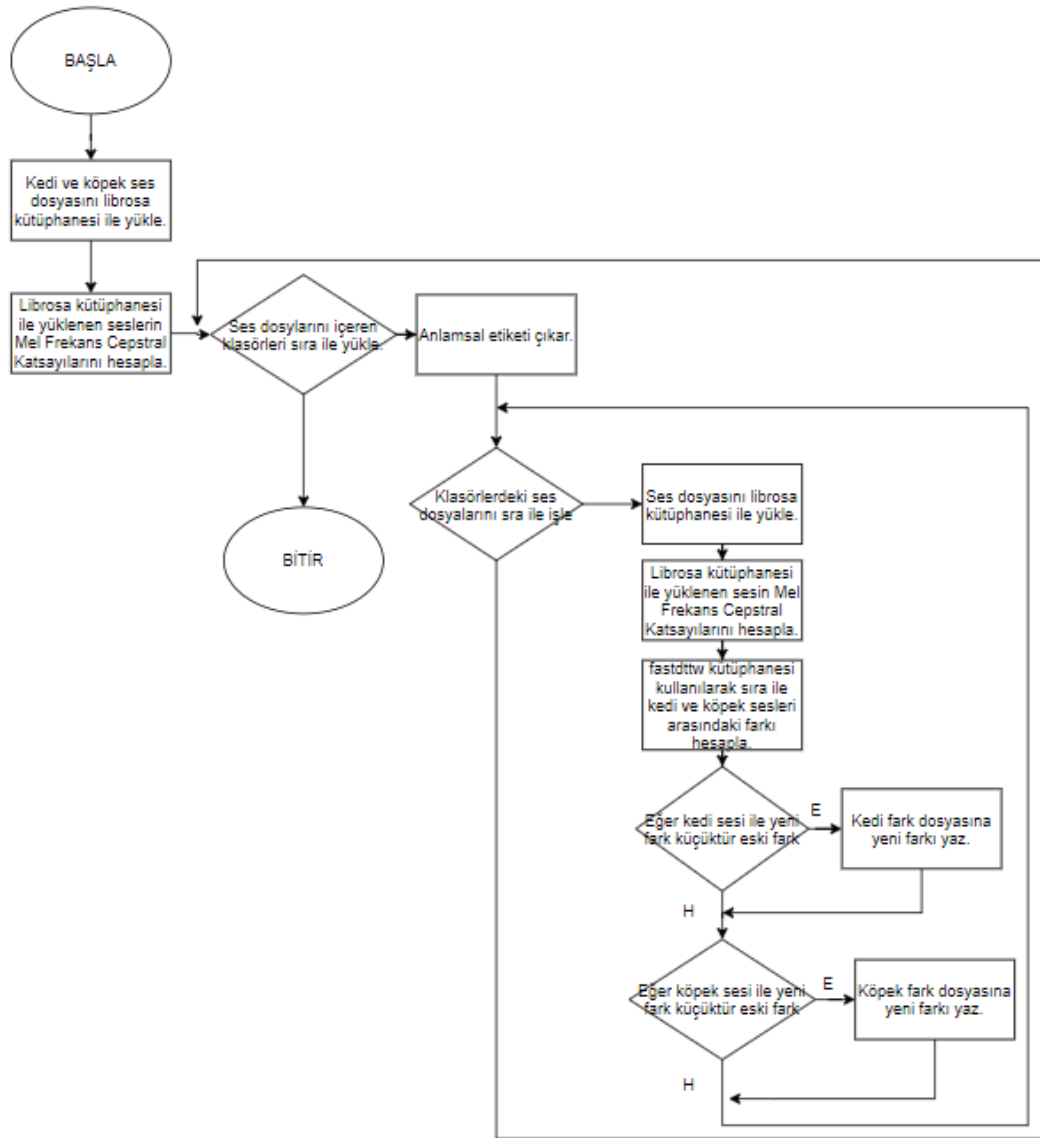
Şekil 49. Veri Setine Karşılık Ses Dosyalarının Üretilmesi

Şekil 49'de Veri Setine Karşılık Ses Dosyalarının Üretilmesi adımları gösterilmiştir.

