



**HAVA TRAFİK AKIŞININ İYİLEŞTİRİLMESİ İÇİN  
RASSAL ORMAN VE DERİN SINIR AĞLARI TABANLI  
VARİŞ ZAMANI VE YÖRÜNGE TAHMİNİ  
Doktora Tezi**

**Onur BAŞTÜRK**

**Eskişehir 2021**

**HAVA TRAFİK AKIŞININ İYİLEŞTİRİLMESİ İÇİN RASSAL ORMAN VE  
DERİN SİNİR AĞLARI TABANLI VARİŞ ZAMANI VE YÖRÜNGE TAHMİNİ**

**Onur BAŞTÜRK**

**Doktora Tezi**

**Hava Trafik Kontrol Anabilim Dalı**

**Danışman: Doç. Dr. Cem ÇETEK**

**Eskişehir**

**Eskişehir Teknik Üniversitesi**

**Lisansüstü Eğitim Enstitüsü**

**Şubat 2021**

## ÖZET

### HAVA TRAFİK AKIŞININ İYİLEŞTİRİLMESİ İÇİN RASSAL ORMAN VE DERİN SİNİR AĞLARI TABANLI VARİŞ ZAMANI VE YÖRÜNGE TAHMİNİ

Onur BAŞTÜRK

Hava Trafik Kontrol Anabilim Dalı

Eskişehir Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Ocak 2021

Danışman: Doç. Dr. Cem ÇETEK

Bu çalışmada Yeşilköy Terminal Kontrol Sahası içerisinde yer alan İstanbul Atatürk Havalimanı'na iniş yapan uçakların varış zamanlarının ve terminal kontrol sahası içerisinde izledikleri yörüngelerin rassal orman ve derin sinir ağırları yöntemleri kullanılarak tahmini gerçekleştirilmiştir. 2018 yılı Nisan-Ağustos ayları arasında İstanbul Atatürk Havalimanı'na iniş yapan uçaklar için genel uçuş bilgileri, yörünge verileri ve havacılık amaçlı rutin hava raporları farklı kaynaklardan elde edilerek makine öğrenmesi yöntemlerinde kullanılacak eğitim ve test kümeleri oluşturulmuştur. Varış zamanının tahmini için rassal orman ve derin sinir ağırları yöntemleri kullanılarak 51.902 uçuşa ait verilerle makine öğrenmesi modelleri elde edilmiştir. Elde edilen makine öğrenmesi modellerinin planlanan varış zamanına göre tahmin hatalarını kalkıştan itibaren %64 oranında, terminal kontrol sahasına girişten itibaren ise %84 oranında azalttığı tespit edilmiştir. Benzer şekilde yörünge tahmini için derin sinir ağırları yöntemi kullanılarak 187.509 adet yörünge noktası için makine öğrenmesi modeli eğitilmiştir. Modelin doğruluğu uyum derecesi ( $R^2$ ) ile ölçülmüş ve tahmin edilen çıktılar için 0.96-0.97 aralığında değerler elde edilmiştir.

**Anahtar Sözcükler:** Hava trafik kontrol, Makine öğrenmesi, Rassal orman, Yapay sinir ağırları, Derin öğrenme

## ABSTRACT

### PREDICTION OF AIRCRAFT ESTIMATED TIME OF ARRIVAL AND TRAJECTORY USING RANDOM FORESTS AND DEEP NEURAL NETWORKS FOR IMPROVEMENT OF AIR TRAFFIC FLOW

Onur BAŞTÜRK

Department of Air Traffic Control

Eskişehir Technical University, Institute of Graduate Programs, January 2021

Supervisor: Assoc. Prof. Cem ÇETEK

In this study, arrival times and trajectories of flights to Istanbul Ataturk International Airport which resides in Yeşilköy Terminal Maneuvering Area, were predicted using random forests and deep neural network methods. General flight information, trajectory data and meteorological aerodrome reports for arrivals to Istanbul Ataturk International Airport between April-August 2018 were obtained from different data sources and training/test data sets to be used in machine learning methods were generated. Random forests and deep neural network methods were used for the prediction of arrival times and machine learning models were obtained using data for 51.902 flights. The machine learning models decreased the prediction errors with respect to the scheduled time of departure in 64% immediately after departure and 84% after entrance to terminal maneuvering area. Similarly, deep learning method was used for trajectory prediction and machine learning model was trained with 187.509 trajectory points. The accuracy of the model was measured using the degree of fitness ( $R^2$ ) and values close to 1 (0.96-0.97) were obtained for the predicted outputs which correspond to a high degree of fitness.

**Keywords:** Air traffic control, Machine learning, Random forest, Neural networks, Deep learning

## TEŞEKKÜR

Çalışmalarımız sırasında her aşamada yardımını ve desteğini hiçbir zaman esirgemeyen, karamsarlığa kapıldığım anlarda beni yüreklendirerek ve çalışmaya teşvik ederek doktora eğitim hayatımı başarıyla tamamlamamı sağlayan değerli danışmanım Doç. Dr. Cem ÇETEK'e sonsuz teşekkürlerimi sunarım.

Tez izleme jürimde yer alan, engin tecrübelerinden faydalanma fırsatı yakaladığım değerli akademisyenler Doç. Dr. Özlem ŞAHİN ve Dr. Öğr. Üy. Emre KOYUNCU'ya bu süreçte verdikleri destek ve önerilerden dolayı teşekkürlerimi sunarım.

Doktora eğitim hayatım boyunca hep yanımda olan ve bu süreci sıkıntısız bir şekilde atlatabilmem için emeğini benden hiçbir zaman esirgemeyen sevgili eşim Sinem ATABEY BAŞTÜRK'e sonsuz teşekkürlerimi sunarım.

Son olarak eğitim hayatımın her aşamasında çok önemli katkıları olan sevgili aileme teşekkürlerimi sunarım.

## **ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ**

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Eskişehir Teknik Üniversitesi tarafından kullanılan “bilimsel intihal tespit programı”yla tarandığını ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

Onur BAŞTÜRK

## İÇİNDEKİLER

|                                                          |     |
|----------------------------------------------------------|-----|
| BAŞLIK SAYFASI .....                                     | i   |
| JÜRİ VE ENSTİTÜ ONAYI .....                              | ii  |
| ÖZET .....                                               | iii |
| ABSTRACT .....                                           | iv  |
| TEŞEKKÜR .....                                           | v   |
| ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ .....        | vi  |
| İÇİNDEKİLER .....                                        | vii |
| TABLolar DİZİNİ .....                                    | ix  |
| ŞEKİLLER DİZİNİ .....                                    | x   |
| SİMGELER VE KISALTMALAR DİZİNİ .....                     | xi  |
| 1. GİRİŞ .....                                           | 1   |
| 1.1. Çalışmanın Konusu ve Önemi.....                     | 4   |
| 1.2. Çalışmanın Organizasyonu .....                      | 6   |
| 2. KAYNAK TARAMASI .....                                 | 7   |
| 2.1. ETA Tahmini.....                                    | 9   |
| 2.2. Yörünge Tahmini .....                               | 11  |
| 2.3. Çalışmanın Özgünlüğü .....                          | 13  |
| 3. PROBLEMİN TANIMI.....                                 | 15  |
| 3.1. Hava Trafik Sistemi .....                           | 15  |
| 3.2. Yeşilköy TMA Yapısı ve Kabuller.....                | 18  |
| 3.3. Problemin Tanımı .....                              | 20  |
| 3.4. Veri Kümesi .....                                   | 20  |
| 4. METODOLOJİ .....                                      | 25  |
| 4.1. Verilerin Veri Tabanına Aktarılması.....            | 26  |
| 4.2. Veri Analizi, Görselleştirme ve Veri Temizliği..... | 28  |
| 4.3. Öznitelik Mühendisliği.....                         | 29  |
| 4.4. Makine Öğrenmesi Yöntemleri.....                    | 31  |
| 4.4.1. Karar ağaçları .....                              | 32  |
| 4.4.2. Budama .....                                      | 35  |

|                                 |    |
|---------------------------------|----|
| 4.4.3. Torbalama .....          | 36 |
| 4.4.4. Rassal orman .....       | 37 |
| 4.4.5. Yapay sinir ağıları..... | 38 |
| 4.4.6. Derin öğrenme.....       | 41 |
| 5. SONUÇLAR .....               | 43 |
| 5.1. ETA Tahmini .....          | 43 |
| 5.2. Yörünge Tahmini .....      | 47 |
| 6. TARTIŞMA .....               | 50 |
| KAYNAKÇA .....                  | 52 |
| ÖZGEÇMİŞ .....                  | 57 |

## TABLÖLAR DİZİNİ

**Tablo 2.1.** Havacılıkta yapay zekâ uygulamalarının amaçları ve faydaları

**Tablo 3.1.** Yeşilköy TMA trafik istatistikleri

**Tablo 3.2.** ETA ve yörünge tahmininde kullanılan ortak veriler

**Tablo 3.3.** Allft+ nokta ve sektör profili

**Tablo 4.1.** Tarih-saat dönüşümü örneği

**Tablo 4.2.** METAR dönüşümü örneği

**Tablo 5.1.** Öznitelikler

**Tablo 5.2.** Kalkış sonrası ETA tahmini (LTBA)

**Tablo 5.3.** TMA girişi sonrası ETA tahmini (LTBA)

**Tablo 5.4.** Kalkış sonrası ETA tahmini (İSG)

**Tablo 5.5.** Öznitelikler

**Tablo 5.6.** Yörünge tahmini için performans ölçütleri

## ŞEKİLLER DİZİNİ

Şekil 1.1. Havacılık/ATM yapay zekâ fırsatları

Şekil 3.1. Yeşilköy TMA

Şekil 3.2. 05 pisti için 2018 yılı standart geliş rotaları

Şekil 3.3. FR24 “.kml” uzantılı dosya yapısı

Şekil 4.1. Metodoloji

Şekil 4.2. Veri tabanı yapısı

Şekil 4.3. Ham verilerin veri tabanına aktarılması

Şekil 4.4. Geliştirilen web uygulamasının yapısı

Şekil 4.5. Verilerin uygulamada görüntülenmesi

Şekil 4.6. Yörüngelerin uygulamada görüntülenmesi

Şekil 4.7. Karar ağacı yapısı

Şekil 4.8. Rassal orman

Şekil 4.9. Nöron yapısı

Şekil 4.10. ANN nöron yapısı

Şekil 4.11. Bir nöron tarafından yapılan hesaplamalar

Şekil 4.12. ANN yapısı

Şekil 4.13. Derin yapay sinir ağı yapısı

Şekil 5.1. RF hata histogramları (a) Kalkıştan sonra (b) TMA’ya girişten sonra

Şekil 5.2. DNN hata histogramları (a) Kalkıştan sonra (b) TMA’ya girişten sonra

Şekil 5.3. ETA tahmin ekranı

Şekil 5.4. Bir uçuş için elde edilen yörünge tahmini

## SİMGELER VE KISALTMALAR DİZİNİ

|       |                                                                                      |
|-------|--------------------------------------------------------------------------------------|
| ABD   | : Amerika Birleşik Devletleri                                                        |
| ACC   | : Area Control Centre (Saha Kontrol Merkezi)                                         |
| ADS-B | : Automatic Dependent Surveillance – Broadcast (Otomatik Bağımlı Gözetim-Yayın)      |
| AFTN  | : Aeronautical Fix Telecommunications Network (Sabit Havacılık Telekomünikasyon Ağı) |
| AI    | : Artificial Intelligence (Yapay Zekâ)                                               |
| AMAN  | : Arrival Management (Varış Yönetimi)                                                |
| ANN   | : Artificial Neural Networks (Yapay Sinir Ağları)                                    |
| APP   | : Approach Control Service (Yaklaşma Kontrol Hizmeti)                                |
| ARIMA | : Autoregressive Integrated Moving Average (Oto regresif Hareketli Ortalamalar)      |
| ATA   | : Actual Time of Arrival (Gerçek Varış Zamanı)                                       |
| ATCo  | : Air Traffic Controller (Hava Trafik Kontrolörü)                                    |
| ATD   | : Actual Time of Departure (Gerçek Kalkış Zamanı)                                    |
| ATIS  | : Automatic Terminal Information Service (Otomatik Terminal Bilgi Hizmeti)           |
| ATM   | : Air Traffic Management (Hava Trafik Yönetimi)                                      |
| BADA  | : Base of Aircraft Data (Uçak Veritabanı)                                            |
| BP    | : Back propagation (Geri yayılım)                                                    |
| CART  | : Classification and Regression Trees (Sınıflandırma ve Regresyon Ağaçları)          |
| CAS   | : Calibrated Airspeed (Kalibre Edilmiş Hava Sürati)                                  |
| CDM   | : Collaborative Decision Making (Ortaklaşa Karar Verme)                              |
| CNN   | : Convolutional Neural Network (Evrişimli Sinir Ağları)                              |
| CPF   | : Correlated Position Reports for a Flight (Uçuşun Bağlantılı Konum Raporları)       |
| CTFM  | : Current Tactical Flight Model (Geçerli Taktik Uçuş Modeli)                         |

|             |                                                                                                       |
|-------------|-------------------------------------------------------------------------------------------------------|
| CTR         | : Control Zone (Kontrol Bölgesi)                                                                      |
| DART        | : Data-driven Aircraft Trajectory Prediction Research (Veri GÜdümlü Uçak Yörünge Tahmini Araştırması) |
| DBSCAN      | : Density-Based Spatial Clustering of Applications with Noise                                         |
| DCI         | : Dynamic Cost Index (Dinamik Maliyet Endeksi)                                                        |
| DDR2        | : Demand Data Repository (Talep Veri Deposu)                                                          |
| DMAN        | : Departure Management (Kalkış Yönetimi)                                                              |
| DNN         | : Deep Neural Network (Derin Sinir Ağı)                                                               |
| EAAI HLG    | : European Aviation High Level Group on AI (Yapay Zekâ Alanında Avrupa Havacılık Yüksek Seviye Grubu) |
| EC          | : European Commission (Avrupa Komisyonu)                                                              |
| ETA         | : Tahmini Varış Zamanı (Estimated Time of Arrival)                                                    |
| ETMS        | : Enhanced Traffic Management System (Gelişmiş Trafik Yönetim Sistemi)                                |
| EU          | : European Union (Avrupa Birliği)                                                                     |
| Eurocontrol | : European Organisation for the Safety of Air Navigation (Avrupa Hava Seyrüsefer Emniyeti Teşkilatı)  |
| FAA         | : Federal Aviation Agency (ABD Federal Havacılık Kurumu)                                              |
| FIC         | : Flight Information Centre (Uçuş Bilgi Merkezi)                                                      |
| FIR         | : Flight Information Region (Uçuş Bilgi Bölgesi)                                                      |
| FTFM        | : Filed Tactical Flight Model (Doldurulmuş Taktik Uçuş Modeli)                                        |
| GBM         | : Gradient Boosting Machines (Gradyan Arttırma Makineleri)                                            |
| GLM         | : Generalized Linear Models (Genelleştirilmiş Doğrusal Modeller)                                      |
| GPS         | : Global Positioning System (Küresel Konumlama Sistemi)                                               |
| HF          | : High Frequency (Yüksek Frekans)                                                                     |
| HMM         | : Hidden Markov Models (Saklı Markov Modelleri)                                                       |
| HSV         | : Hue-Saturation-Value (Renk Doygunluk Değeri)                                                        |
| IEM         | : IOWA State University Iowa Environmental Mesonet (IOWA Devlet Üniversitesi Iowa Çevresel Veri Ağı)  |

|       |                                                                                                       |
|-------|-------------------------------------------------------------------------------------------------------|
| ISG   | : Istanbul Sabiha Gokcen International Airport (İstanbul Sabiha Gökçen Uluslararası Havalimanı)       |
| KML   | : Keyhole Markup Language (Anahtar Deliği Biçimlendirme Dili)                                         |
| LSTM  | : Long Short-Term Memory (Uzun Kısa-Süreli Bellek)                                                    |
| LTBA  | : Istanbul Ataturk International Airport (İstanbul Atatürk Uluslararası Havalimanı)                   |
| MAE   | : Mean Absolute Error (Mutlak Ortalama Hata)                                                          |
| MCNN  | : Multi-cell Neural Networks (Çok Hücreli Sinir Ağları)                                               |
| METAR | : Meteorological Aerodrome Report (Havacılık Amaçlı Rutin Hava Raporu)                                |
| ML    | : Machine Learning (Makine Öğrenmesi)                                                                 |
| MLAT  | : Multilateration (Hava Aracı Konum Belirleme)                                                        |
| MSE   | : Mean Squared error (Ortalama Kare Hata)                                                             |
| NOAA  | : National Oceanic and Atmospheric Administration (Ulusal Okyanus ve Atmosfer Dairesi)                |
| OGC   | : Open Geospatial Consortium (Açık Coğrafi Bilgi Konsorsiyumu)                                        |
| OOB   | : Out-of-Bag (Torba Dışı)                                                                             |
| PEK   | : Beijing Capital International Airport (Pekin Başkent Uluslararası Havalimanı)                       |
| QRF   | : Quantile Regression Forests (Quantile Regresyon Ormanı)                                             |
| ReLU  | : Rectified Linear Unit (Doğrultulmuş Doğrusal Birim)                                                 |
| RF    | : Random Forest (Rassal Orman)                                                                        |
| RGB   | : Red-Green-Blue (Kırmızı Yeşil Mavi)                                                                 |
| RMSE  | : Root Mean Square Error (Kök Ortalama Kare Hata)                                                     |
| RNN   | : Recurrent Neural Network (Tekrarlayan Sinir Ağı)                                                    |
| RTFM  | : Regulated Tactical Flight Model (Düzenlenmiş Taktik Uçuş Modeli)                                    |
| SESAR | : Single European Sky ATM Research (Tek Avrupa Hava Sahası Hava Trafik Yönetimi Araştırma Geliştirme) |
| SPO   | : Single Pilot Operations (Tek Pilotlu Operasyonlar)                                                  |
| STA   | : Scheduled Time of Arrival (Planlanan Varış Zamanı)                                                  |

|      |                                                                                   |
|------|-----------------------------------------------------------------------------------|
| STAR | : Standard Arrival Route (Standart Geliş Rotası)                                  |
| STD  | : Scheduled Time of Departure (Planlanan Kalkış Zamanı)                           |
| SVR  | : Support Vector Regression (Destek Vektör Regresyonu)                            |
| SWIM | : System-Wide Information Management (Sistem Geneline Bilgi Paylaşımı)            |
| TAM  | : Total Airport Management (Toplam Havaalanı Yönetimi)                            |
| TAS  | : True Airspeed (Hakiki Hava Sürati)                                              |
| TBO  | : Trajectory Based Operations (Yörünge Temelli Operasyonlar)                      |
| TMA  | : Terminal Maneuvring Area (Terminal Kontrol Sahası)                              |
| TWR  | : Aerodrome Control Service (Meydan Kontrol Hizmeti)                              |
| VHF  | : Very High Frequency (Çok Yüksek Frekans)                                        |
| XML  | : Extensible Markup Language (Genişletilebilir İşaretleme Dili)                   |
| ZGSZ | : Shenzhen Bao'an International Airport (Shenzhen Bao'an Uluslararası Havalimanı) |

## 1. GİRİŞ

Son on yıllık dönem içinde (2010-2019), dünya ücretli-tarifeli yolcu trafiği yıllık ortalama %6,1 artış ile 2019 yılında 4,5 milyarı geçmiştir (Devlet Hava Meydanları İşletmesi [DHMİ], 2020). Ayrıca, küresel boyutta gerçekleşen ticari uçuşların 2010-2018 döneminde %27 artış ile 30 milyondan 38 milyona yükseldiği görülmüştür (<http-1>). Avrupa’da ise 2019 yılında toplam uçuş sayısı 11,1 milyon ile rekor kırmıştır. 2018 yılında yapılan tahminlere göre, Avrupa hava trafiğinin 2035 yılından itibaren, yıllık 15 milyon uçuşun üzerine çıkacağı öngörülmüş (Single European Sky ATM Research [SESAR], 2020), ciddi önlemler alınmadığı takdirde, 2040 yılında havaalanlarının yaklaşık 1,5 milyon uçuşu karşılayamayacağı ve 160 milyon yolcunun uçamayacağı ifade edilmiştir (European Organisation for the Safety of Air Navigation [Eurocontrol], 2018). Buna rağmen yol ve havaalanı ağının son derece yoğun olacağı ve günün çoğunda yaklaşık tam kapasiteyle çalışan büyük havaalanlarının sayısının 6’dan 16’ya çıkacağı belirtilmiştir. Bu durumda 2016 yılında uçuş başına 12 dk. olan gecikmelerin 2040 yılında 20 dk. civarında olacağı ve günümüzde her gün 1-2 saat gecikme yaşayan 50 bin yolcunun, 2040 yılında 470 bine ulaşacağı vurgulanmıştır. İlgili dönemlerde yapılan bu tahminlere göre hava trafik sisteminin en öncelikli sorununun mevcut kapasitenin geliştirilmesi olduğu ifade edilmiştir.

Ancak 2019 yılının sonlarında ortaya çıkan yeni tip Koronavirüs (Covid-19) havacılık sektörünü ciddi biçimde etkilemiştir. 2020 yılı için tahminler incelendiğinde, yolcu sayısında 2019 yılına kıyasla %60’a yakın bir azalma olacağı beklenmektedir (International Civil Aviation Organization [ICAO], 2020). Pandemi sonrası için birçok farklı senaryo geliştirilmekle birlikte, havayolu şirketlerinin son verilerine ve beklentilerine göre, referans senaryoda pandemi sonrası sektörde hızlı bir iyileşme olacağı ve hava trafik hacminin 2019 değerlerini 2024 yılında tekrar yakalayacağı öngörülmektedir (Eurocontrol, 2020). Hava trafiğindeki son 50 yıllık artışa bakıldığında çeşitli küresel siyasal, toplumsal ve ekonomik krizlere rağmen sektörün kendini kısa sürede topladığı ve eski seviyelerine geri döndüğü bilinmektedir (<http-2>). Dolayısıyla, 2024 sonrasında kapasite problemlerinin yine önemli bir sorun olacağı ve halihazırda kapasite problemleri yaşayan havaalanlarının, artan talebi karşılayabilmeleri için gerek altyapılarında gerekse operasyonel anlamda iyileştirmeler yapmaları gerektiği açıktır. Artan hava trafiğiyle başa çıkabilecek, ölçeklenebilir ve modern

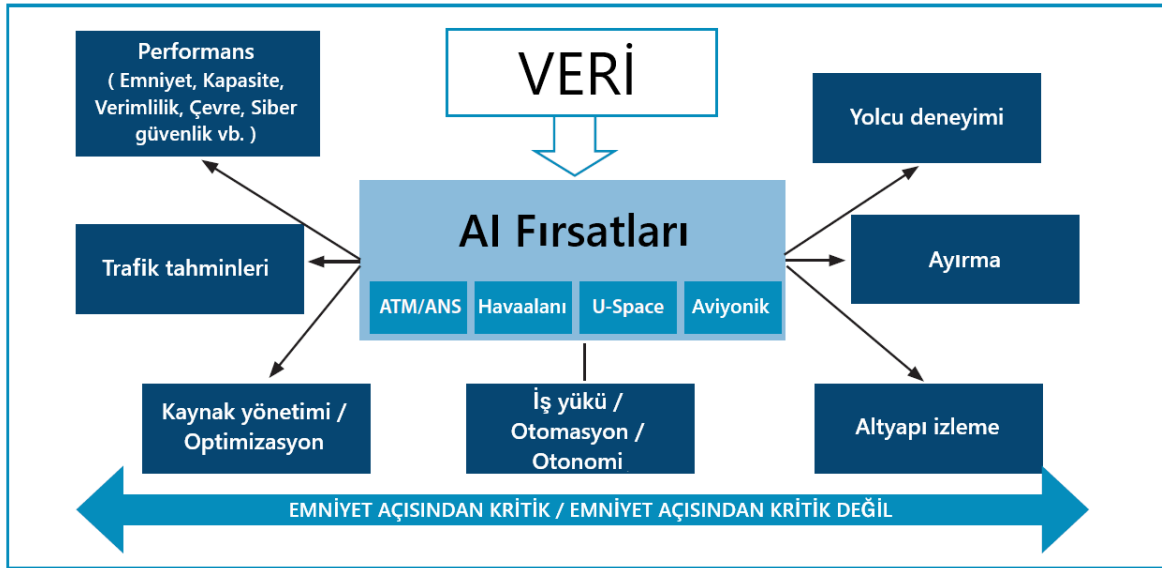
bir Hava Trafik Yönetimi (ATM) sisteminin geliştirilmesi vizyonuyla çalışmalar yürüten Tek Avrupa Hava Sahası ATM Araştırma Geliştirme (SESAR) projesi, uçuşun bir bütün olarak ele alınmasını ve her aşamasında iyileştirmeler yapılması gerektiğini ifade etmektedir (SESAR, 2020). Bu iyileştirmelerden ve geliştirilen konseptlerden bazıları şunlardır:

- Gelişmiş uçuş yörüngeleri (Yörünge Temelli Operasyonlar, TBO)
- Gelişmiş havaalanı performansı (Toplam Havaalanı Yönetimi, TAM)
- Yüksek seviye otomasyon (Tek Pilotlu Operasyonlar, SPO)
- Gelişmiş seyrüsefer hizmetleri (sesli iletişimden veri iletişimine geçiş, yer sistemleri arasında veri paylaşımının artırılması vb.)
- Hava seyrüsefer hizmetleri altyapısının ve kısıtlı kaynakların optimal kullanımı
- Gelişmiş küresel iş birliği (Sistem Genelinde Bilgi Paylaşımı, SWIM)
- Gelişmiş emniyet ve güvenlik

Bu vizyonun gerçekleştirilmesinde dijital dönüşümün önemli bir rol oynadığı ifade edilmekte; sistem genelinde veri paylaşımı ve birbirleriyle bağlantılı otomasyon sistemleri aracılığıyla emniyet, verimlilik ve maliyet açısından sistem bütününde performansın iyileştirileceği vurgulanmaktadır. Eğilimlerin belirlenebilmesi için veri madenciliği ve büyük veri yönetimi (gerçek zamanlı bilgi sağlama) becerilerine sahip karmaşık algoritmaların ve tahmin modellerinin, bütün paydaşlar arasında verimli ve etkili ortaklaşa karar verme sürecini mümkün kılacağı ifade edilmektedir. Ayrıca, makine öğrenmesi ve yapay zekâ (AI) alanındaki gelişmelerin hava trafik yönetiminde (ATM) birçok yenilikçi uygulamanın kapılarını aralayacağı tahmin edilmektedir. AI, Avrupa Birliği (EU) tarafından da önemli bir stratejik öncelik olarak görülmektedir. Bu nedenle, AI potansiyelinin havacılık ve ATM aktörleri tarafından anlaşılmasını sağlamak ve sektörde kullanımını hızlandırmak amacıyla Avrupa Hava Seyrüsefer Emniyeti Teşkilatı (Eurocontrol), Avrupa Komisyonu (EC) ve birçok organizasyonun iş birliğiyle European Aviation High Level Group on AI (EAAI HLG) kurulmuştur (EAAI HLG, 2020).

AI sistemleri, veri toplama ve verileri yorumlama yoluyla çevrelerini algılayan, bu verilerden türetilen bilgiyi işleyerek verilen amacı gerçekleştirebilmek için en iyi aksiyona karar veren ve harekete geçen akıllı yazılım ve/veya donanım sistemleridir (EAAI HLG, 2020). Yapay zekânın bir alt kümesi olan makine öğrenmesi (ML), algoritmaların giriş ve

çıkış verilerinden sistemlerin çalışma prensiplerini öğrenebilme ve modelleyebilme yeteneğine karşılık gelmektedir. ML, insanların veya geleneksel bilgisayar destekli sistemlerin belirlemekte zorlandıkları, karmaşık verilerdeki şablonları belirleyebilme yeteneğine sahiptir. Bu durum makine öğrenmesini, hava ve yer sistemleri arasındaki veri akışına giderek daha bağımlı hale gelen havacılık sektörüne son derece uygun hale getirmektedir. Şekil 1.1 havacılık ve ATM sistemlerinde AI/ML tarafından katkı sağlanan fırsatları göstermektedir. Bu fırsatlar havacılığın geleceği için önem arz eden bazı konuları ön plana çıkarmaktadır: kapasite problemlerinin çözümü, yüksek seviye karmaşıklıkların hafifletilmesi, çevresel etkilerin azaltılması ve dijital dönüşümün gerçekleştirilmesi.



Şekil 1.1. Havacılık/ATM yapay zekâ fırsatları (EAAI HLG, 2020)

Kapasite problemlerinin çözümünde; verilerin daha faydalı bir şekilde kullanımı sağlanarak, daha doğru tahminlerin ve daha gelişmiş araçların elde edilmesi, hava sahası ve havaalanlarının daha etkin ve verimli kullanılması hedeflenmektedir. Yüksek seviye karmaşıklıkların hafifletilmesinde, planlamadan operasyonlara bütün ATM aktörlerine fayda sağlayacağı, yeni çakışma tespiti, trafik öneri sistemleri ve çözüm araçları ile emniyeti arttıracacağı ifade edilmektedir. Optimal yörüngelerin elde edilmesi, yeşil rotaların üretilmesi ve tahmin doğruluğunun artırılması ile çevresel etkilerin hafifletilmesinde de rol oynayacağı düşünülmektedir. Dijital dönüşüm sonucunda, arttırılmış otomasyon seviyesi, güvenli veri

bağlantısı ve paylaşımı, altyapının ve hava trafik hizmetlerinin görselleştirilmesi ile sistemin daha modüler ve daha çevik hale geleceği öngörülmektedir (EAAI HLG, 2020).

### **1.1. Çalışmanın Konusu ve Önemi**

Terminal kontrol sahaları (TMA) kapasite problemlerinden en olumsuz etkilenen hava trafik sistemi bileşenlerinden biridir. İniş ve kalkış yapan uçakların toplandıkları ve yeniden dağıldıkları TMA'lar, uçuş yoğunluğu ve operasyon çeşitliliği nedeniyle hava trafiğinin en karmaşık olduğu yerlerdir. Dolayısıyla, artan talep karşısında en büyük darboğaz TMA'larda yaşanmaktadır (Cetek, 2015).

Hava sahası kapasitesi ve verimliliğinin artırılmasındaki önemli faktörlerden birisi tahmini varış zamanının (ETA) yüksek doğrulukla tahmin edilebilmesidir (Ayhan, Costas and Samet, 2018). Varış zamanının doğru tahmini, Ortaklaşa Karar Verme (CDM) (International Air Transport Association [IATA], Eurocontrol and Civil Air Navigation Services Organization [CANSO], 2008) ve gecikme yönetiminin önemli bileşenlerinden biridir. CDM, hava trafik yönetiminin dört ana alanında (emniyet, kapasite, verim ve çevresel etki) gelişme sağlayarak maliyetlerin kontrol altında tutulabilmesini sağlar. Gecikme yönetimi ise, havayolu firmalarının gecikmeyi telafi etmek için yakıt tüketimini artırmaları veya yakıt tasarrufu yaparak gecikme maliyetlerini kabullenmeleri konusunda karar vermelerini sağlayan dinamik maliyet endeksi (DCI) (Cook, Tanner and Lawes, 2007) kavramının arkasındaki yaklaşımdır.

İyileştirilmiş ETA tahmini, uçuş gecikmelerinin etkilerini hafifleteceği gibi, havayolu firmalarının yer hizmetleri personel ve ekipman faaliyetlerinin koordinasyonunu sağlamalarında, zaman ve enerji kayıplarının önüne geçmede faydalıdır. Benzer şekilde hava trafik akış yönetimi, pist ve kapı atamaları ve yer destek elemanlarının optimizasyonu da daha doğru ETA tahmininden faydalanabilecek süreçlerdir (Kern, Medeiros and Yoneyama, 2015).

ETA tahmininin yanı sıra, ATM süreçlerinin birçoğunun performansı, yörüngenin de doğru tahminine dayalıdır. Kalkış ve varış yönetimi (DMAN/AMAN), çatışma tespiti, uçuş planlama ve akış yönetimi bu süreçlere örnek olarak verilebilir (Leege, Paassen and Mulder, 2013).

Yörünge temelde bir dizi durum değişkeni ile ifade edilen uçak durumlarının kronolojik olarak sıralı bir listesidir (Fernandez vd., 2017). Bu durum değişkenlerinin en önemlileri hız (Hakiki Hava Sürati (TAS), Kalibre Edilmiş Hava Sürati (CAS) veya Mach), 3 boyutlu konum bilgisi (enlem, boylam ve irtifa), baş açısı ve uçağın anlık kütlesidir. 4 boyutlu yörünge tahmini, var olan verilere bağlı olarak gelecekte uçağın bulunacağı enlem, boylam, irtifa ve zaman bilgilerinin tahmini olarak tanımlanabilir

Eurocontrol tarafından sağlanan BADA (Base of Aircraft Data) modelleri gibi son derece doğru uçak performans modelleri ve NOAA (National Oceanic and Atmospheric Administration) tarafından sağlanan hava tahminleri mevcut olmasına rağmen gerçek ve tahmini yörüngeler arasında önüne geçilemeyen hatalar bulunmaktadır. Bu hatalar stokastik bir sürecin deterministik bir yaklaşımla ifade edilmesinin sonucudur.

Hava trafik kontrolörlerinin (ATCo) amaçları, uçakların emniyetli ve verimli biçimde yönetimidir. Bu amaçlar havadaki uçakların ayrımını ve sıralanmasını gerektirir. Kontrolörler her uçağın durumunu izlemek ve gerektiğinde çatışmaları çözmek için karar destek sistemlerini kullanırlar. Ayrıca kontrolörler, trafik yoğunluğunun emniyetli sınırlar içerisinde kalmasını sağlamak için de karar destek sistemlerini kullanarak akış yönetimi ile ilgili kararlar verirler. Bu sistemlerin merkezinde ise yörünge tahmin araçları yer alır. Bu araçlar uçağın başlangıç durumu, uçuş planı, rüzgâr ve sıcaklık verileri gibi birçok veriyi kullanarak uçağın gelecekteki 4 boyutlu yörüngesini hesaplar. Yanlış yörünge tahmininin ATM sistemi üzerinde önemli etkileri bulunmaktadır (Ayhan and Samet, 2016). Bunların bazıları aşağıdaki gibidir:

- Daha yüksek ayırım değerleri nedeniyle sektördeki uçak sayısını kısıtlar.
- Çatışma sayısındaki artışa bağlı olarak kontrolörün çalışma yükünü artırır ve sektör kapasitesini düşürür.
- Verimsiz yakıt tüketimine neden olur.
- Artan emisyonlar nedeniyle çevreye etkisi artar.

Bu problemlerin üstesinden gelmek amacıyla ATM sisteminde, TBO adı verilen ve yörünge verilerinin ATM sisteminin dayandığı temel bilgi olduğu yeni bir konseptte geçiş yapılması planlanmaktadır (SESAR, 2012). Ancak bu dönüşüm, operasyon öncesi aşamada yüksek doğruluğa sahip yörünge tahmini becerilerine ihtiyaç duymaktadır.

Bu alıřmanın konusu, uakların ETA'larının ve terminal kontrol sahaları ierisinde izleyecekleri yrngelerin, ML yntemleri kullanarak tahmin edilmesi ve elde edilen ML modelleri aracılıėıyla bir karar destek sistemi nerisi sunulmasıdır.

## **1.2. alıřmanın Organizasyonu**

İkinci blmde ML tanımı yapılarak havacılık sektrnde makine ėrenmesinin kullanıldıėı bazı alanlardan bahsedilmiř ve ML ile ETA tahmini ve yrnge tahmini literatr incelenmiřtir. nc blmde kısaca hava trafik sistemi aıklanmıř, problemin tanımı yapılarak veri kaynakları ve elde edilen veri kmesinin yapısından ve boyutundan bahsedilmiřtir. Drdnc blmde problemin zmnde izlenen metodolojiden bahsedilmiř, kullanılan ML yntemleri teorik olarak anlatılmıřtır. Beřinci blmde ETA tahmini ve yrnge tahmininde kullanılacak olan ML modelleri eėitilmiř, test edilmiř ve performansları llmřtir. Son blm ise tartıřma blmdr.

## 2. KAYNAK TARAMASI

Verilerdeki örüntülerin belirlenmesi problemi çok uzun yıllardır çalışılan önemli bir problemidir. Örüntü tanıma, verilerdeki düzenin bilgisayar algoritmalarıyla otomatik olarak ortaya çıkarılması ve bu örüntülerin sınıflandırma gibi amaçlar doğrultusunda kullanılması ile ilgilenir (Bishop, 2006). Örneğin el yazısı karakter tanıma problemi el yazısındaki büyük değişkenlikler nedeniyle çözülmesi zor bir problemidir. Bu tarz problemlerin çözümünde makine öğrenmesi yaklaşımının kullanılması başarılı sonuçlar vermektedir.