Modellemesi yapılan kedi ve köpek seslerini ayırt edebilen makinelerin daha önce seçmiş olduğu Latin alfabesinden türetilmiş kelimeler arasından kedi ve köpek seslerine en yakın gördüğü seçmesi sağlanmıştır.

Öncelikle, pytsx3 kütüphanesi Latin alfabesinde kelimeler türetilmiştir. Makinemin konuşma dili İngilizce olduğu için İngiliz alfabesinde okunduğu şekli ile WAV dosyaları oluşmuştur. Makine performansı dolayısı ile 38.3 Gigabyte boyutunda 495959 adet ses dosyası türetilenmiştir.

Türetilen ses dosyalarının Mel Frekans Cepstral Katsayısı hesaplandıktan sonra kedi ve köpek seslerinden seçilen örneklemin Mel Frekans Cepstral Katsayısı ile karşılaştırılmış ve hesaplanan sonuçlar Kedi ve Köpek için ayrı ayrı dosyalara yazılmıştır.



Şekil 50. Mel Frekans Cepstral Katsayısı ile Ses Karşılaştırması

Şekil 50' de Mel Frekans Cepstral Katsayısı ile Ses Karşılaştırması algoritması gösterilmiştir.

Karşılaştırmaların sonucunda makine köpek sesine karşılık “*xoxo*” kelimesini, kedi sesine karşılık ise “*qtkp*” kelimesini seçmiştir.

SONUÇ VE ÖNERİLER

Çalışma için Kaggle Audio Cats and Dogs veri seti kullanılmıştır. Makineye kedi ve köpek sesleri ile eğitilmiştir. Klasik yaklaşımlarda olduğu gibi öncelikle ses dosyalarından özellik çıkarımı yapılmıştır. DTW ile kedi ve köpek seslerinin grafikleri üretilmiştir. Sesler eğitilirken bazı sorunlar ile karşılaşmıştır bunlar, eğitim verisi hatası ile test hatası arasındaki farkın artması, yapay sinir ağlarının düzensizliği, aşırı uyumluluk ve kısa ve uzun dönem hafıza sorunları olarak özetlenebilir. Bunların çözmek için erken durma, nöron bırakma, evrişimli yapay sinir ağları kullanımı ve uzun-kısa dönem hafıza algoritmaları kullanılmıştır.

Eğitim tamamlandıktan sonra makineye Latin alfabesi ile kelimeler türetilmiştir ve bu kelimelerin Mel Frekans Cepstral Katsayısı hesaplanmıştır.

Latin alfabesinden türetilen seslerin Mel Frekans Cepstral Katsayıları ile veri setinin Mel Frekans Cepstral Katsayıları, DTW algoritması kullanılarak karşılaştırılmıştır. DTW’de fark sıfıra ne kadar yakın ise sesler de birbirlerine o kadar benzer sesler olarak kabul edilmektedir.

Model oluşturulurken Karmaşık periyodik sesin farklı frekanslarda salınan sinüs dalgalarının toplamına ayrıştırılmasını sağlamak için Fourier Dönüşümü kullanılmıştır ancak Fourier Dönüşümü yapıldığında ses dosyalarındaki zaman bilgisi kaybolmaktadır. Bu yüzden farklı aralıklar ile birden çok Fourier Dönüşümü uygulanarak Kısa Zaman Fourier Dönüşümü yaklaşımı benimsenmiştir. Modelimiz artık zaman, frekans ve desibel bilgilerine sahip olmuştur. Sesin tını/dokusal yönlerini yakalamak için MFCC bilgisi çıkartılmıştır. Ancak burada model eğitilirken degradenin kaybolma olasılığı çok yüksek olduğu için daha iyi bir yakınsama sağlayan ReLU katmanları modele uygulanmıştır. ReLU katmanları uygulandıktan sonra sinir ağında aşırı uyum sorunu olabileceği görülmüştür. Nöron bırakma, erken durma ve düzenlilik fonksiyonları uygulanarak aşırı uyumdan kaçınılmıştır.