ML şu şekilde tanımlanabilir:  $T$  bir dizi görevi,  $P$  bu görevlerdeki performans ölçütünü ve  $E$  deneyimi ifade etmek üzere, bir bilgisayar programının  $T$  görevlerindeki  $P$  performansı, artan  $E$  deneyimi ile iyileşiyorsa bu program *deneyimle öğreniyor* denir (Mitchell, 1997). Örneğin satranç oynamayı öğrenen bir bilgisayar programı, kendine karşı oyunlar oynayarak elde ettiği deneyimle, kazanma becerisiyle ölçülen performansını iyileştirilebilir. Burada  $T$ ,  $P$  ve  $E$  şu şekilde tanımlanabilir:

T: satranç oynamak

P: Rakipler karşısında kazanılan oyunların yüzdesi

E: Kendine karşı oynadığı eğitim amaçlı oyunlar

ML yönteminde, eğitim kümesi adı verilen büyük bir veri kümesi, modelin parametrelerinin ayarlanmasında kullanılır. Bu aşama eğitim veya öğrenme aşaması olarak adlandırılır. Model eğitildikten sonra test kümesi adı verilen yeni girdiler kullanılarak modelin performansı ölçülür.

ML algoritmaları geniş anlamda denetimli öğrenme, denetimsiz öğrenme, takviyeli öğrenme ve derin öğrenme olarak sınıflandırılabilir. Veri kümesi girdi vektörleriyle birlikte bu girdilere karşılık gelen çıktıları da (etiketleri) içeriyorsa bu bir denetimli öğrenme problemidir. Her bir girdi vektörünün sonlu sayıda ayrık kategorilere atanmasının amaçlandığı problemlere sınıflandırma problemleri adı verilir. Eğer istenilen çıktı bir veya daha fazla sürekli değişkenden meydana geliyorsa bu problemlere regresyon problemleri adı verilir.

Bazı örüntü tanıma problemlerinde eğitim verileri, girdilere karşılık gelen etiket değerlerini içermemektedir. Denetimsiz öğrenme adı verilen bu problemlerde amaç, veri içerisinde birbirine benzeyen örnek gruplarını keşfetmektir. Buna kümeleme adı verilir.

Takviyeli öğrenmede model sadece bir veri kümesinden öğrenmekle kalmaz, aynı anda ortamla da etkileşime girerek deneyimiyle ilgili bir geribildirim alır (Ghatak, 2017). Bu geribildirim, ödül adı verilen ve gerçekleştirilen eylemin ne kadar iyi olduğunu ifade eden sayısal bir değerdir. Takviyeli öğrenmenin amacı bu ödül değerini maksimize eden eylemler dizisinin deneme yanılma yoluyla belirlenmesidir (Sutton and Barto, 2015). Makine öğrenmesinin spesifik bir alt alanı olan derin öğrenmede, verilerin ardışık katmanlarda giderek daha anlamlı hale gelmesi hedeflenmektedir. Modele katkı sağlayan katmanların sayısı modelin derinliği olarak tanımlanır (Chollet, 2018).

Çoğu uygulamada, problemin çözümünü kolaylaştırmak veya hesaplamayı hızlandırmak için orijinal girdi değişkenleri bir önışlemden geçirilerek bir dizi yeni değişkene dönüştürülür. Bu önışlem aşaması öznitelik çıkarma olarak da adlandırılır. Test verileri de eğitim verileriyle aynı şekilde önışleme tutulmalıdır.

Havacılıkta en sık karşılaşılan AI yöntemi makine öğrenmesidir (EAAI HLG, 2020). Havacılık paydaşları, ATM, havaalanı sistemleri ve aviyonik alanlarında gerçek dünya performansını iyileştirmek için birçok uygulama geliştirmektedir. Tablo 2.1 geliştirilen AI uygulamalarının ana hedeflerini ve faydalarını özetlemektedir.

**Tablo 2.1.** Havacılıkta yapay zekâ uygulamalarının amaçları ve faydaları

| Çalışma Alanı                | Fayda                                                                                                                        |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------|
| Trafik tahmini/modelleme     | 1. Uçak yörünge tahminlerinin iyileştirilmesi, belirsizliğin azaltılması ve kapasitenin artırılması                          |
| Kaynak yönetimi/Optimizasyon | 1. Sektör konfigürasyonlarının ve kapasitenin optimizasyonu<br>2. ATM talep ve kapasite dengesinin desteklenmesi             |
| İş yükü/Otomasyon/Otonomi    | 1. ATCo iş yükünün azaltılması (örneğin ses tanıma)<br>2. Akıllı emniyet araçlarıyla risklerin azaltılması                   |
| Havaalanı performansı        | 1. Pist kapasitesinin artırılması<br>2. Havaalanı gecikmelerinin azaltılması                                                 |
| Yolcu deneyimi               | 1. Yolcu transferinin ve müşteri memnuniyetinin iyileştirilmesi<br>2. Emniyetli binişi hızlandırmak için biyometri kullanımı |
| Uçuş kabiliyetleri           | 1. Otomatik ses tanıma ile pilot ve ATCo asistanlığı                                                                         |

|                             |                                                                                |
|-----------------------------|--------------------------------------------------------------------------------|
|                             | 2. Bilgisayarlı görü aracılığıyla otomatik taksi, kalkış ve iniş operasyonları |
| Havayolu şirket performansı | 1. Yakıt tüketiminin optimizasyonu<br>2. Rota önerme                           |

ETA tahmini ve yörünge tahmini ile ilgili kaynaklar iki ayrı başlıkta incelenmiştir.

## 2.1. ETA Tahmini

ETA tahmininde geleneksel yöntemlerde genellikle uçak performans modellerine ve parametrik veya fizik tabanlı yörünge modellerine bağlı, deterministik bir yaklaşım kullanılır. Bu yöntemlerde ilk olarak uçuş rotası ile irtifa ve hız profilleri tahmin edilerek, tahmini yörüngeye uygun zaman değerleri hesaplanır (Slattery and Zhao, 1997; Mondoloni, 2006; Murphy vd., 2003). Ancak varış zamanı sadece uçağın uçuş hızına ve irtifa profiline değil, rota boyunca uçuşa etki eden hava olayları, hava sahası sektör yoğunlukları ve hava alanı yoğunluğu gibi uçakların gerçek uçuş profillerini doğrudan etkileyen operasyonel koşullara da bağlıdır.

ML yöntemleri ile ETA tahmini için literatürde sınırlı sayıda kaynak bulunmaktadır. ABD Federal Havacılık Dairesi'nin (FAA) Gelişmiş Trafik Yönetim Sistemi (ETMS) tarafından üretilen ETA tahminlerinin iyileştirilmesi için Kern ve arkadaşları tarafından geliştirilen ML tabanlı yaklaşımda, modelin eğitimi için genel uçuş bilgileri, meteorolojik veriler ve hava trafiği verileri kullanılmıştır (Kern, Medeiros and Yoneyama, 2015). ML yöntemi olarak ağaç tabanlı bir yöntem olan Rassal Orman (RF) seçilmiştir. Bu yöntemin seçilmesinin sebepleri farklı yapıdaki (sürekli, kategorik vb.) öznitelikleri kullanabilme becerisi ve çok büyük veri kümeleri ve fazla sayıdaki özniteliklerle bile etkili biçimde çalışabilmesidir. Kullanılan veri kümesi ABD hava sahasındaki 24.787 ticari uçuştan meydana gelmektedir. Öznitelik değerleri uçuş süresince seyir aşamasında rastgele bir noktadan çıkarılmıştır. Öznitelik seçimi için izlenen yaklaşım önce az sayıda öznitelik ile başlayıp yenilerinin modele eklenmesi ve tahmin doğruluğunu arttıran özniteliklerin modelde tutulması şeklindedir. Eğitilen model, ETMS tarafından sağlanan ETA tahmininin mutlak ortalama hatasını (MAE) %42,7 oranında azaltmıştır.

Takacs çalışmasında, varış zamanı tahminlerinin iyileştirilerek uçuşların daha verimli hale gelmesi için General Electric tarafından düzenlenen Flight Quest yarışmasında üçüncülüğü kazandığı ML yönteminden bahsetmiştir (Takacs, 2014). Yarışmanın amacı ABD iç hatlar uçuşlarının pist ve kapı varış zamanlarının; uçuş geçmişi, hava durumu, hava trafiği ve diğer verilere bağlı olarak, olabildiğince doğru tahmin edilmesidir. Yarışmada sağlanan veri kümesi, 12.11.2012-28.02.2013 tarihleri arasındaki 109 günlük uçuş bilgilerini kapsayan ve 34 tabloda toplam 252 kolonda “csv” formatında tutulan 128 GB boyutlu ham veri şeklindedir. Takacs’ın yaklaşımı, ham veriden türetilmiş 56 adet öznitelik üzerinde oluşturulan, ardışık ridge regresyon ve gradyan arttırma makinelerinden (GBM) meydana gelen 6 aşamalı bir modeli içermektedir. Eğitim ve modelin koşturulması için 1 TB hafızaya sahip 64 çekirdekli bir bilgisayar kullanılmıştır. Takacs yarışmanın son aşamasında yaklaşık 81 saat toplam CPU zamanı çalışan uygulamasıyla 5,3926 kök ortalama kare hata (RMSE) değerine ulaşarak üçüncülüğü elde etmiştir.

Glina vd. ağaç tabanlı bir toplu öğrenme yöntemi olan Quantile Regression Forests (QRF) (Meinshausen, 2006) yöntemini kullanarak varış zamanının tahmini için yeni bir yaklaşımı önermişlerdir (Glina, Jordan and Ishutkina, 2012). Modellerini Dallas/Fort Worth Uluslararası Havalimanı’ndan elde ettikleri veriler üzerinde doğrulamış ve havaalanına 20 nm’den daha yakın uçakların pist varış zamanını 60 sn’den daha küçük MAE ile tahmin etmişlerdir. Ancak verileri sadece 5 günlük bir periyodu kapsamaktadır.

Ayhan vd. sadece varış havaalanı için değil potansiyel rota üzerindeki hava trafiği ve hava durumu parametrelerini de toplayarak daha zengin bir öznitelik kümesi elde etmişler, bir dizi regresyon modeli ve Tekrarlayan Sinir Ağı (RNN) yöntemi ile Eurocontrol tarafından yapılan ETA tahminlerinden daha başarılı tahminler elde etmişlerdir (Ayhan, Costas and Samet, 2018). Yaptıkları deneylerde geliştirdikleri sistemin İspanya’daki herhangi bir ticari uçuşun varış zamanını 4 dk’lık bir RMSE ile belirleyebildiklerini doğrulamışlardır.

Wang vd. TMA’ya giriş noktasında uçakların piste varış zamanlarının tahminini gerçekleştirebilmek için yeni, veri güdümlü otomatik bir sistem geliştirmişlerdir (Wang, Liang and Delahaye, 2020). Önerilen sistem temelde veri ön işleme ve makine öğrenmesi modellerinden meydana gelmektedir. Uçuşlar iniş yaptıkları pistlere göre ayrıldıktan sonra birçok farklı makine öğrenmesi yöntemi kullanılarak modeller eğitilmiş ve Beijing TMA’da gerçek Otomatik Bağımlı Gözetim-Yayın (ADS-B) verileri üzerinde testler

gerçekleştirilmiştir. Son olarak, bireysel modellerden daha karmaşık yapıda bir yığın modeli geliştirilmiş ve en iyi ETA tahminleri bu model kullanılarak gerçekleştirilmiştir.

Wang vd. zaman serileri ve RNN kullanarak ETA tahmini gerçekleştirmişlerdir (Wang, Li and Mao, 2020). Tahmin problemi seyir seyahat süresi tahmini ve TMA seyahat süresi tahmini şeklinde iki alt probleme bölünmüştür. Seyir seyahat süresi tahmini için bir zaman serileri tahmin modeli olan otoregresif hareketli ortalamalar (ARIMA) kullanılmıştır. TMA yolculuk süresinin tahmini için ise bir RNN türü olan çift yönlü Uzun Kısa-Süreli Bellek (LSTM) kullanılmıştır. TMA seyahat süresi tahmini, 120 km içerisinde 5 büyük havaalanı içeren Guangdong-Hong Kong-Macao Greater Bay Area'da gerçek 4 boyutlu yörünge verileri kullanılmıştır. Son olarak Beijing Capital Uluslararası Havalimanı'ndan (PEK) Shenzhen Bao'an Uluslararası Havalimanı'na (ZGSZ) yapılan uçuşların 2 aylık 4 boyutlu yörünge verileri kullanılarak, uzun dönemli ETA tahmini yapılmış ve 6.09 dk'lık bir MAE değeri elde edilmiştir.

## **2.2. Yörünge Tahmini**

Son yıllarda birçok alanda olduğu gibi ATM'de de ML çalışmalarında önemli bir artış yaşanmıştır. Ancak ML yöntemleri ile yörünge tahmini çalışmaları çok daha eskilere dayanmaktadır.

Leege vd. sabit varış rotalarını takip eden trafiğin sıralanması için ML tabanlı bir yörünge tahmini yaklaşımını önermişler, gerçek uçak yörüngelerini ve meteorolojik verileri kullanmışlardır (Leege, Paassen and Mulder, 2013). GLM (Generalized Linear Model), SVR (Support Vector Regression) ve Yapay Sinir Ağları (ANN) yöntemlerinin hepsini yörünge tahmini problemine uygulamışlar ancak implementasyon ve yorumlama kolaylığı nedeniyle sadece GLM'den bahsetmişlerdir. Son noktası pist olan 45 nm uzunluğundaki sabit varış rotasında geliştirdikleri model, varış zamanında gözlemlenen varyansın %63'ünü açıklayabilmiştir. Ortalama mutlak hata 18 sn olarak belirlenmiştir. Tahmin modeli kullanılarak pist kapasitesi saatte 4 uçak artırılmıştır.

Ayhan ve Samet çalışmalarında yörüngeleri, hava parametrelerini ve diğer verileri içeren 4 boyutlu bitişik küpler şeklinde modellemiş, zaman serileri kümeleme ve HMM (Hidden Markov Models) yöntemlerini kullanarak yörünge tahmini yapmışlardır (Ayhan and

Samet, 2016). Atlanta-Miami seferini yapan Delta Airlines uçuşunun (DAL2173) Mayıs 2010-Aralık 2011 dönemine ait 594 yörünge ve 56.752 nokta verisini kullanarak 12,6 km'lik yatay doğruluk değerini elde etmişlerdir.

Shi vd. LSTM bazlı bir yörünge tahmin modelini önermişler ve Haziran 2017-Kasım 2017 dönemini kapsayan 5 aylık süre için ADS-B yer istasyonlarından alınan verilerle Markov modellerine göre daha iyi sonuçlar elde ettiklerini ifade etmişlerdir (Shi vd., 2018).

Fernandez vd. DART (Data-driven Aircraft Trajectory Prediction Research) adını verdikleri SESAR 2020 projesiyle veri madenciliği ve makine öğrenmesi yöntemlerini kullanarak veri güdümlü yörünge tahmini yapmaya çalışmışlardır (Fernandez vd., 2017). HMM, kümeleme ve regresyon olmak üzere üç temel yaklaşımı ele almış ve tek bir yörünge tahmini için iki aşamalı (kümeleme/HMM) hibrit bir yöntem kullanmışlardır. Önerilen yöntem İspanya hava sahasında bir aylık bir periyodu kapsayan uçuşlardan elde edilen veri kümesi (yörünge ve NOAA hava durumu verileri) kullanılarak test edilmiş ve uçuş planlarında tanımlı noktalarda birkaç yüz metrelik tahmin doğruluğu elde edilmiştir.

Wang vd. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) ve Çok Hücreli Sinir Ağları (MCNN) yöntemlerini kullanarak terminal kontrol sahasında kısa dönem yörünge tahmini yapmışlar ve PEK TMA'daki Temmuz 2017 ADS-B verileri için (36.288 uçuş, 3.242.384 yörünge noktası) yaptıkları uygulamada çoğu yörünge için 100 sn'nin altında tahmin hatası elde etmişlerdir (Wang, Liang and Delahaye, 2017).

Basora vd. yörünge verilerini ve kümeleme yöntemlerini kullanarak hava trafiği akışını otomatik olarak belirleyebilmek için bir yaklaşım önermişler ve Fransız Saha Kontrol Merkezi (ACC) içerisindeki gerçek yörüngelere uygulamışlardır (Basora, Morio and Mailhot, 2017). 26 Haziran 2015 gününe ait Reims ACC Demand Data Repository (DDR2) trafik verilerini (9.442 yörünge) kullanarak yaptıkları uygulamada elde edilen kümelerin mevcut hava rotalarıyla uyumlu olduğunu ifade etmişlerdir.

Zeng vd. TMA'da yörünge tahmininin doğruluğunu arttırmak için yörünge verilerini kullanarak veri güdümlü bir yörünge tahmin modeli önermişlerdir (Zeng vd., 2020). Eşit zaman aralıklarına sahip, gürültüsüz ve yüksek kaliteli yörüngeler elde etmek için yörüngeleri yeniden oluşturabilecek bir düzeltme yöntemi geliştirmişlerdir. 4 boyutlu yörünge tahmini problemini, yörüngeler arasındaki yineleyen ve uzun/kısa süreli geçici bağımlılıkları yakalayabilecek sıradan sıraya LSTM problemi olarak ifade etmişlerdir.

Önerilen model Guangzhou TMA için elde edilen bir veri kümesine uygulanmış ve elde edilen sonuçların doğruluk anlamında birçok veri güdümlü yörünge tahmini yönteminden üstün olduğunu belirtmişlerdir.

Ma ve Tian, 4 boyutlu yörünge tahmini için Evrişimli Sinir Ağları (CNN) ve LSTM yöntemlerini birlikte kullanarak hibrit yapıya sahip, yeni bir derin öğrenme yaklaşımı önermişlerdir (Ma and Tian, 2020). Test için ADS-B verileri kullanılmış ve aynı veri kümesi üzerinde sadece LSTM ve BP modeli kullanılarak karşılaştırma yapılmıştır. Tahmin hatası LSTM modeline kıyasla %21,62, BP modeline kıyasla ise %52,45 azaltılmıştır.