Geleneksel yaklaşımların aksine ses dosyalarından elde edilen MFCC bir görsel olarak ele alınmıştır ve çok katmanlı algılayıcıdan daha iyi performans gösteren, yoğunluk katmanlarına göre daha az parametre gereksinimi duyan Evrişimli Sinir Ağları kullanılmıştır. MFCC katmanlarında bulunan ardışık verilerin işlenmesi için

RNN katmanları modele eklenmiştir. RNN katmanları uzak geçmiş bilgisini tutmaz ve uzun dönem hafızaya sahip değildir. Uzun dönem hafıza bilgisine sahip olabilmek için RNN'nin özel türü olan LSTM tercih edilmiştir ve modelin eğitilmesi tamamlanmıştır.

Çalışma sonucunda ise QTKP harflerinden oluşan kelime kedi sesine en yakın ses olarak belirlenmiş ve fark 56823.36 olarak hesaplanmıştır. Köpek sesi için ise XOXO harflerinden oluşan kelime en yakın ses olarak belirlenmiş ve fark 26862.43 olarak hesaplanmıştır. Farklar çok fazla olsa dahi makinenin doğal dilinin ilk sesleri olarak kabul edilebilir. Makine performansı arttırıldığında daha fazla ses ile karşılaştırma yapılabilir veya makinenin kendine özgü sesler ve alfabe üretilmesi sağlanabilir.

Bu çalışmada makine DTW, RNN ve MFCC kullanarak kendi diline ait kelimeleri oluşturmayı başarmıştır.

Çalışma sonucunda makine kedi ve köpek seslerini tanımlayıp, sınıflandırabilmiş ve onlara karşılık sesler üretebilmiştir. Böylece makine artık canlıları sınıflandırabilme yeteneğine sahipken onlara karşılık sesler de üretebilmektedir.

Makinenin ürettiği sesler çok ilkel olup kendisi yeni sesler üretmek yerine var olan diller ile etkileşimde bulunarak var olan doğal dillerden yeni kelimeler türetmiştir. Daha fazla veri seti ile başka canlılar ve sesler için de doğal dile katkıda bulunulabilir. Çalışmanın sonucunda görülmüştür ki doğal diller de öncelikle başka canlıların çıkardığı seslere karşılık yanıtlar üretilip canlıları sınıflandırarak oluşmuş ve diğer toplum ve diller ile etkileşime geçtiği zaman diller zenginleşmeye başlamıştır.

Makinelerin kendine özgü doğal dillerini geliştirmesi konusunda yapılmış çok fazla çalışma bulunmamakla birlikte bu alanda yayın da bulmak çok güçtür. Çalışma, bu alandaki ender çalışmalar arasındadır ve gelecek çalışmalara ışık olması umulmaktadır.

Bu çalışma, insanların ilk dillerinin diğer canlılar ve nesneler ile etkileşime geçip çıkan sesleri taklit ederek oluşturduğu tezlerini savunmakta ve bu yöntemle makinenin kendi dilini oluşturmasını sağlamıştır. Çalışma sonunda ilk seslerin diğer nesnelerin ve canlıların çıkardığı seslerin taklit edilmesi ile oluşabildiği

gözlemlenmiştir ve makinenin ilk ilkel kelimelerini üretmesi dil bilimi alanında insanların doğal dillerini nasıl oluşturduğu konusunda yön gösterecektir.

Çalışma sonucunda doğada bulunan canlıların, ilk dillerde çıkardıkları seslere benzer sesler üretilmesi ile seslendirildiği ve diğer toplumlar ve diller ile etkileşime geçildikçe seslerin daha da anlaşılır hale geldiği gösterilmiştir.

Çalışma ile makineler doğadaki sesleri sınıflandırma yetisi kazanmıştır. İnsanlar da canlıları sınıflandırma yetisine sahiptir. İlk doğal diller olduğu zaman makinelerde yaptığımız çalışma sonucunda bu sınıflandırmalar yardımı ile sınıflandırdığı canlıların seslerini taklit ettiği ortaya konulmuştur. Diğer diller ile etkileşime geçmeye başladığında ise diller zenginleşmiştir.

Çalışmada var olan metotlar ile yeni yaklaşım türetilmiş, sestten konuşmaya, konuşmadan MFCC'ye dönüşüm yapılmış ve doğadaki sesler ile karşılaştırmalar yapılmıştır. Bebeklerin de ilk çıkardığı seslere baktığımızda anlamsız sesler çıkarıyor gibi görünseler de etrafındaki nesneleri, canlıları veya dertlerini ifade etmeye çalışmaktadırlar. Makineler de ilk olarak anlamsız görünen kelimeler üretse de ilerleyen çalışmalarda diğer diller ile de etkileşime geçerek daha anlamlı kelimeler ve cümleler üretebilecektir.