### 2.3. Çalışmanın Özgünlüğü

Bu çalışmanın literatürdeki benzer çalışmalardan farkları şu şekilde sıralanabilir:

- 1) ETA tahmini çalışmaları genellikle belirli iniş/kalkış havaalanı çiftleri arasında veya iç hat uçuşlarında sınırlı bölgelerde yapılmaktadır. Bu çalışmada sadece iç hat uçuşları ele alınmamış, İstanbul Atatürk Havalimanı'na (LTBA) iniş yapan bütün tarifeli uçuşlar kalkış havaalanı ayrımı gözetilmeksizin modellerin eğitiminde kullanılmıştır. Dolayısıyla tahmin alanı diğer çalışmalara oranla çok daha geniştir.
- 2) ETA tahmini tek pist konfigürasyonu için değil, havaalanında kullanımda olan bütün pistler için gerçekleştirilmiştir.
- 3) Hem ETA tahmini hem de yörünge tahmini konusunda literatürdeki birçok çalışma veriler üzerinde ciddi ön işleme (koordinat dönüşümü, boyut azaltma vb.) gerektirmektedir. Ayrıca hatalı veya eksik olmasa dahi modelin doğruluğunu düşüren bazı verilerin (örneğin holding yapan uçaklar veya Standart Geliş Rotası'ndan (STAR) sapma gösteren uçuşlar) elendiği görülmektedir. Bu çalışmada veriler üzerinde basit dönüşümler yapılmış (tarih-zaman, mesafe vb.) ve holding yapan veya STAR'lardan sapma gösteren uçaklar eğitim verilerinden çıkarılmamıştır.
- 4) ETA tahmini çalışmalarında genellikle kalkış/iniş havaalanı çiftleri kategorik öznitelikler olarak kullanılmakta, dolayısıyla elde edilen modellerde aynı havaalanından kalkış yapan bütün uçaklar için kalkış havaalanının etkisi benzer olmaktadır. Bu çalışmada yörünge verilerinden uçuş mesafeleri hesaplanmış,

böylece aynı havaalanından kalkış yapan bütün uçaklar için kalkış havaalanının aynı etkiye sahip olmaması sağlanmıştır. Ayrıca bu durum aynı modelin benzer pist konfigürasyonlarına sahip farklı havaalanları için de herhangi bir düzeltme yapılmaksızın kullanılabilmesini sağlamaktadır.

- 5) ETA tahmini ve yörünge tahmini sonucunda elde edilen modeller, geliştirilen bir web uygulamasına entegre edilerek, dinamik olarak test edilebilmesi sağlanmıştır. Böylece model performansları sadece sayısal değerlerle ifade edilmeyip tahmin değerlerinin ve yörüngelerin görselleştirilmesi sağlanmıştır.
- 6) Literatürdeki çalışmalarda genellikle makine öğrenmesi modellerinin eğitiminde R ve Python programlama dilleri kullanılmaktadır. Bu çalışmada R programlama diline ek olarak Java programlama dili, h2o platformu ve DeepLearning4J kütüphanesi kullanılmıştır.

### **3. PROBLEMİN TANIMI**

Bu bölümde öncelikle hava trafik sistemi ve hava trafik hizmetleri gibi havacılık kavramlarından bahsedilecek, daha sonra çalışma kapsamında ETA ve yörünge tahminin yapıldığı Yeşilköy TMA ve LTBA ile ilgili bilgiler verilerek problemin tanımı yapılacaktır. Son olarak çalışmada kullanılacak veri kaynakları ve elde edilen veri kümesinden bahsedilecektir.

#### **3.1. Hava Trafik Sistemi**

Hava trafik sistemi, hava sahası, hava seyrüseferi teknik donanımı, hava araçları ve insan gücü bileşenlerinden meydana gelen ve havacılık faaliyetlerinin emniyetli, verimli ve ekonomik bir şekilde yürütülmesini amaçlayan bir sistemdir (Cetek, 2015).

Hava araçlarının uçuş operasyonlarını gerçekleştirdikleri yer olarak tanımlanan hava sahaları temelde kontrollü ve kontrolsüz hava sahaları olarak ikiye ayrılmaktadır. Kontrollü hava sahaları; içerisinde hava trafik kontrol hizmetlerinin verildiği, sınırları ve boyutları belirlenmiş hava sahalarıdır. Kontrolsüz hava sahalarında ise hava trafik hizmeti sağlanmamaktadır. Kontrollü hava sahaları, içerisinde gerçekleşen operasyonlara ve verilen hizmetlere göre havaalanı, kontrol bölgesi, terminal kontrol sahası, hava yolları ve uçuş bilgi bölgesi şeklinde sınıflandırılmaktadır (ICAO, 2018a).

Havaalanı, hava seyrüseferinin başlangıç ve bitiş noktasını oluşturan, hava araçlarının iniş, kalkış ve uçuşa hazırlık operasyonları için gerekli olan bina tesis ve donanımları barındıran sahalardır (ICAO, 2018b). Kontrol Bölgesi (CTR-Control Zone), genellikle havaalanını merkezine alacak şekilde çevreleyen, boyutları tanımlanmış silindirik hava sahasıdır. TMA bir ya da daha çok büyük havaalanı etrafında kurulu olan ve hava alanları ile hava trafik yol ağının arasındaki geçişin sağlandığı hava sahasıdır. İniş veya kalkış yapan uçakların alçalma veya tırmanma hareketlerini gerçekleştirdiği ve aynı zamanda iniş uçaklarının trafik sıralamasının yapıldığı TMA'lar, yön, seviye ve hız gibi çok boyutlu değişikliklerin aynı anda meydana geldiği karmaşık hava sahalarıdır. Koridor şeklinde tesis edilmiş ve seyrüsefer cihazlarıyla donatılmış hava yolları, birleşerek hava trafik yol ağı sistemini oluşturur. Uçuşlar, hava yolları üzerinden planlanır ve gerçekleştirilir. Uçuş Bilgi Bölgesi (FIR- Flight Information Region) içerisinde uçuş bilgi hizmeti ve ikaz hizmetinin

sağlandığı, sınırları belirli hava sahasıdır. Ülkemizde, İstanbul ve Ankara FIR olmak üzere iki adet uçuş bilgi bölgesi bulunmaktadır.

Hava araçları, aerodinamik ve fizik kurallarına göre havada tutunabilen ve taşıma işlemini gerçekleştirebilen havadan ağır araçlardır. Hava trafiği, uçuş halindeki veya bir havaalanının manevra sahasındaki bütün uçakların meydana getirdiği trafiktir (ICAO, 2005).

Havacılık faaliyetlerinde kullanılan teknik donanımlar haberleşme, seyrüsefer ve gözetim olarak üç sınıfa ayrılmaktadır. Haberleşme sistemleri, telsiz yoluyla Çok Yüksek Frekans (VHF) veya Yüksek Frekans (HF) bantları üzerinden, uçaklarla ATCo arasındaki veya pilotlarla uçak işleticileri arasındaki sözlü iletişimi sağlar. Hava trafik kontrol üniteleri arasındaki veya meteoroloji istasyonu gibi birimlerle yapılan sözlü iletişim ise direk telefon hatları ile sağlanmaktadır. Ayrıca bütün havaalanları, hava trafik kontrol merkezleri ve meteoroloji ofisleri Sabit Havacılık Telekomünikasyon Ağı (AFTN-Aeronautical Fix Telecommunications Network) ile birbirine bağlı olup aralarında uçuşun tüm aşamaları boyunca yazılı iletişim sağlayabilmektedir (Cook, 2007).

Hava yollarının tasarımında yolların sınırları, boyutları ve doğrultuları seyrüsefer sistemlerinden sağlanan pozisyon bilgileri ile belirlenmekte ve uçakların bu yollar üzerinde uçuşlarını gerçekleştirmelerinde radyo seyrüsefer yardımcıları kullanılmaktadır. Seyrüsefer sistemleri yer tabanlı veya uydu tabanlı olabilmektedir (GPS: Küresel Konumlama Sistemi).

Uçağın seyrüseferinin yerden takibini sağlayan gözetim sistemleri aracılığıyla, uçaklardan alınan hız, konum ve irtifa bilgileri hava trafik kontrolörlerine iletilir ve olası tehlikelerin önüne geçilmesi sağlanır.

Hava trafik sisteminde insan gücü, ortak bir amaç doğrultusunda çalışan, alanında özel eğitimden geçmiş, yüksek derecede sorumluluğa sahip pilot, ATCo, uçak bakım teknisyenleri, uçuş mürettebatı ve kabin ekibi gibi birçok personelden meydana gelmektedir. Emniyetli bir uçuş için her bir personelin görevini eksiksiz ve hatasız yerine getirmesi gerekmektedir.

Hava trafik hizmetleri; uçuş bilgi hizmeti, ikaz (uyarı) hizmeti ve hava trafik kontrol hizmetleri olarak üç gruba ayrılmaktadır. Bir uçağın uçuşuna başlamak üzere motor çalıştırmısından, varış meydanındaki park pozisyonuna yanaşmasına kadar geçen tüm süreçte sağlanan hava trafik kontrol hizmeti, uçaklar arasındaki ve uçaklarla manevra sahası

üzerindeki manialar arasındaki çarpışmaları önlemeyi, emniyetli, düzenli ve hızlı bir trafik akışı sağlamayı amaçlamaktadır (ICAO, 2018a). Üç tip hava kontrol hizmeti bulunmaktadır:

- i. Meydan Kontrol Hizmeti (Aerodrome Control Service: TWR)
- ii. Yaklaşma Kontrol Hizmeti (Approach Control Service: APP)
- iii. Saha Kontrol Hizmeti (Area Control Service: ACC)

Meydan kontrol hizmeti, meydan trafiğini oluşturan uçaklar için (havaalanının manevra sahasında bulunan ve meydan civarında uçan) verilen hava trafik kontrol hizmetidir. Yaklaşma kontrol hizmeti TMA içinde yer alan uçaklara verilen hizmettir. Yaklaşma kontrol ünitesi, havaalanına inmek üzere saha kontrol ünitesinden gelen uçakların sıralamalarının yapılarak, gerekli ayırma standartlarına göre havaalanına yönlendirilmesinden ve havaalanından kalkan uçakların uçuş seviyelerini almış bir şekilde sıralanarak uçuş planındaki uçuş rotalarına yönlendirilmesinden sorumludur. Yol kontrol hizmeti olarak da isimlendirilen saha kontrol hizmeti, FIR sınırları içerisindeki transit uçakların seyrüseferlerinin emniyetli bir şekilde tamamlanmasından ve FIR içerisinde bir havaalanına iniş veya bir havaalanından kalkış operasyonu düzenleyecek uçakların alçalma veya tırmanma hareketlerinin koordine edilmesinden sorumludur.

Uçuş bilgi hizmeti FIR içerisinde verilen, uçuşların etkili ve emniyetli bir şekilde yapılabilmesi için yararlı tavsiye ve güncel bilgileri vermek amacıyla sağlanan hizmettir. Uçuş bilgi hizmeti hava trafik kontrol hizmeti veren ünite veya Uçuş Bilgi Merkezi (FIC: Flight Information Centre) tarafından sağlanabilir. FIC, uçuş bilgi hizmeti ve ikaz hizmeti sağlamak için tesis edilmiş bir ünedir.

Uçuş bilgi hizmetleri için bilgi akışı HF, VHF ve Otomatik Terminal Bilgi Hizmeti (ATIS: Automatic Terminal Information Service) olmak üzere üç şekilde yapılmaktadır. Bu yayınlarla havadaki ve yerdeki uçaklara, ilgili meydanlar ve hava yollarındaki meteorolojik durum ve iniş, kalkış ve uçuş için gerekli diğer bilgiler belirli frekanslardan aktarılır. Meteorolojik durumda ve diğer bilgilerde önemli değişiklikler olduğunda bu yayınlar yenilenerek verilmeye devam edilir (ICAO, 2018a).

İkaz hizmeti de FIR içerisinde verilen, arama ve kurtarmaya ihtiyaç duyan uçaklar hakkında ilgili kuruluşları haberdar etmek ve gerektiğinde bu kuruluşlara yardımcı olmak amacıyla sağlanan bir hizmettir.

Hava trafik sisteminin çıktısı hava trafik hizmetleri aracılığı ile ortaya çıkan hava trafik akışıdır.

### 3.2. Yeşilköy TMA Yapısı ve Kabuller

Yeşilköy TMA içerisinde İstanbul Atatürk Havalimanı, İstanbul Sabiha Gökçen Havalimanı, Tekirdağ Çorlu Atatürk Havalimanı ve İstanbul Havalimanı olmak üzere 4 uluslararası havalimanına hizmet veren, yoğun trafik talebinin olduğu bir hava sahasıdır. Yeşilköy TMA'nın çevresi 458 nm, kapladığı alan ise 30,837  $km^2$ 'dir (Şekil 3.1). Son 5 yıl içerisinde bu havalimanlarına düzenlenen toplam operasyonların yıllara göre dağılımı Çizelge 3.1'de verilmiştir (http-3).

**Tablo 3.1.** *Yeşilköy TMA Trafik İstatistikleri*

|                               | 2016      | 2017      | 2018      | 2019      | 2020    |
|-------------------------------|-----------|-----------|-----------|-----------|---------|
| <b>İstanbul Atatürk</b>       | 466.396   | 460.785   | 464.646   | 138.279   | 34.218  |
| <b>İstanbul Sabiha Gökçen</b> | 231.927   | 220.171   | 232.275   | 235.717   | 116.725 |
| <b>Tekirdağ Çorlu Atatürk</b> | 28.252    | 35.906    | 22.906    | 29.106    | 19.151  |
| <b>İstanbul</b>               | -         | -         | 775       | 329.900   | 170.699 |
| <b>Toplam Yeşilköy</b>        | 726.575   | 716.862   | 720.602   | 733.002   | 340.793 |
| <b>Toplam Türkiye</b>         | 1.452.995 | 1.500.457 | 1.544.169 | 1.556.417 | 793.899 |

İstanbul Atatürk Havalimanı'nın 05/23, 35L/17R ve 35R/17L pisti olmak üzere üç adet pisti bulunmaktadır. Günümüzde sadece 05/23 pisti kullanılmaktadır. Sabiha Gökçen Havalimanı'nın 06/24 doğrultusunda bir adet pisti, Çorlu Havalimanı'nın 04/22 doğrultusunda bir adet pisti ve 29 Ekim 2018 tarihinde hizmete giren İstanbul Havalimanı'nın halihazırda 17R/35L, 17L/35R ve 18/36 olmak üzere 3 adet aktif kullanılan pisti bulunmaktadır.

Çalışma için gerekli veriler 2018 yılı Nisan ayında toplanmaya başlanmıştır. Bu tarihlerde İstanbul Atatürk Havalimanı Türkiye'nin en yoğun hava trafiğine sahip havaalanı olduğundan çalışma alanı olarak seçilmiştir. Üç pistin de kullanımda olduğu bu tarihlerde

19

### 3.3. Problemin Tanımı

Bu çalışmada Yeşilköy TMA içerisinde bulunan İstanbul Atatürk Havalimanı'na 01.04.2018-01.08.2018 tarihleri arasında iniş yapan uçakların uçuş bilgileri, yörünge verileri ve İstanbul Atatürk Havalimanı METAR raporları kullanılarak, kalkıştan itibaren ve TMA'ya girişten itibaren ETA tahminleri ile TMA içerisinde izledikleri yörünge tahmini makine öğrenmesi yöntemleri kullanılarak gerçekleştirilmiştir.

### 3.4. Veri Kümesi

ML modellerinin eğitimi için kullanılacak verilerin ve veri kaynaklarının seçiminde alan bilgisi, daha önce konuyla ilgili yapılmış çalışmalar, verilerin erişilebilirliği, doğruluğu vb. birçok etken etkili olmaktadır. ML ile ETA ve yörünge tahmini alanlarında yapılan çalışmalar incelendiğinde genellikle uçuşla ilgili genel bilgilerin, meteorolojik verilerin, uçak performans verilerinin ve 4 boyutlu yörünge verilerinin kullanıldığı görülmektedir (Ayhan, Costas and Samet, 2018; Kern, Medeiros and Yoneyama, 2015; Legee, Paassen and Mulder, 2013; Takacs, 2014; Wang, Liang and Delahaye, 2020; Ma and Tian, 2020). Çalışmaların birçoğunda ortak olarak kullanılan veriler Tablo 3.2'de gösterilmektedir.

**Tablo 3.2.** *ETA ve yörünge tahmininde kullanılan ortak veriler*

| Genel uçuş bilgileri          | Yörünge verileri | Hava durumu    |
|-------------------------------|------------------|----------------|
| Çağrı numarası (ICAO/IATA)    | Zaman            | Zaman          |
| Kalkış havaalanı              | Enlem            | Rüzgâr yönü    |
| Variş havaalanı               | Boylam           | Rüzgâr hızı    |
| Havayolu firması              | İrtifa           | Sıcaklık       |
| Uçak tipi                     | Hız (CAS/TAS)    | Basınç         |
| Planlanan kalkış zamanı (STD) | Baş açısı        | Görüş mesafesi |
| Gerçek kalkış zamanı (ATD)    |                  |                |
| Planlanan varış zamanı (STA)  |                  |                |
| Gerçek varış zamanı (ATA)     |                  |                |

Bu verilere ilave olarak bazı çalışmalarda, trafik yoğunluğu (kalkış ve iniş sayıları), kapı ve pist gecikmeleri, varışa kalan mesafe, yüksek irtifa rüzgâr ve sıcaklık verileri gibi verilerin de kullanıldığı görülmektedir.

Çalışma kapsamında 2018 yılı Nisan-Ağustos ayları arasında İstanbul Atatürk Havalimanı'na (LTBA) iniş yapan uçakların genel uçuş bilgileri, yörünge verileri ve aynı tarih aralığındaki LTBA Havacılık Amaçlı Rutin Hava Raporu (METAR) verileri “kml”, “allft+” ve düz metin gibi farklı formatlarda “FlightRadar24 (FR24)” ([http-4](#)), “DDR2” ([http-5](#)) ve “IOWA State University Iowa Environmental Mesonet (IEM)” ([http-6](#)) kaynaklarından elde edilmiştir.

FR24 dünya çapında hava trafiğini canlı olarak takip eden ve görsel hale getiren bir uçuş takip uygulamasıdır. FR24, ADS-B, Multilateration (MLAT) ve radar gibi birçok veri kaynağından elde ettiği verileri bir araya getirir. Bu verilere ek olarak havayolu firmalarından ve havaalanlarından elde edilen tarife ve uçuş durumu bilgileri de FR24 tarafından sağlanır. FR24, dünya genelinde 20 binden fazla alıcı ile en büyük ADS-B ağına sahiptir ve seyir irtifasında Avrupa ve ABD hava sahasının tamamı kapsama alanındadır. ADS-B donanımına sahip olmayan uçakların konumları MLAT yardımıyla hesaplanır. Ayrıca 2 yıl öncesine kadar tüm uçuş verileri “.kml” ve “.csv” uzantılı olarak elde edilebilir.

Keyhole Markup Language (KML), Open Geospatial Consortium (OGC) tarafından geliştirilen ve coğrafi verilerin Google Earth gibi uygulamalarda görüntülenebildiği bir formattır ([http-7](#)). Extensible Markup Language (XML) standardında olan KML dosyaları, etiket tabanlı yapıya sahip, iç içe geçmiş elemanlardan ve bu elemanların özelliklerinden oluşur. Şekil 3.3'te FR24'ten indirilen KML dosyalarının genel yapısı gösterilmektedir.