Bu çalışma sayesinde makinelerin kendi dillerini üretebilmesi için yaklaşım geliştirilmiştir. Böylece insansı makinelerin oluşturulma sürecinde bu makinelerin de tıpkı insanların sahip olduğu gibi kendi dilleri ile iletişim kurabilme süreçlerine katkıda bulunulmuştur. Doğal dillerin oluşumu sürecinde de dil bilimcilere yol göstermesi hedeflenmiştir.

Çalışma sonucunda makine evrimsel sinir ağlarını kullanarak canlıları ayırt etme yetisi kazanmıştır. Makine yeni sesler türetme yetisi kazanmıştır. Makine sesleri analiz edip, bir bebeğin ilk kelimeleri duyarak kopyalaması misali sınıflandırdığı canlıları isimlendirebilmiştir. Makine, bir bebek gibi söyleyeni tekrarlayarak yeni kelimeler üretebilme yetisi kazanmıştır. Dil bilimde insanların nasıl konuşmayı öğrendiği sorusuna yapılan çalışma ile sesleri taklit ederek karşılık vermesi ile bulmuş olabilir teorisi oluşturulabilmektedir. Ses işleme algoritmalarına, görüntü işleme algoritmaları eklendiğinde daha etkin sonuçlar elde edildiği görülmüştür. Yapay Sinir Ağları, MFCC, DTW bir arada kullanılarak yeni model geliştirilmiştir. Veri seti için

en yüksek doğruluk ve en düşük kayıp oranı LSTM'e sahip RNN ile desteklenen ReLU modeli ile sağlanmıştır.

Bu çalışma sonucunda makineye limitli insan bilişinden sıyrılıp kendi geliştirdiği anahtar ifadeler ile karmaşık problemleri mukayese yetkinliğine kavuşmasına fırsat verilmiştir.

Bu çalışmadaki limitli makine performansı sebebi ile 38.3 Gigabyte boyutundaki veri ile ses dosyalarının karşılaştırılması ve var olan canlıların isimlendirilmesi sağlanmıştır.

Kullanılan makinenin dili İngilizce olduğu için Python dilinin ilgili kütüphanesi sadece İngiliz alfabesi ile kelimeler üretebilmiştir. Gelecek çalışmalarda türetilen kelimeler geliştirilebilir.

Çalışmada tekil canlı isimleri türetilmiş olup birden fazla canlı için kullanılacak ekler belirlenmemiştir. Örneğin Türk dilinde eğer birden fazla kedi varsa kediler, İngiliz dilinde ise cats şeklinde ifade edilmektedir. Makinenin şu anda ses dosyalarındaki canlıların sayısını tespit etme yeteneği bulunmamaktadır. Ayrıca dillerin belirgin özelliklerinden olan bağlaçlar, sıfatlar, zarf ve zamirler de bulunmamaktadır.

Gelecek çalışmalarda görüntü işleme yöntemleri de kullanılarak canlı sesleri bu sesler ile eşleştirilip renk, uzaklık, yakınlık, boy, hacim gibi kavramlar da makineler tarafından türetilir.

Bu tür çalışmalar ilerledikçe insansı makineler, insanlar gibi düşünüp kendi dilleri aracılığı ile kendi aralarında anlaşabilecek ve bunun yanı sıra ilk doğal dillerin de nasıl ortaya çıktığı netlik kazanabilecektir.

Bu çalışmanın kısıtları olarak sadece kedi ve köpek seslerinden seçilen örnekler ile çalışılmış bunlara ait birer kelime türetmesi sağlanmış ve türetilen kelimenin geliştirilmesine müsaade edilmeyip ilkel hali ile bırakılmıştır. Birden çok canlı içeren sesler tek bir canlı olarak işlem gördürülmüş ve çoğul/tekil canlılar/nesneler arasında ayırım yaptırılmamıştır. Bu sebeple makine çoğul canlılar/nesneler için yeni kelime veya çoğul eki türetmemiştir. Daha tutarlı sonuçlar için veri boyutunun küçük tutulmasına önem verilmiştir.