Eurocontrol tarafından geliştirilen ve 2013 yılında kullanıma açılan DDR2 web portalı basit bir ara yüze ve kapsamlı bir işlevselliğe sahiptir. Eurocontrol, geçmiş Avrupa hava trafiği verilerini, stratejik trafik tahminlerini, çevresel verileri, analiz raporlarını ve araçlarını DDR2 hizmeti aracılığıyla ATM paydaşlarına sağlamaktadır. Bütün bu bilgiler erişimi kısıtlı olan DDR2 tarafından yönetilmekte ve veriler so6, allft+ vb. farklı formatlarda sunulmaktadır.

Allft+ günlük olarak Avrupa hava sahasındaki bütün uçuşların bilgilerini içeren, yüksek dosya boyutuna sahip (>1GB) “.txt” uzantılı bir veri dosyasıdır. Her kayıt bir satırda

tutulur ve bu kayda ait bütün alanlar özel bir karakterle birbirinden ayrılır. Bu alanlar üç ana başlık altında toplanabilir:

- 1 – uçuşla ilgili genel bilgileri içeren alanlar,
- 2 – CDM ile ilgili alanlar,
- 3 – uçuş profili ile ilgili alanlar (nokta profili, sektör profili vb.)

Özellikle nokta profili ve sektör profili FR24 verilerinde yer almayan faydalı verilerdir. Bu verilerden uçağın TMA'ya giriş zamanı, uçuş seviyesi, konumu ve uçağın yaklaşma sırasında hangi STAR'ı izlediği de belirlenebilir. Bu veriler Filed Tactical Flight Model (FTFM), Regulated Tactical Flight Model (RTFM), Current Tactical Flight Model (CTFM) ve Correlated Position Reports for a Flight (CPF) gibi farklı versiyonlarda sunulmaktadır. Bu versiyonlar arasında yörüngenin en doğru biçimde ifade edildiği versiyon CPF'tir ancak çoğu uçuş için CPF verisi bulunmamaktadır. Bu durumda yörünge profiline en yakın versiyon CTFM olarak kabul edilmektedir (Horst and Schindler, 2012). Nokta profilinde ve sektör profilinde yer alan bilgiler Tablo 3.3'te gösterilmektedir.

Havacılık Amaçlı Rutin Hava Raporu (METAR) bilgileri, pilotun uçuş öncesinde veya uçuş esnasında kalkış, varış ve yedek havaalanlarındaki meteorolojik şartlardan haberdar olmasını sağlamaktadır ([http-8](http://8)).

```

<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Document>
    <name>Uçuş Numarası</name>
    <description>
      Kalkış/Varış Havaalanları
      Planlanan ve Gerçekleşen Kalkış/Varış Havaalanları vb.
    </description>
    <Folder>
      <name>Route</name>
      <Placemark>
        Enlem/Boylam/İrtifa/Zaman/Hız/Baş Açısı
      </Placemark>
      .
      .
      .
      <Placemark>
        Enlem/Boylam/İrtifa/Zaman/Hız/Baş Açısı
      </Placemark>
    </Folder>
    <Folder>
      <name>Trail</name>
      <Placemark>
        ...
      </Placemark>
      .
      .
      .
      <Placemark>
        ...
      </Placemark>
    </Folder>
  </Document>
</kml>

```

Şekil 3.3. FR24 “.kml” uzantılı dosya yapısı

Tablo 3.3. Allft+ nokta ve sektör profili

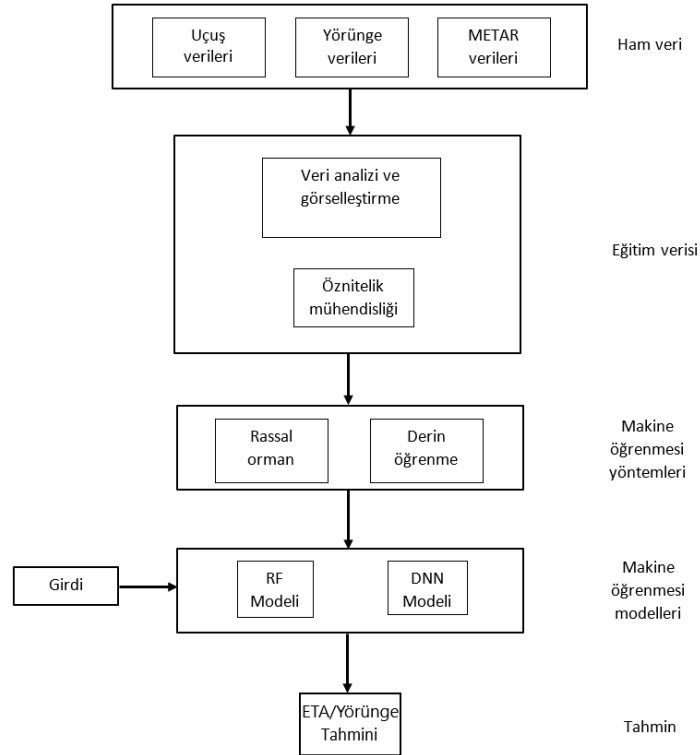
| Nokta Profili                                       | Sektör Profili                             |
|-----------------------------------------------------|--------------------------------------------|
| Zaman                                               | Hava sahasının adı                         |
| Nokta adı                                           | Hava sahasına giriş zamanı                 |
| Rota adı                                            | Hava sahasından çıkış zamanı               |
| Nokta üzerindeki uçuş seviyesi                      | Giriş noktasının enlem/boylam bilgileri    |
| Uçuş başlangıcından itibaren mesafe                 | Çıkış noktasının enlem/boylam bilgileri    |
| Nokta türü (hava alanı, tanımlı nokta, coğrafi vb.) | Giriş noktasındaki uçuş seviyesi           |
| Enlem                                               | Çıkış noktasındaki uçuş seviyesi           |
| Boylam                                              | Uçuş başlangıcından giriş noktasına mesafe |
| Bağlı mesafe                                        | Uçuş başlangıcından çıkış noktasına mesafe |

Sonuç olarak bu çalışma kapsamında, 01.04.2018-01.08.2018 tarihleri arasında İstanbul Atatürk Havalimanı'na iniş yapan 63.460 uçuşa ait uçuş bilgileri, 18.869.049 adet yörünge noktasına ait bilgiler ve 6.236 adet METAR raporu elde edilmiştir.



#### 4. METODOLOJİ

ML uygulamalarında problemin tanımı yapıldıktan ve problemle ilişkili veriler elde edildikten sonra izlenmesi gereken bir süreç bulunmaktadır (http-13). Öncelikle düzenli veya düzensiz yapıda ve farklı formatlarda olabilen bu veriler tablo formatına dönüştürülerek düzenli hale getirilmeli ve aralarındaki ilişkiler kurulmalıdır. Veriler incelenerek hatalı, eksik veya uygulama kısıtlarını sağlamayan veriler elenmeli ve veri temizliği yapılmalıdır. Veri temizliğinden sonra öznitelik mühendisliği yapılarak makine öğrenmesi modellerinde kullanılacak öznitelikler (değişkenler) belirlenmeli ve verilerin özniteliklere dönüştürülmesi için gerekli işlemler yapılmalıdır. Daha sonra problemin tanımına uygun olarak, kullanılacak ML yöntemleri belirlenmeli, veriler eğitim ve test verileri şeklinde ayrıldıktan sonra eğitim verileri kullanarak model eğitilmelidir. Eğitilen model test verileri kullanılarak test edilmeli ve modelin performansı ölçülmelidir. Modellerin performansının artırılması için gerekirse bütün süreç tekrarlanabilir. Bu çalışmada izlenen metodolojinin özeti Şekil 4.1’de gösterilmektedir.



Şekil 4.1. Metodoloji

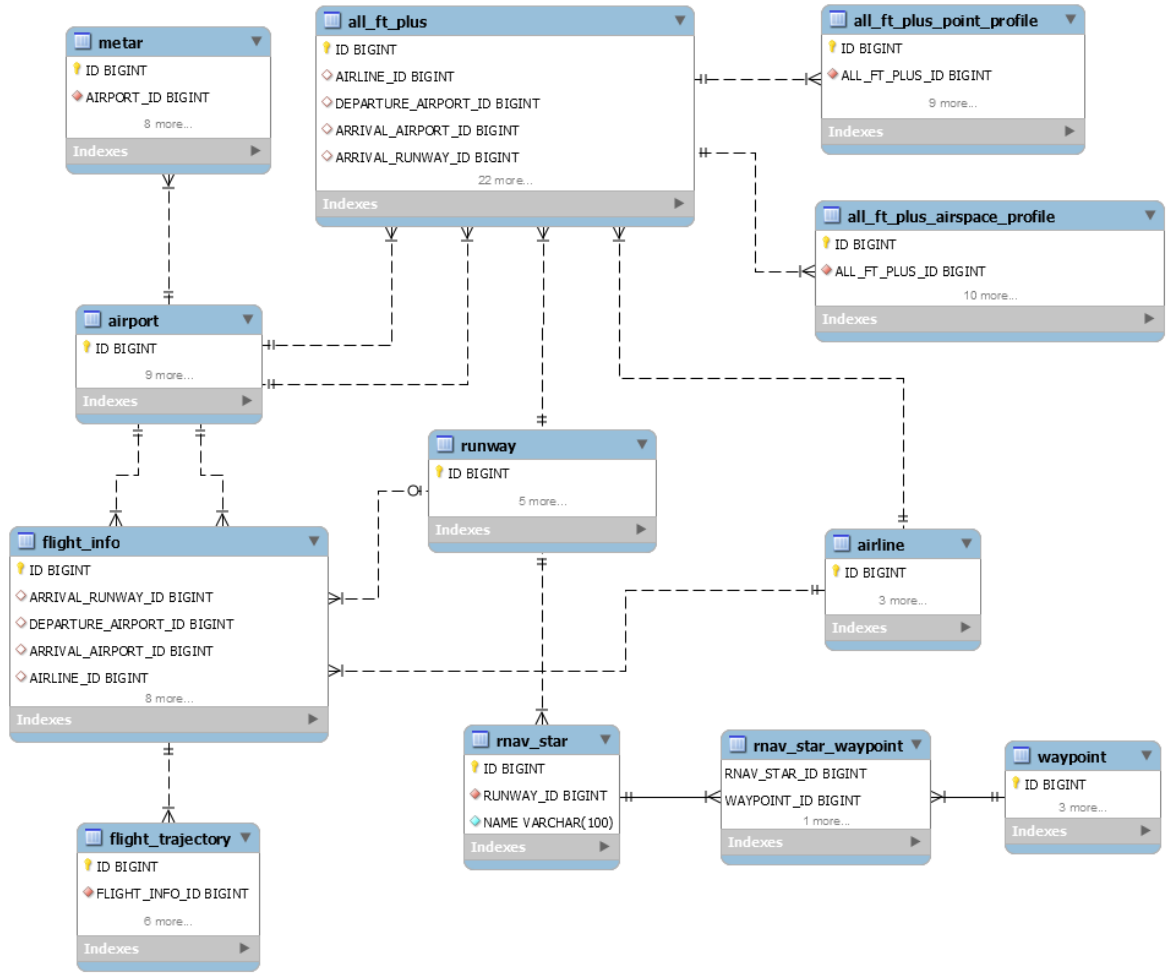
Bu aşamalara ek olarak elde edilen ML modellerinin geliştirilen uygulamaya entegrasyonu sağlanarak, modellerin yeni girdilerle dinamik olarak kullanılabilmesi sağlanmıştır.

#### **4.1. Verilerin Veri Tabanına Aktarılması**

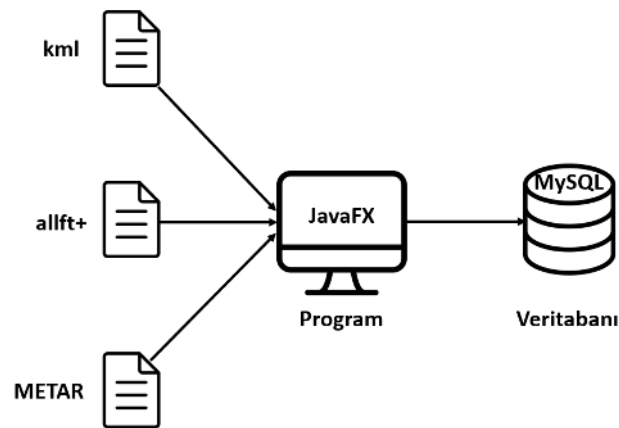
ML yöntemleri genellikle düzenli veri kümeleri üzerinde çalışır. Düzenli veri kümeleri yorumlanması, yönetimi ve görselleştirilmesi kolay yapılardır ve aşağıdaki özelliklere sahiptir (http-9):

- Her değişken bir sütuna karşılık gelir
- Her gözlem bir satıra karşılık gelir
- Gözlemlenen her birim bir tabloya karşılık gelir

Dolayısıyla ilk olarak “.kml” ve “.txt” uzantılı dosyalarda saklanan verilerin, ML uygulamalarında kullanılacak, yorumlanması ve görselleştirilmesi kolay, düzenli ve ilişkili veriler haline dönüştürülmesi gerekmektedir. Bunun için ilişkisel bir veri tabanı tasarlanmış (Şekil 4.2), geliştirilen bir program aracılığıyla, elde edilen ham veriler farklı dosya tiplerinden okunarak veri tabanı tasarımına uygun olarak dönüştürülmüş ve veri tabanına aktarılmıştır (Şekil 4.3). Bu aşamada boş veya hatalı verilerle ilgili herhangi bir düzeltme veya filtreleme yapılmamış, verilerin olduğu gibi veri tabanına aktarılması sağlanmıştır. Uçakların yörüngelerinin analizinde ve hatalı, eksik verilerin belirlenmesinde önemli bir bilgi olan iniş pisti bilgisi FR24 verilerinden direk olarak elde edilemediği için, iniş pisti uçağın son yaklaşımadaki yörüngesinin baş açılarının ortalaması alınarak belirlenmiştir.



Şekil 4.2. Veri tabanı yapısı

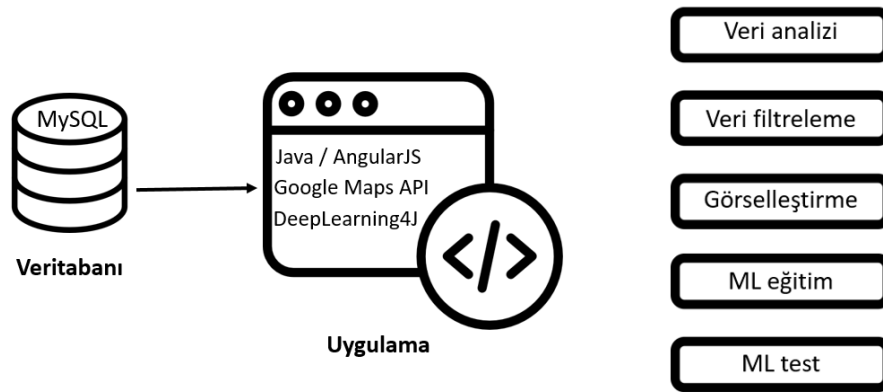


Şekil 4.3. Ham verilerin veri tabanına aktarılması

#### 4.2. Veri Analizi, Görselleştirme ve Veri Temizliği

Veri tabanına aktarılan uçuş, yörünge ve METAR verilerinin analizi, yörüngelerin harita üzerinde görselleştirilmesi, hatalı verilerin tespiti ve temizlenmesi, özniteliklerin ve eğitim verilerinin oluşturulması, ML modellerinin eğitimi ve elde edilen ML modelleri kullanılarak yeni tahminler yapılabilmesi için web tabanlı bir uygulama geliştirilmiştir (Şekil 4.4). Bu uygulama ile uçuş bilgileri çağrı işareti, iniş zamanı vb. filtreler ile filtrelenerek uçuş, yörünge ve METAR verileri tablo formatında elde edilebilmekte ve yörüngeler harita üzerinde görüntülenebilmektedir (Şekil 4.5, Şekil 4.6). Ayrıca uçuş, yörünge ve METAR bilgilerinin çıktıları alınabilmektedir. Geliştirilen programda İstanbul Atatürk Havalimanı için standart geliş rotaları (STAR) da görüntülenebilmektedir.

Uygulama kullanılarak farklı günlerdeki uçuş bilgileri liste şeklinde elde edilmiş, bu listeler incelendiğinde bazı verilerde STD, STA, ATD, iniş pisti, uçak tipi, havayolu firması vb. önemli bazı bilgilerin eksik veya hatalı olduğu görülmüştür. Ayrıca yörüngeler harita üzerinde görsel olarak incelendiğinde bazı uçuşlarda yörüngelerin iniş/kalkış kısımlarının eksik olduğu veya yörüngelerdeki zamanlarla uçağın iniş/kalkış zamanlarının uyuşmadığı gözlemlenmiştir. Bazı uçuş verilerinin tekrarladığı, bazılarının ise hatalı eşleştirildiği gözlemlenmiştir. Eğitim verileri oluşturulmadan önce öznitelik mühendisliği aşamasında bu verilerin ayıklanması veya düzeltilmesi sağlanmıştır.

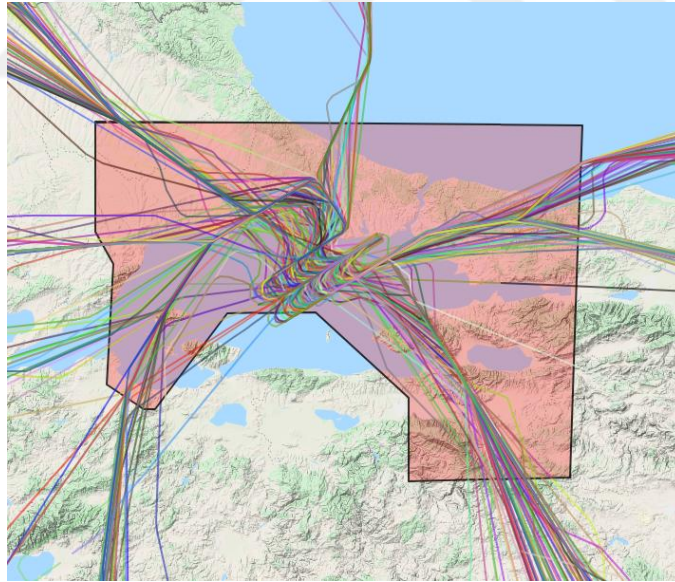


Şekil 4.4. Geliştirilen web uygulamasının yapısı

| CALLSIGN    | DEP. | STD                       | ATD                       | STA                       | ETA                       | TMA ENTRY           | TMA COORD.             | TMA FL | RWY | AIRCRAFT        |
|-------------|------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------|------------------------|--------|-----|-----------------|
| TK2403<br>⏚ | AYT  | 14-05-2018<br>00:20 +0300 | 14-05-2018<br>00:25 +0300 | 14-05-2018<br>01:35 +0300 | 14-05-2018<br>01:20 +0300 | 14-05-2018<br>00:53 | 39.911667<br>29.667778 | 236    | 05  | Airbus A321-231 |
| TK2105<br>⏚ | ESB  | 14-05-2018<br>00:35 +0300 | 14-05-2018<br>00:32 +0300 | 14-05-2018<br>01:45 +0300 | 14-05-2018<br>01:20 +0300 | 14-05-2018<br>00:52 | 40.877778<br>30.597500 | 244    | 05  | Airbus A321-231 |
| TK2305<br>⏚ | ADB  | 14-05-2018<br>00:45 +0300 | 14-05-2018<br>00:45 +0300 | 14-05-2018<br>01:55 +0300 | 14-05-2018<br>01:41 +0300 | 14-05-2018<br>01:07 | 40.300278<br>27.375000 | 244    | 05  | Airbus A321-231 |
| TK1070<br>⏚ | VAR  | 14-05-2018<br>00:50 +0300 | 14-05-2018<br>00:55 +0300 | 14-05-2018<br>01:45 +0300 | 14-05-2018<br>01:42 +0300 | 14-05-2018<br>01:10 | 42.082222<br>28.766111 | 245    | 05  | Boeing 737-8F2  |
| TK2451<br>⏚ | ADA  | 14-05-2018<br>00:50 +0300 | 14-05-2018<br>00:48 +0300 | 14-05-2018<br>02:25 +0300 | 14-05-2018<br>02:07 +0300 | 14-05-2018<br>01:41 | 40.871944<br>30.630556 | 244    | 05  | Boeing 737-8F2  |

First Prev 1 2 3 4 5 ... 105 Next Last

Şekil 4.5. Verilerin uygulamada görüntülenmesi



Şekil 4.6. Yörüngelerin uygulamada görüntülenmesi

#### 4.3. Öznitelik Mühendisliği

Öznitelik mühendisliği ham verilerden özniteliklerin çıkartılması ve ML yöntemlerine uygun formatlara dönüştürülmesi işlemidir (Zheng and Casari, 2018).

Bu aşamada öncelikle veri analizi ve görselleştirme aşamasında tespit edilen hatalı ve tekrarlayan verilerin ayıklanması sağlanmıştır. Yörünge noktaları, STD, ATD, STA, ATA, iniş pisti, uçak tipi, havayolu firması vb. eksik olan veriler ayıklanmış, tekrarlayan veriler tespit edilerek tekrarlardan bir tanesi ayıklanmıştır. Özellikle allft+ verilerinde kalkış ve iniş günleri arasında gün farkı olan verilerin her iki gün için oluşturulan allft+ dosyasında da yer aldığı gözlemlenmiştir; ayrıca aynı uçuş bilgilerinin uçuş durumundaki değişiklikler nedeniyle de birden çok kaydedildiği belirlenmiştir. Örneğin iptal edilen bir uçuş operasyonu daha sonra gerçekleştirildiğinde uçuş durumu ‘Cancelled’ olan bir veri ile aynı uçuşa ait uçuş durumu ‘Terminated’ olan bir başka veri daha bulunmaktadır. Bu gibi durumlarda uçuş durumu ‘Terminated’ olan veriler dikkate alınmıştır.

Tarih saat formatları genellikle ML uygulamalarında kullanıma uygun değildir. Bu nedenle bu tarz veriler ay, gün, saat, dakika vb. şeklinde ayrı değerlere bölünerek farklı öznitelikler şeklinde kullanılmaktadır. STD, STA, ATD ve TMA giriş saati gibi veriler bunlara örnek olarak verilebilir. Bütün veriler 2018 yılına ait olduğu için yıl verisi dikkate alınmamış, uçuşun yapıldığı gün 1-365 arasında bir değer olarak ifade edilmiştir. Benzer biçimde 00:00-23:59 aralığındaki saat verisi, gece yarısından itibaren geçen dakika olarak hesaplanarak 0 – 1439 aralığına dönüştürülmüştür. Tablo 4.1’de örnek bir tarih saat dönüşümü gösterilmektedir.

Veri kümesindeki uçuşların %90’ından fazlasının ulusal havayolu firmaları tarafından gerçekleştirildiği gözlemlenmiş, bu nedenle havayolu firmaları ‘Türk Hava Yolları’, ‘Atlas Global’, ‘Onur Air’, ‘Pegasus’ ve ‘Diğer’ olarak 5 kategoriye ayrılmıştır.

Veride üretici ve model şeklinde detaylı tanımlanan uçak tipleri ICAO kuyruk türbülansı kategorilerine göre ‘Hafif’, ‘Orta’ veya ‘Ağır’ kategorilerine dönüştürülmüştür.

Diğer birçok çalışmadan farklı şekilde kalkış/iniş havaalanı çiftleri kategorik olarak kullanılmak yerine, yörüngelerden toplam uçuş mesafeleri hesaplanmış, böylece aynı havaalanından kalkış yapan bütün uçaklar için kalkış havaalanının aynı etkiye sahip olmaması sağlanmıştır.

Allft+ ve FR24 verileri eşleştirilerek TMA giriş zamanı ve koordinatları belirlenmiş, TMA’ya girişten itibaren inişe kadar olan mesafe yine yörünge noktalarından hesaplanmıştır.

METAR verileri orijinal formatından dönüştürülerek birçok farklı öznitelik elde edilmiş, uçağın gerçek iniş zamanına en yakın METAR raporu ilgili uçuşla eşleştirilerek, iniş

zamanında havaalanı civarındaki gerçek hava koşulları modellenmeye çalışılmıştır. Tablo 4.1’de örnek bir METAR dönüşümü gösterilmektedir.

*Tablo 4.1. Tarih-saat dönüşümü örneği*

| Tarih saat       | Yılın günü | Gece yarısından itibaren geçen dakika |
|------------------|------------|---------------------------------------|
| 2018-04-01 09:50 | 91         | 590                                   |
| 2018-05-28 17:15 | 148        | 1035                                  |

*Tablo 4.2. METAR dönüşümü örneği*

|                              |                                              |
|------------------------------|----------------------------------------------|
| <b>Orijinal METAR raporu</b> | LTBA 010050Z 16004KT CAVOK 11/08 Q1010 NOSIG |
| <b>Gözlem zamanı</b>         | 01-04-2018 00:50                             |
| <b>Rüzgâr yönü (derece)</b>  | 160                                          |
| <b>Rüzgâr hızı (kt)</b>      | 4                                            |
| <b>Sıcaklık (derece)</b>     | 11                                           |
| <b>Çiy noktası</b>           | 8                                            |
| <b>Basınç</b>                | 1010                                         |
| <b>Görüş mesafesi</b>        | 9999                                         |

#### 4.4. Makine Öğrenmesi Yöntemleri

Gerek ETA tahmini gerekse yörünge tahmini probleminde elimizdeki veriler hem girdi vektörlerini hem de bunlara karşılık gelen çıktıları içermektedir. Dolayısıyla bu problemlerin ML yöntemleriyle çözümünde denetimli öğrenme yöntemleri kullanılmalıdır. Ayrıca tahmin edilmek istenen değerler sayısal değerler olduğundan bu problemler regresyon problemleridir.

RF yöntemi hem kategorik hem de sürekli öznitelikleri bir arada kullanabilen, çok büyük veri kümelerinde çok fazla sayıda öznitelikle dahi hızlı ve etkili bir biçimde çalışabilen ve aşırı uyum eğilimi düşük olan bir yöntemdir (Leege, Paassen and Mulder, 2013). Ayrıca RF hem sınıflandırma hem de regresyon problemlerinde kullanılabilir.

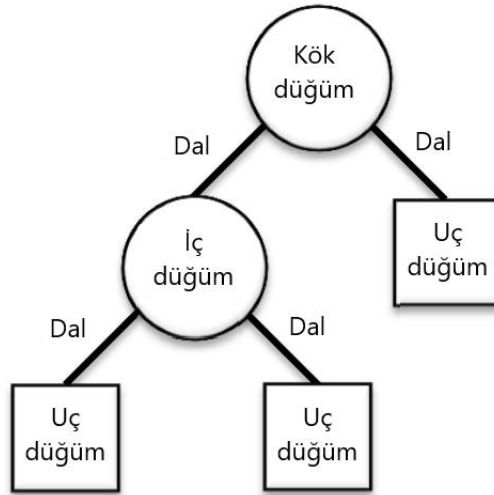
RF, birçok karar ağacının birleşiminden meydana gelen, bireysel öğrenme modellerini bir arada kullanarak tahmin doğruluğunu arttırmayı amaçlayan bir toplu öğrenme yöntemidir

(Breiman, 2001). Dolayısıyla RF yöntemini anlayabilmek için öncelikle karar ağaçları açıklanmalıdır.

#### 4.4.1. Karar ağaçları

Karar ağaçları uygulama ve yorumlama açısından kolay bir yöntemdir. Hem regresyon hem de sınıflandırma problemlerine uygulanabilen karar ağaçları, grafiksel olarak ifade edilebilir ve uzman olmayan kişiler tarafından dahi kolayca yorumlanabilir. (James vd., 2017).

Karar ağaçları, kök düğümden başlayarak uç düğümlerde (yapraklarda) sonlanan, dallarla birbirine bağlı karar düğümlerinden meydana gelir (Şekil 4.7) (Ghatak, 2017).



Şekil 4.7. Karar ağacı yapısı

Her bir  $m$  karar düğümünde  $x$  girdisine bir  $f_m(x)$  fonksiyonu uygulanır ve bu fonksiyonun çıktısına bağlı olarak karar düğümünün dallarından birisi seçilir. Bu süreç kök düğümde başlar ve bir uç düğüme erişilene kadar devam eder. Erişilen uç düğümün değeri ağacın ilgili girdiye karşılık gelen çıktısını ifade eder (Alpaydın, 2010).

Her bir  $f_m(x)$  fonksiyonu  $d$  boyutlu girdi uzayını daha küçük bölgelere ayırır. Böylece ağaç büyüdükçe karmaşık bir fonksiyon basit kararlar şekline dönüşür. Sınıflandırma problemlerinde her bir uç düğüm bir sınıfa karşılık gelirken, regresyon problemlerinde uç düğümler sayısal değerlere karşılık gelir. Uç düğümler, girdi uzayındaki yerel bölgeleri

tanımlar. Yerel bölgeler içerisinde kalan örnekler aynı sınıfa veya çok benzer sayısal değerlere sahiptir.

Tek değişkenli bir ağaçta, her bir ara düğümde verilecek kararlarda özniteliklerden sadece birisi kullanılır. Eğer  $x_j$  özniteliği ayırık değerli ise  $n$  adet değerden birini alır. Bu durumda, karar düğümünde  $x_j$  değeri kontrol edilerek  $n$  adet olası daldan  $x_j$  değerine karşılık gelen dal seçilir.  $x_j$  özniteliği sayısal değerli ise ayırık değere dönüştürülmelidir. Bu durumda  $f_m(x)$  fonksiyonu bir karşılaştırma işlemi olarak tanımlanır:

$$f_m(x) : x_j > w_{m0} \quad (4.1)$$

Burada  $w_{m0}$  uygun bir eşik değerdir. Karar düğümü girdi uzayını ikiye böler:

$$L_m = \{x | x_j > w_{m0}\} \quad (4.2)$$

$$R_m = \{x | x_j \leq w_{m0}\} \quad (4.3)$$

Buna ikili ayırma adı verilir. Kök düğümünden uç düğüme kadar izlenen yol üzerindeki ardışık karar düğümleri, farklı öznitelikleri kullanarak girdi uzayında birbirine dik ayrımlar meydana getirir.

Sınıflandırma ağaçlarında ayırmanın başarısı safsızlık (katışıklık) ölçüsü ile ölçülür. Eğer ayırmadan sonra bütün dallar için, ilgili daldaki bütün örnekler aynı sınıfa aitse o ayırma saftır. Örneğin  $m$  düğümü için,  $N_m$ , eğitim kümesinde  $m$  düğüme ulaşan gözlemler olsun. Bu durumda bütün  $N_m^i, C_i$  sınıfına aittir:

$$\sum_i N_m^i = N_m \quad (4.4)$$

$m$  düğüme ulaşan bir örnek verildiğinde,  $C_i$  sınıfının olasılığı şu denklemlerle ifade edilir:

$$\hat{P}(C_i | x, m) \equiv p_m^i = \frac{N_m^i}{N_m} \quad (4.5)$$

Bütün  $i$  değerleri için  $p_m^i$  0 veya 1 ise  $m$  düğümü saftır.  $m$  düğüme ulaşan gözlemlerin hiçbirisi  $C_i$  sınıfına ait değilse  $p_m^i = 0$ , hepsi  $C_i$  sınıfına ait ise  $p_m^i = 1$  olur. Eğer ayırma saf ise daha fazla ayırma yapılmaz ve  $p_m^i = 1$  olan sınıf etiketine sahip bir uç düğüm eklenir. Safsızlığı ölçmek için entropi, Gini endeksi (Breiman vd., 1984) veya yanlış sınıflandırma hatası kullanılabilir. Entropi şu şekilde tanımlanır:

$$I_m = - \sum_{i=1}^K p_m^i \log_2 p_m^i \quad (4.6)$$

Eğer  $m$  düğümü saf değilse, örnekler safsızlığı azaltacak şekilde ayrılmalıdır. Ancak burada önemli olan ayırma yapılabilecek birçok öznitelik arasından safsızlığı minimize eden öz niteliğin belirlenmesidir. Ayırmadan sonra elde edilecek alt kümeler ne kadar saf ise, sonrasında çok daha az ayırma işlemine gerek duyulacak ve olası en küçük ağacı elde etmek mümkün olacaktır. Bu nedenle, bütün öznitelikler için safsızlık hesaplanır ve en küçük entropiye sahip olan ayırma seçilir. Ağaç oluşturma işlemi, bütün dallar saf olana kadar yinelemeli şekilde devam eder. Bu işlemler, Sınıflandırma ve Regresyon Ağaçları (CART) algoritmasının temelini oluşturur (Breiman vd., 1984).

Buradaki problemlerden birisi, bu ayırma algoritmasının kategori sayısı yüksek olan özniteliklere ayrıcalık göstermesidir. Çok kategori çok dal anlamına gelir. Bu durumda safsızlıktaki azalmanın az kategorili özniteliklere göre daha fazla olması olasıdır. Çok dallı düğümler karmaşıktır ve bu durum karar ağaçlarının karmaşık fonksiyonları basit kararlara dönüştürme düşüncesiyle çelişmektedir. Buna benzer durumların önüne geçmek için çok kategorili öznitelikleri cezalandıracak yöntemler önerilmiştir.

Gürültünün varlığı durumunda, ağaçları en saf hallerine kadar büyütme çok büyük bir ağacın oluşturulmasına ve aşırı uyuma neden olur. Örneğin doğru etiketlenmiş örneklerin arasında yanlış etiketlenmiş bir örnek olduğu durumda, aşırı uyumu önlemek için, ağaç oluşturma işlemi düğümler yeterince saf olduğunda, yani entropi değeri belli bir eşik değerden küçük olduğunda ( $I < \theta_I$ ) sonlandırılır.

Regresyon ağaçları sınıflandırma ağaçlarına benzer şekilde oluşturulur. Tek fark safsızlık ölçüsünün regresyona uygun bir ölçü ile değiştirilmesidir.  $m$  düğümü için  $X_m$ , karar düğümlerinde kök düğümünden  $m$  düğümüne kadarki bütün koşulları sağlayan  $x \in X$  kümesidir.

$$b_m(x) = \begin{cases} 1 & x \in X_m \text{ ise} \\ 0 & \text{diğer} \end{cases} \quad (4.7)$$

Regresyonda, ayırmanın iyiliği ortalama kare hata (MSE) ile ölçülür.  $g_m$ ,  $m$  düğümündeki tahmini değer olmak üzere

$$E_m = \frac{1}{N_m} \sum_t (r^t - g_m)^2 b_m(x^t) \quad (4.8)$$

Burada

$$N_m = |X_m| = \sum_t b_m(x^t) \quad (4.9)$$

Bir düğümde, düğüme ulaşan gözlemlerin ortalaması (gürültü çoksa medyanı) kullanılır.

$$g_m = \frac{\sum_t b_m(x^t) r^t}{\sum_t b_m(x^t)} \quad (4.9)$$

Bu durumda Denklem 4.8  $m$  düğümündeki varyansa karşılık gelir. Eğer bir düğümdeki hata kabul edilebilir ise, yani  $E_m < \theta_r$  ise, bu durumda bir uç düğüm oluşturulur ve bu düğüme  $g_m$  değeri atanır. Eğer hata kabul edilebilir değilse,  $m$  düğümüne ulaşan veriler dallardaki hataların toplamı minimum olacak şekilde bölünür. Sınıflandırmada olduğu gibi, her düğümde, hatayı minimize edecek öznelilik aranır ve daha sonra süreç tekrarlanır.

#### 4.4.2. Budama

Genellikle, bir düğüme ulaşan eğitim verilerinin sayısı, toplam eğitim kümesinin belli bir yüzdesinden küçük olduğunda (örneğin %5), ilgili düğüm, safsızlığa veya hataya bakılmaksızın daha fazla bölünmez. Buradaki düşünce çok az sayıdaki gözleme bağlı herhangi bir kararın varyansa ve dolayısıyla genelleme hatasına neden olacağıdır. Bir ağacın tamamen oluşturulmadan önce sonlandırılmasına ön budama adı verilir.

Daha basit ağaçlar elde etmenin bir diğer yolu da son budamadır. Son budama genellikle ön budamadan daha başarılı sonuçlar verir. Son budamada gereksiz alt ağaçların bulunarak budanmasına çalışılır.

Son budamada, bütün uç düğümler saf olana ve eğitim hatası olmayana dek ağaç büyütülür. Daha sonra aşırı uyuma neden olan ağaçlar belirlenerek budanır. Bunun için eğitim kümesinden, eğitim sırasında kullanılmayan bir budama kümesi ayrılır. Her bir alt ağaç, o ağaç tarafından kapsanan eğitim verileri ile etiketlenerek bir uç düğümlerle değiştirilir. Eğer uç düğüm, budama kümesi ile denenen alt ağaçtan daha iyi sonuç verirse alt ağaç budanır, aksi taktirde alt ağaç budanmaz.

#### 4.4.3. Torbalama

Karar ağaçları yüksek varyansa sahiptir. Eğitim verisi rastgele iki parçaya bölünüp her parça için birer karar ağacı oluşturulursa, sonuçlar birbirinden çok farklı olabilir. Breiman tarafından öne sürülen torbalama yöntemi, bir istatistiksel öğrenme yönteminin varyansını düşürmek ve doğruluğunu arttırmak için kullanılan genel amaçlı bir prosedürdür (Breiman, 1996).