KAYNAKÇA

- Abhimanyu C., Abhinav P., Chandersh S., Natural Language Processing, International journal of technology enhancements and emerging engineering research, PennState Üniversitesi, 2013.
- Acarer, T. İ., SEEK ve Citymatrix Projeleri Üzerinden Mimarlık ve Yapay Zeka Çalışmaları Üzerine Bir Tartışma, Yüksek Lisans, Yeditepe Üniversitesi 2021.
- Almahdi, M. E., Doğal Dillerin Zaman ve Konuma Bağlı Değişimlerinin Nicel Olarak Ölçülmesi, Yüksek Lisans, Gebze Teknik Üniversitesi, 2020.
- Altun, H. E., Kuantum Yapay Zeka Algoritmalarının Analizi: Grover Araması Ve Doğal Dil İşleme İle Kuantum Haberleşmesi, Yüksek Lisans, Bahçeşehir Üniversitesi, 2021.
- Ayata, D., Türkçe ve İngilizce Twitter Duygu Sınıflandırması İçin Makine Öğrenmesi ve Doğal Dil İşleme Tekniklerinin Uygulanması, Yüksek Lisans, Boğaziçi Üniversitesi, 2018.
- B.H. Juang, L. R., Automatic Speech Recognition - A Brief History of the Technology Development, Georgia Teknoloji Enstitüsü. Atlanta Rutgers Üniversitesi ve California Üniversitesi, 2005.
- Bozyiğit, F., Doğal Dil İşleme Kullanımı İle Nesne Tabanlı Analiz Ve Kaynak Kod Değerlendirmesi, Doktora, Dokuz Eylül Üniversitesi, 2020.
- Brill, E., Mooney, R. J, An overview of empirical natural language processing, AI Magazine 18-4, 1997.
- Büyüköz, B., Doğal Dil İşlemede Bağlam Odaklı Derin Öğrenme Yöntemlerinin Transfer Kapasitesinin Metin Sınıflandırması Problemi Üzerinden İncelenmesi, Yüksek Lisans, Boğaziçi Üniversitesi, 2016.
- Cengiz, M., Doğal Dil İşleme Temelli Görsel Senaryo Taslağı Motoru Geliştirilmesi, Doktora, Dokuz Eylül Üniversitesi, 2020.
- Chowdhary, Natural language processing, Fundamentals of Artificial Intelligence, Springer, 2020.

- Danışman, Ö., Farklı bağımlılık varsayımlarına dayalı saklı Markov modelleri, Doktora, Dokuz Eylül Üniversitesi, 2021.
- Dave, V. S., Dutta, K., Neural network based models for software effort estimation: a review, Artificial Intelligence Review, Springer, 2014.
- Demir, F., Doğal Dil İşleme Kullanarak Türkçe Yazılım Gereksinimlerinden Sınıf Diyagramları Oluşturma, Yüksek Lisans, Atılım Üniversitesi, 2021.
- Deneç, G., Canlı Ses İşleme ve Geribesleme Kullanılarak Sempatik Titreşen Teller ile Sound Mass Tarzı Sesler Yaratmak, Doktora, İstanbul Teknik Üniversitesi, 2015.
- Doğan, İ., Yeni Çağ Pazarlamada Yapay Zeka Uygulamaları: Yapay Zeka Mobil Bankacılık Uygulamaları Hakkında Bir Analiz, Yüksek Lisans, Bilgi Üniversitesi, 2018.
- Durna, M., Doğal Dil İşlemesi ve Makine Öğrenmesi ile Türkçe Haberlerde Olay Göstergesi Tespiti, Yüksek Lisans, Boğaziçi Üniversitesi, 2019.
- Duşcu, M., Makine Öğrenmesi ve Ses İşleme Kullanılarak Sosyal Ağlarda Duygu Analizi, Yüksek Lisans, Özyeğin Üniversitesi, 2019.
- Erdal, C., Bulanık mantık ve firmaların başarı kriterlerinin tanımlanarak bulanık mantık ile ölçülmesinin bir uygulaması, Yüksek Lisans, Kırıkkale Üniversitesi, 2008.
- Ergüven, Ö. S., Doğal dil işlemede semantik gösterimlerin sistematik değerlendirilmesi, Yüksek Lisans, İzmir Yüksek Teknoloji Üniversitesi, 2018.
- Fine, S., Singer, Y., Tishby, N., The hierarchical hidden Markov model: Analysis and applications, Machine learning, Springer, 1998.
- Hekimoğlu, B., Değişken yük durumunda paralel aktif güç filtresinin uyarlamalı yapay sinir ağları ile denetimi, Doktora, Kocaeli Üniversitesi, 2010.
- Indurkha, N., Damerau, F. J., Handbook of natural language processing, Chapman and Hall/CRC, 2010.
- Jahnavi, M., Introduction to Neural Networks, Advantages and Applications, Towards Data Science, 2017.

- Keklik, O., Doğal Dil İşleme Tekniklerini Kullanarak Otomatik Soru Üretme , Yüksek Lisans, İzmir Yüksek Teknoloji Üniversitesi, 2018.
- Kothari, P. R., Processing natural language requirement to extract basic elements of a class, International Journal of Applied Information Systems, PennState Üniversitesi, 2012.
- Lanzetta, M., Machine Learning, Deep Learning, and Artificial Intelligence, Artificial Intelligence for Autonomous Networks, Chapman and Hall/CRC, 2018.
- Liddy, E. D., Natural language processing, Syracuse Üniversitesi, 2001.
- Maind, S. B., Research paper on basic of artificial neural network, International Journal on Recent and Innovation Trends in Computing and Communication 2.1, 2014.
- Novichkova, S., Egorov, S., & Daraselia, N., MedScan, a natural language processing engine for MEDLINE abstracts, Bioinformatics 19.13, Oxford University Press, 2003.
- Özcan, D. G., Doğal Dil İşleme Tekniklerinin İncelenmesi ve Seyahat-Turizm Sesli Asistanı için Türkçe Varlık İsmi Tanıma Aracı Geliştirilmesi, Yüksek Lisans, Akdeniz Üniversitesi, 2020.
- Özkan, B., Doğal Dil İşleme Yöntemleri İle Diyalog Niyet Sınıflandırması, Yüksek Lisans, Bahçeşehir Üniversitesi, 2021.
- Piczak, K. J., ESC: Dataset for Environmental Sound Classification, Proceedings of the 23rd ACM international conference on Multimedia 2015.
- Pirim, H., Yapay Zeka, Yaşar Üniversitesi Dergisi 1.1, 2006.
- Saygılı, H., Yapay Zeka Kamu Politikaları: ABD ve Çin'in Yapay Zeka Kamu Politikalarının Karşılaştırılması, Yüksek Lisans, Ankara Yıldırım Beyazıt Üniversitesi, 2020.
- Searle, J., Minds, Brains, and Programs, The Behavioral and Brain Sciences, 1980.
- Senin, P., Dynamic time warping algorithm review, Hawaii Üniversitesi, Bilişim ve Bilgisayar Bilimleri Departmanı, 2008.

- Stuart J. R., Peter N., Artificial Intelligence: A Modern Approach, Pearson, 3. Basım, 2009.
- Toprak, A., Doğal Dil İşleme İle İngilizce Otomatik Sözlük Oluşturma, Yüksek Lisans, İstanbul Ticaret Üniversitesi, 2019.
- Widrow, B., & Lehr, M. A., Perceptrons, Adalines, and backpropagation, Handbook of Brain Theory and Neural Networks, MIT Press, 1995.
- Winograd, T., Five Lectures on Artificial Intelligence, Standford Üniversitesi, Bilgisayar Bilimleri Bölümü, 1974.
- Yazgaç, B. G., Süne ve Kımıl Zararlılarının Ses İşleme Yöntemleri ile Sınıflandırılması ve Bir Gömülü Sistem Gerçeklemesi, Yüksek Lisans, İstanbul Teknik Üniversitesi, 2017.
- Yegnanarayana, B., Artificial neural networks, PHI Learning Pvt. Ltd, 2009.
- Yılmaz, A., Sinirsel Bulanık Mantık Modeliyle Kanser Risk Analizi, Doktora, Sakarya Üniversitesi, 2015.
- Yılmaz, A., Yapay Zekâ, Kodlab, 8. Baskı, 2020.
- Zadeh, L., Fuzzy Sets, Information and Control 8.3, Elsevier, 1965.