Her birinin varyansı  $\sigma^2$  olan  $n$  bağımsız gözlemden  $(Z_1, Z_2, \dots, Z_n)$  meydana gelen bir veri kümesinde gözlemlerin ortalamasının  $(\bar{Z})$  varyansı  $\sigma^2/n$  ile ifade edilir. Yani bir dizi gözlemin ortalamasını almak varyansı düşürür (James vd., 2017). Dolayısıyla varyansı düşürmenin ve tahmin doğruluğunu arttırmanın yolu, popülasyondan birçok eğitim kümesi alarak, her eğitim kümesi için ayrı bir tahmin modeli üretmek ve sonuçların ortalamasını almaktır. Ancak genellikle ML uygulamalarında, birden fazla eğitim kümesine erişim sağlamak mümkün değildir. Bu problemin üstesinden gelebilmek için tek bir eğitim veri kümesinden tekrarlı örnekler alınarak  $B$  farklı eğitim veri kümesi üretilir. Diğer bir deyişle  $B$  adet farklı eğitim kümesi kullanılarak  $\hat{f}^1(x), \hat{f}^2(x), \dots, \hat{f}^B(x)$  hesaplanabilir ve bunların ortalaması alınarak düşük varyanslı bir istatistiksel öğrenme modeli elde edilebilir:

$$\hat{f}_{ort}(x) = \frac{1}{B} \sum_{b=1}^B \hat{f}^b(x) \quad (4.10)$$

Bu işleme “torbalama” adı verilir.

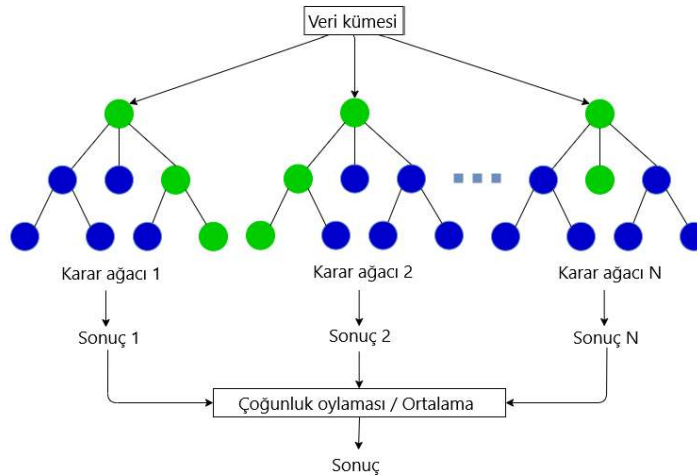
Torbalama yöntemi regresyon ağaçlarına uygulandığında  $B$  adet veri kümesinden  $B$  adet regresyon ağacı oluşturulur ve tahmin sonuçlarının ortalaması alınır. Ağaçlar budanmaz, dolayısıyla her bir ağacın varyansı yüksek ama yanlılığı düşüktür.  $B$  adet ağacın ortalaması alınarak varyans düşürülür.

Torbalama yönteminde modelin test hatasının tahmini için çapraz doğrulama veya doğrulama kümesi yaklaşımlarına gerek bulunmamaktadır. Ortalama olarak her ağaç gözlemlerin yaklaşık üçte ikisini kullanır. Ağaç oluşturmak için kullanılmayan gözlemlere “torba dışı (OOB)” gözlemler denir.  $i$ . gözlem için tahmin sonucu, o gözlemin torba dışı olduğu ağaçlar kullanılarak tahmin edilebilir. Bu da yaklaşık  $B/3$  tahmin anlamına gelir.  $i$ . gözlem için bir tek tahmin elde etmek istendiğinde bu  $B/3$  tahmin sonucunun ortalaması alınır. Böylece veri kümesinde yer alan bütün gözlemler için bir OOB gözlem tahmini elde

edilebilir ve buradan da toplam OOB MSE elde edilmiş olur. Sonuç olarak hesaplanan OOB hatası, modelin test hatası olarak kabul edilebilir çünkü her gözlem için elde edilen tahmin değeri sadece o gözlemin kullanılmadığı ağaçlardan elde edilen tahmin değeridir.

#### 4.4.4. Rassal orman

Birçok karar ağacının birleşiminden meydana gelen RF yönteminde, torbalama yönteminde olduğu gibi farklı eğitim verileri kullanılarak bir dizi karar ağacı üretilir. Ancak torbalamadan farklı olarak bu ağaçlar oluşturulurken, her bir ayırmada  $p$  adet öznelik içerisinde  $m$  adet öznelik rastgele seçilerek ağaçların birbirleriyle ilintisiz hale gelmesi sağlanır (James vd., 2017). Burada amaç veri kümesinde çok güçlü bir özneliğin varlığı durumunda, ağaçların çoğunun en üst ayırma bu özneliği kullanarak birbirlerine benzemelerini ve yüksek korelasyona sahip olmalarını engellemektir. Her bir ayırmada yeniden  $m$  adet öznelik rastgele seçilir ve sadece bu  $m$  öznelikten biri kullanılır. Uygulamada  $m \approx \sqrt{p}$  olarak kabul edilir. Çok sayıda ağaç oluşturulduktan sonra regresyon problemleri için bu ağaçların çıktılarının ortalama değeri alınırken, sınıflandırma problemlerinde çoğunluk oylaması yapılarak ağaçların çıktılarında en çok tekrar eden sınıf çıktı olarak belirlenir (Şekil 4.8).

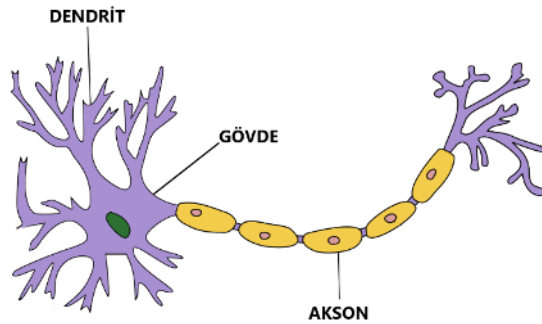


Şekil 4.8. Rassal orman

#### 4.4.5. Yapay sinir ağıları

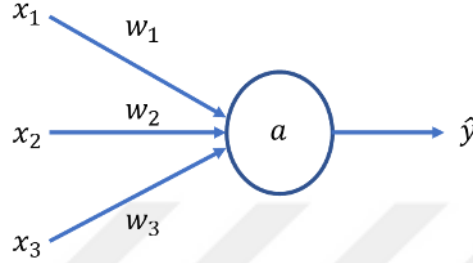
ANN, insan beyninin öğrenme algoritmasını taklit etmeyi amaçlayan bir ML yöntemidir (http-10). İnsan beyni görüntü işleme, ses tanıma, nesne tanıma, matematik, fizik vb. binlerce karmaşık görevi başarılı bir şekilde öğrenebilmektedir. Beynin çalışma mantığını inceleyen bilim adamları, “neural rewiring” adı verilen deneylerle, bazı organların sinirlerinin beyindeki bağlantı noktalarını değiştirmiş ve beynin farklı işlevleri gerçekleştirmekle görevli olduğu düşünülen bölgelerin yeni bir işlevi de öğrenerek yerine getirebildiğini gözlemlemişlerdir. Örneğin yapılan bir çalışmada kulaklarla işitme korteksi arasındaki sinirsel bağlantı koparılarak, gözdeki optik sinirler bu bölgeye bağlanmış ve işitme korteksinin görmeyi öğrendiği ifade edilmiştir (Roe vd, 1992). Ayrıca görme engelli bireylerin, insan sonarı veya insan ekolokasyonu olarak adlandırılan bir yöntemle, ayak sesleri veya ağızla yapılan seslerle ses dalgaları yaratarak ve bu dalgaların cisimlerden yansımaları dinleyerek, cisimlerin boyut ve yerini saptayabildikleri bilinmektedir. (http-11, http-12). Dolayısıyla beynin görüntüleri, dokunma duyusunu ve sesleri tek bir algoritmayla işleyebildiği düşünülmüş ve insan beyninin öğrenme algoritmasını taklit edebilmek için ANN yöntemi geliştirilmiştir.

ANN, insan beynindeki nöronların ve nöron ağlarının benzetimini yapmaktadır. İnsan beyni, temelde bir hücre gövdesinden, nörona gelen sinyalleri gövdeye ileten dendritlerden ve işlenen sinyalleri diğer nöronlara ileten aksonlardan oluşan milyarlarca sinir hücresi içermektedir (Şekil 4.9). Basit düzeyde nöron, dendritler aracılığıyla bir dizi girdiyi alarak bir hesaplama yapan ve çıktıları aksonlar aracılığıyla diğer nöronlara ileten bir hesaplama birimidir (http-13).



Şekil 4.9. Nöron yapısı

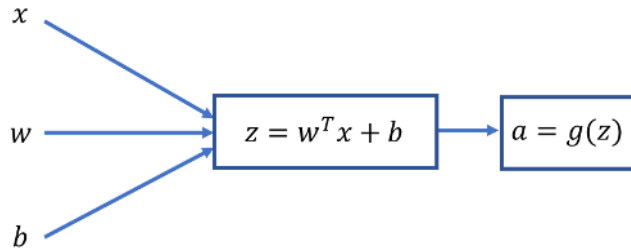
Yapay sinir ağlarında nöronlar, girdileri aktivasyon fonksiyonu adı verilen doğrusal olmayan bir fonksiyonla işleyerek bir çıktı elde eden birimler şeklinde modellenebilir. Şekil 4.10'da üç girdi ve tek çıktıya sahip örnek bir nöron gösterilmektedir.



Şekil 4.10. ANN nöron yapısı

Burada  $x = [x_1 \ x_2 \ x_3]^T$  öznitelik vektörü,  $w = [w_1 \ w_2 \ w_3]^T$  ağırlık vektörü,  $a$  aktivasyon fonksiyonu ve  $\hat{y}$  çıktı (tahmin) olarak adlandırılır.

Bu durumda tek bir nöron tarafından gerçekleştirilen işlemler Şekil 4.11'de gösterilmektedir.



Şekil 4.11. Bir nöron tarafından yapılan hesaplamalar

Burada  $b$  yanlılık (bias) vektörüdür.

Yapay sinir ağlarında sıklıkla kullanılan bazı aktivasyon fonksiyonları doğrultulmuş doğrusal birim (ReLU), tanjant hiperbolik (tanh), sigmoid (lojistik) ve kimlik (identity) fonksiyonlarıdır:

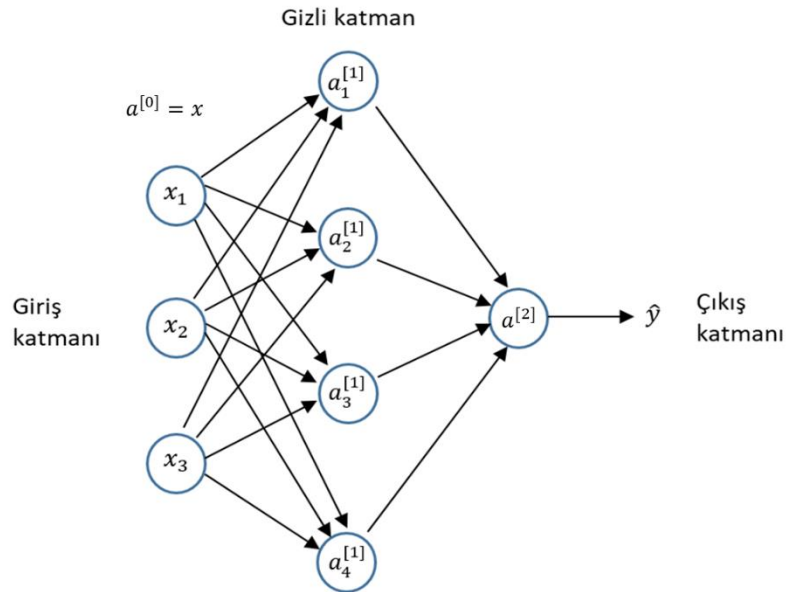
$$g(z) = \max(0, z) \text{ (ReLU)} \quad (4.11)$$

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} (\tanh) \quad (4.12)$$

$$g(z) = \frac{1}{1 + e^{-z}} (\text{sigmoid}) \quad (4.13)$$

$$g(z) = z (\text{identity}) \quad (4.14)$$

ANN bu şekilde birçok nöronun birleşmesinden ve katmanlardan meydana gelir. Yapay sinir ağlarında ilk katman girdi katmanı son katman ise çıktı katmanı olarak adlandırılır. Girdi veya çıktı katmanı olmayan aradaki tüm katmanlara ise gizli katman adı verilir. Gizli katman dışındaki değerler, eğitim kümesinde gözlemlenebilen veya çıktı olarak elde edilebilen değerler iken bu katmandaki değerler gözlemlenememektedir. Bu katmana gizli katman adı verilmesinin sebebi budur. Şekil 4.12’de örnek bir ANN gösterilmektedir.



Şekil 4.12. ANN yapısı

Burada  $a_i^{[j]}$ ,  $j$ . katmandaki  $i$ . nöronun aktivasyonunu ifade etmektedir.

Şekildeki ANN tarafından gerçekleştirilen işlemler vektör formunda şu şekilde ifade edilebilir:

$$z^{[1]} = W^{[1]}a^{[0]} + b^{[1]} \quad (4.11)$$

$$a^{[1]} = g(z^{[1]}) \quad (4.12)$$

$$z^{[2]} = W^{[2]}a^{[1]} + b^{[2]} \quad (4.13)$$

$$a^{[2]} = g(z^{[2]}) = \hat{y} \quad (4.14)$$

Bu şekilde katmanlarda ilerleyerek çıktının hesaplanması sürecine ileri yayılım adı verilir. Yapay sinir ağlarında, ileri yayılım sonucunda elde edilen çıktı değerleri (tahminler) ile gerçek çıktılar arasındaki hatanın minimize edilmesi için geri yayılım algoritması kullanılarak hatayı minimize edecek ağırlık matrisi ve yanlılık vektörü parametreleri hesaplanır. Burada hata fonksiyonu şu şekilde tanımlanır:

$$L(\hat{y}, y) = -(y \log \hat{y} + (1 - y) \log(1 - \hat{y})) \quad (4.15)$$

Bu durumda  $m$  satırdan oluşan bütün veri kümesi için minimize edilmesi gereken maliyet fonksiyonu şu şekilde tanımlanır:

$$J(W^{[1]}, b^{[1]}, W^{[2]}, b^{[2]}) = \frac{1}{m} \sum_{i=1}^m L(\hat{y}^{(i)}, y^{(i)}) \quad (4.16)$$

Şekil 4.12'deki ANN için geri yayılım algoritması şu adımları içermektedir:

$$dz^{[2]} = a^{[2]} - y \quad (4.17)$$

$$dW^{[2]} = dz^{[2]}a^{[1]T} \quad (4.18)$$

$$db^{[2]} = dz^{[2]} \quad (4.19)$$

$$dz^{[1]} = W^{[2]T} dz^{[2]} * g^{[1]'}(z^{[1]}) \quad (4.20)$$

$$dW^{[1]} = dz^{[1]}x^T \quad (4.21)$$

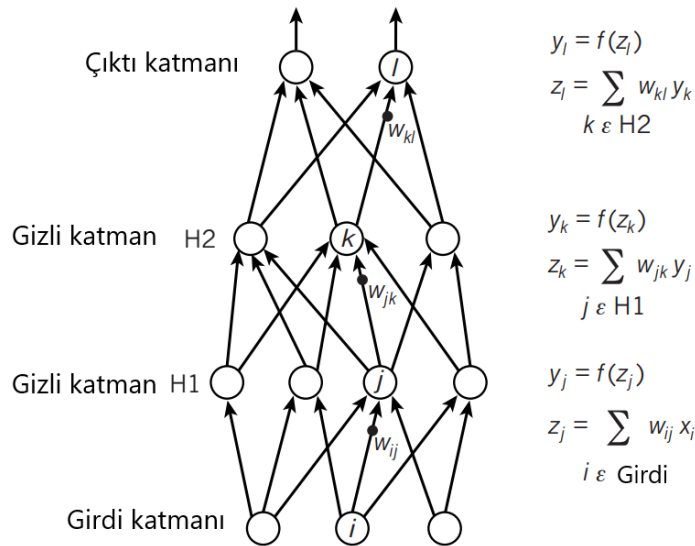
$$db^{[1]} = dz^{[1]} \quad (4.22)$$

#### 4.4.6. Derin öğrenme

Makine öğrenmesinde amaç girdilerin faydalı temsillerini öğrenerek çıktıyı tahmin edebilmektir. Burada temsilden kasıt aynı verinin farklı şekillerde ifade edilmesidir. Örneğin renkli bir görüntüyü ele aldığımızda aynı görüntü hem RGB (red-green-blue) hem de HSV (hue-saturation-value) formatlarında ifade edilebilir (Chollet, 2018). Bazı problemler için RGB formatı daha uygun iken bazıları için HSV daha uygundur. Dolayısıyla, ML yöntemleri girdilerin ilgili probleme uygun temsillerini bulmaya çalışır.

Geleneksel ML yöntemleri özellikle ham verileri işleme kabiliyeti açısından sınırlı kalmaktadır. Uzun bir süre boyunca, ML sistemlerinin oluşturulmasında, ham verileri uygun bir temsile dönüştürebilecek bir öznetelik çıkarma işlemi için çok dikkatli bir mühendislik sürecine ve ciddi bir alan bilgisine ihtiyaç duyulmuştur (LeCunn, Bengio and Hinton, 2015). Derin öğrenme yöntemleri ise, basit ancak doğrusal olmayan modüllerin birçok katmanda bir araya gelmesiyle, ham veriden başlamak üzere verinin her katmanda daha soyut bir temsilini oluşturmayı amaçlayan yöntemlerdir. Şekil 4.13'te girdi katmanı, iki gizli katman ve çıktı katmanından meydana gelen örnek bir derin sinir ağı (DNN) gösterilmektedir.

İlk kez Dechter tarafından 1986 yılında kullanılan *derin öğrenme* teriminden (Dechter, 1986) ANN bağlamında ilk olarak 2000 yılında Aizenberg tarafından bahsedilmiştir (Aizenberg, I. N., Aizenberg, N. N. and Vandewalle, 2000). Günümüzde derin öğrenme yöntemleri görüntü tanıma (Krizhevsky, Sutskever and Hinton, 2012), ses tanıma (Hinton vd., 2012) ve çeviri (Sutskever, Vinyals and Le, 2014) gibi birçok alanlarda en başarılı yöntemlerdendir.



Şekil 4.13. DNN Yapısı

## 5. SONUÇLAR

Bu bölümde ETA tahmini ve yörünge tahmini için rassal orman ve derin yapay sinir ağları yöntemleri kullanılarak ML modellerinin eğitiminden ve elde edilen makine öğrenmesi modellerinin performans sonuçlarından bahsedilmiştir.

### 5.1. ETA Tahmini

FR24 ve allft+ verilerinin eşleştirilmesinden ve hatalı/eksik verilerin ayıklama işlemlerinden sonra veri kümesinde 51.902 uçuşa ait öznitelik değerleri elde edilmiştir. ETA tahmini için kullanılan öznitelikler Tablo 5.1’de gösterilmektedir. Uçakların gerçek iniş zamanları etiket olarak kullanılmıştır.

ML modellerinin eğitimi ve testi için RF ve DNN yöntemleri kullanılmıştır. Bu yöntemler, R programlama dilinde h2o paketi kullanılarak geliştirilmiştir. h2o açık kaynak kodlu, dağıtık ve ölçeklenebilir bir ML platformudur ([http-14](http://h2o.ai)).

**Tablo 5.1.** *Öznitelikler*

| Öznitelik                             | Veri Tipi              |
|---------------------------------------|------------------------|
| Planlanan Kalkış Zamanı (STD)         | Nümerik                |
| Gerçekleşen Kalkış Zamanı (ATD)       | Nümerik                |
| Planlanan İniş Zamanı (STA)           | Nümerik                |
| Toplam Mesafe                         | Nümerik (Türetilmiş)   |
| TMA Girişinden Piste Olan Mesafe      | Nümerik (Türetilmiş)   |
| TMA Giriş Zamanı                      | Nümerik                |
| Havayolu Firması                      | Kategorik              |
| Uçak Tipi                             | Kategorik (Türetilmiş) |
| İniş Pisti                            | Kategorik (Türetilmiş) |
| Havalimanı Civarındaki Rüzgâr Hızı    | Nümerik                |
| Havalimanı Civarındaki Rüzgâr Yönü    | Nümerik                |
| Havalimanı Civarındaki Sıcaklık       | Nümerik                |
| Havalimanı Civarındaki Görüş Mesafesi | Nümerik                |

Veriler %80 eğitim ve %20 test verisi olacak şekilde rastgele bölünmüştür. Deneme yanılma yöntemiyle ML modellerinin parametreleri için farklı değerler denenmiş ve modellerin bu farklı parametre değerleriyle performansları ölçülmüştür. Denemelerde en iyi başarıyı gösteren parametre değerleri RF için ayırma özneliği sayısı = 6 ve ağaç sayısı = 200 şeklinde, DNN için gizli katman sayısı = 2, epoch sayısı = 100 ve gizli katmanlardaki nöron sayısı = 40 şeklinde belirlenmiştir. Gizli katmanlarda ReLU (Denklem 4.11), çıkış katmanında ise Identity (Denklem 4.14) aktivasyon fonksiyonları kullanılmıştır. Modellerin eğitiminde 2.80 GHz işlemciye ve 6GB hafızaya sahip 4 çekirdekli bir bilgisayar kullanılmış ve eğitim 3 dakikadan daha kısa sürmüştür.

ETA tahmin modellerinin performansları MAE, MSE ve RMSE ile ölçülmüştür:

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |\hat{y}_i - y_i| \quad (5.1)$$

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2 \quad (5.2)$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2} \quad (5.3)$$

Tablo 5.2 ve Tablo 5.3 sırasıyla kalkıştan hemen sonra ve TMA girişinde RF ve DNN yöntemleri için performans değerlerini göstermektedir.

**Tablo 5.2.** Kalkış sonrası ETA tahmini (LTBA)

|             | RF    | DNN   |
|-------------|-------|-------|
| <b>MAE</b>  | 5.68  | 5.91  |
| <b>MSE</b>  | 66.40 | 68.14 |
| <b>RMSE</b> | 8.14  | 8.25  |

**Tablo 5.3.** TMA giriři sonrası ETA tahmini (LTBA)

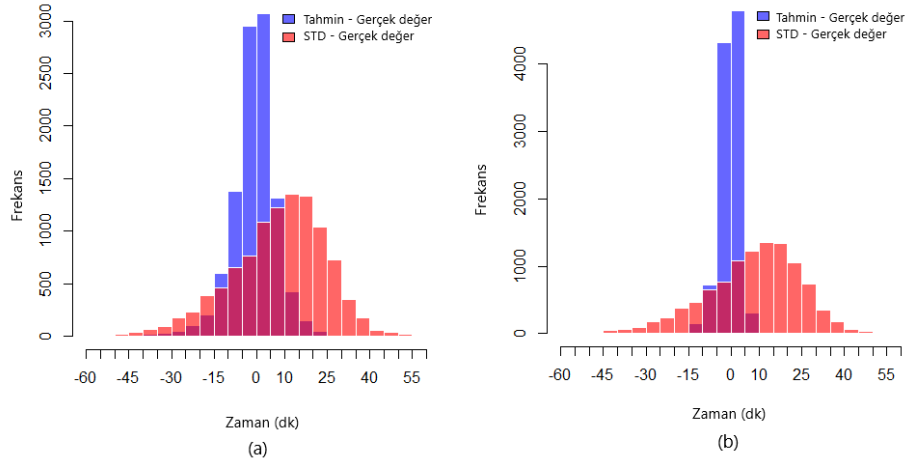
|      | RF    | DNN   |
|------|-------|-------|
| MAE  | 2.60  | 2.66  |
| MSE  | 13.35 | 12.36 |
| RMSE | 3.65  | 3.52  |

Tablo 5.2 ve Tablo 5.3 incelendiğinde her iki ML yönteminde de ETA'nın, TMA giriřinden itibaren 3 dk'nın altında, kalkıştan hemen sonra ise 6 dk'nın altında bir MAE ile tahmin edilebildiği görölmektedir.

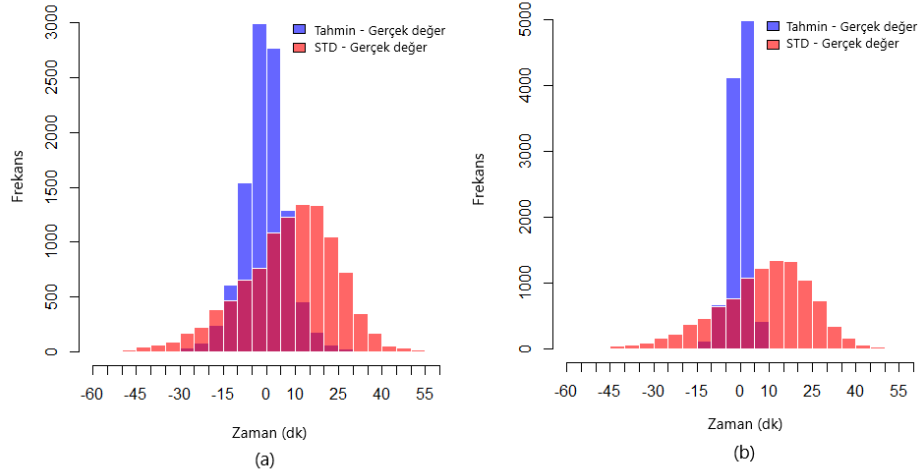
Her iki ML yöntemi için elde edilen modellerde, özniteliklerin bağıl önem dereceleri incelenmiştir. Buna göre kalkış sonrası tahmin için kullanılan RF modelinde, en önemli öznitelikler sırasıyla planlanan varış zamanı (1.0), gerçek kalkış zamanı (0.36) ve planlanan kalkış zamanıdır (0.11). TMA'ya giriř sonrası tahmin için kullanılan RF modelinde ise en önemli öznitelikler sırasıyla TMA'ya giriř zamanı (1.0), planlanan varış zamanı (0.5) ve gerçek kalkış zamanıdır (0.16). Bu durum insan karar verme mekanizmasıyla benzer şekilde çalışan RF için beklenen bir durumdur. Çünkü uçağın kalkış zamanı veya TMA'ya giriř zamanı gibi öznitelikler iniř zamanı için de belirleyici olacaktır.

DNN modelinde RF modelinden farklı olarak birçok özniteliğin bağıl önem derecesinin yüksek olduđu gözlemlenmiştir. Kalkış sonrası tahmin için kullanılan RF modelinde, en önemli öznitelikler sırasıyla gerçekleşen kalkış zamanı (1.0), planlanan varış zamanı (0.88), planlanan kalkış zamanı (0.71), hava yolu firması (THY, 0.54), uçak tipi (M, 0.54) ve toplam mesafedir (0.43). TMA'ya giriř sonrası tahmin için kullanılan DNN modelinde ise en önemli öznitelikler sırasıyla TMA'ya giriř zamanı (1.0), hava yolu firması (THY, 0.39), uçak tipi (M, 0.35), planlanan varış zamanı (0.31) ve gerçek kalkış zamanıdır (0.29).

Ayrıca her iki yöntem için, test verileri kullanılarak STD ve ATA ile ML yöntemleriyle elde edilen ETA ve ATA arasındaki hata histogramları elde edilmiştir. Şekil 5.1 ve Şekil 5.2 RF ve DNN yöntemleri için ETA tahmininin hata histogramlarını göstermektedir. Şekil 5.1 ve Şekil 5.2 incelendiğinde tahmin hatalarının 0 civarında yoğunlaştığı görölmektedir.



**Şekil 5.1.** RF hata histogramları (a) Kalkıştan sonra (b) TMA'ya girişten sonra



**Şekil 5.2.** DNN hata histogramları (a) Kalkıştan sonra (b) TMA'ya girişten sonra

ETA tahmin modellerinin planlanan varış zamanına göre tahmin hatalarını kalkıştan itibaren %64 oranında, TMA'ya girişten itibaren ise %84 oranında azalttığı tespit edilmiştir.

Elde edilen ML modelleri “h2o MOJO” formatında kaydedilerek geliştirilen uygulamaya entegre edilmiş ve öznitelik değerleri girilerek ETA tahminlerinin elde edilebileceği bir ekran geliştirilmiştir (Şekil 5.3).

The screenshot shows a web-based interface for predicting Estimated Time of Arrival (ETA). It features several input fields: 'AIRLINE' (Turkish Airlines), 'DEPARTURE AIRPORT' (Sarajevo International Airport), 'AIRCRAFT' (Airbus A318), and 'ARRIVAL RUNWAY' (05). Below these are sections for 'SCHEDULED TIME OF DEPARTURE' (14.05.2018, 19:10) and 'ACTUAL TIME OF DEPARTURE' (14.05.2018, 19:23). The 'SCHEDULED TIME OF ARRIVAL' is 14.05.2018, 21:05. A 'METAR' field contains the text 'LTBA 142350Z 08003KT 060V170 CAVOK 19/11 Q1016 NOSIG'. Two prediction buttons, 'Random Forests' and 'Deep Learning', are at the bottom. The results show 'ETA-Random Forests' and 'ETA-Deep Learning' both predicting '20:53 +0300'.

**Şekil 5.3.** ETA tahmin ekranı

DNN MOJO modelinin boyutu 250 KB civarında iken RF MOJO modelinin boyutu 90 MB civarındadır. Her iki model için tahminler 5 sn'den kısa bir sürede elde edilebilmektedir.

Geliştirilen modellerde kalkış/iniş havaalanları yerine uçuş mesafeleri dikkate alındığından modeller ilave bir düzeltme gerekmez. Benzer pist konfigürasyonlarına sahip havaalanlarında da kullanılabilir. Modeller İstanbul Sabiha Gökçen Uluslararası Havalimanı'nda (ISG) kalkış sonrası ETA tahmini için test edilmiş ve her iki model için 9 dk'dan daha küçük MAE değerleri elde edilmiştir (Tablo 5.4).

**Tablo 5.4.** Kalkış sonrası ETA tahmini (İSG)

|      | RF     | DNN    |
|------|--------|--------|
| MAE  | 8.83   | 8.36   |
| MSE  | 188.33 | 128.82 |
| RMSE | 13.72  | 11.35  |

## 5.2. Yörünge Tahmini

Çalışma kapsamında elde edilen veriler incelendiğinde uçakların %90'a yakın bir bölümünün 05 pistine iniş yaptığı gözlemlenmiştir. Dolayısıyla yörünge tahmininde 05 pistine iniş yapan uçakların yörünge verilerinden faydalanılmıştır. Bütün standart geliş rotaları için yörüngeler son yaklaşımda iç içe girdiğinden her bir STAR için ayrı bir ML modelinin geliştirilmesi daha uygundur. Bunun için DEKEK1A örnek olarak seçilmiş ve bu STAR üzerinde uygulama yapılmıştır. Yörünge tahmininde kullanılan öznitelikler Tablo

5.5'te gösterilmektedir. Çıktı olarak ise her bir yörünge noktasından bir sonraki yörünge noktasının enlem, boylam, irtifa ve hız verileri kullanılmıştır.

**Tablo 5.5. Öz nitelikler**

| Öznitelik | Veri Tipi |
|-----------|-----------|
| Enlem     | Nümerik   |
| Boylam    | Nümerik   |
| İrtifa    | Nümerik   |
| Hız       | Nümerik   |
| Baş Açısı | Nümerik   |

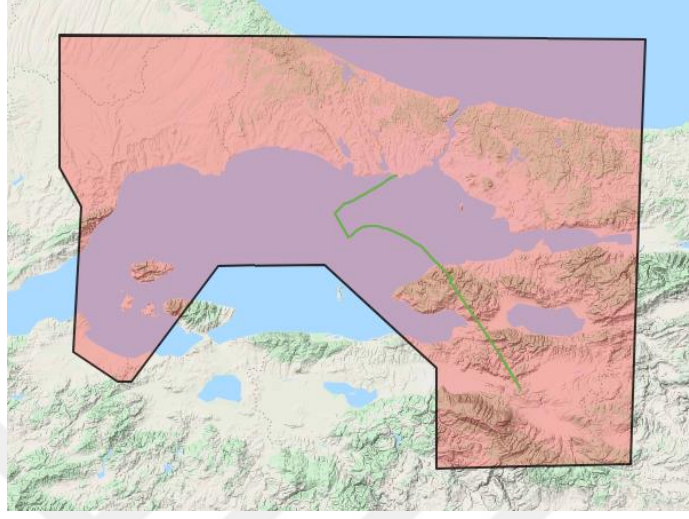
01.06.2018-30.06.2018 tarihleri arasında DEKEK1A standart geliş rotasından 05 pistine inen uçakların yörünge verilerinden eğitim verisi oluşturulmuş, bu verilerden 187.509 adet yörünge noktası elde edilmiştir. 30.04.2018 tarihli uçuşlara karşılık gelen 6.573 adet yörünge noktası test kümesi olarak kullanılmıştır.

Derin öğrenme yöntemi DeepLearning4J Java kütüphanesi ([http-15](http://15)) kullanılarak geliştirilmiş ve uygulamaya entegre edilmiştir. DeepLearning4J, Java ve Scala programlama dilleri için geliştirilen açık kaynak kodlu bir derin öğrenme kütüphanesidir.

Geliştirilen DNN bir girdi katmanı, iki gizli katman ve bir çıktı katmanından meydana gelmektedir. Çıktı katmanının aktivasyon fonksiyonu Identity (Denklem 4.14), diğer katmanların aktivasyon fonksiyonları ise ReLU (Denklem 4.11) olarak seçilmiştir. Her bir gizli katmandaki nöron sayısı 20, eğitim verisi üzerinden kaç tekrar geçileceğini ifade eden epoch değeri ise 1000 olarak seçilmiştir. Modellerin eğitiminde 2.80 GHz işlemciye ve 6GB hafızaya sahip 4 çekirdekli bir bilgisayar kullanılmış ve eğitim 10 dakikadan daha kısa sürmüştür.

Modelin performansı MAE, MSE ve belirlilik katsayısı ( $R^2$ ) ile ölçülmüştür. 0-1 aralığında değer alan  $R^2$ , tahmin edilen regresyon eğrisinin dağılıma uyum derecesi şeklinde düşünülebilir (II Beysolow, 2017). Değer 1'e yaklaştıkça uyum artmaktadır.

$$R^2 = 1 - \sum_{i=1}^m \frac{(\hat{y}_i - y)^2}{(\hat{y}_i - \bar{y})^2} \quad (5.4)$$



**Şekil 5.4.** Bir uçuş için elde edilen yörünge tahmini

Yörünge tahmini için performans ölçütleri Tablo 5.6’da gösterilmiştir.

**Tablo 5.6.** Yörünge tahmini için performans ölçütleri

|               | MAE   | MSE   | $R^2$ |
|---------------|-------|-------|-------|
| <b>Enlem</b>  | 0.117 | 0.035 | 0.97  |
| <b>Boylam</b> | 0.107 | 0.022 | 0.97  |
| <b>İrtifa</b> | 0.141 | 0.04  | 0.96  |
| <b>Hız</b>    | 0.141 | 0.036 | 0.96  |

Tablo 5.6 incelendiğinde özellikle düşük MAE ve MSE değerleriyle birlikte yüksek belirlilik katsayısı ( $R^2$ ) değeri modelin yörünge tahmininde başarılı olduğu anlamına gelmektedir. Şekil 5.4’te DEKEK1A STAR’ını izleyerek iniş yapan bir uçağın tahmini yörüngesi, elde edilen ML modeli kullanılarak çizdirilmiştir. Bu yörünge Şekil 3.2’de DEKEK giriş noktasında başlayıp BA588 noktasında sonlanan DEKEK1A STAR’ıyla uyumu görsel olarak da doğrulanmıştır. Ayrıca yörünge son yaklaşmada 05 pistine doğru yöneldiği ve 05 pistine yakın bir noktada sonlandığı da görülmektedir.

## 6. TARTIŞMA

Dünya genelinde gelecek yıllarda hava trafiğinin artışıyla birlikte ortaya çıkacak problemlerin çözümünde, dijital dönüşümle ve sistem genelinde veri paylaşımı ile birlikte veri odaklı bir yapıya geçileceği, AI ve ML uygulamalarının ATM’de birçok alanda kullanılacağı açıktır.

Bu çalışmada gecikme yönetiminde, hava sahası kapasitesinin ve verimliliğinin artırılmasında önemli bir rolü olan ETA tahmin doğruluğunun, özellikle kapasite problemlerinden en olumsuz şekilde etkilenen TMA’lar içerisinde iyileştirilmesine yönelik bir ML uygulaması geliştirilmiştir. Geliştirilen ML modelleri incelendiğinde ETA tahmininin, planlanan varış zamanına göre kalkıştan itibaren %64, TMA’ya girişten itibaren ise %84 oranında iyileştirildiği gözlemlenmiştir. ETA tahmin modeli, hava trafiğiyle ilgili diğer veriler de elde edilerek (kalkış uçakları, hava sahası ve havaalanı yoğunluğu, yüksek irtifa rüzgâr bilgileri vb.) ve farklı yöntemler denenerek daha da iyileştirilebilir.

Bu çalışmada TMA içerisinde uçakların izleyeceği yörüngelerin tahmini de gerçekleştirilmiştir. DMAN/AMAN, çatışma tespiti vb. ATM süreçlerinin birçoğu yörünge tahminine dayalıdır. Ayrıca yörünge verilerinin ATM sisteminin dayandığı temel bilgi olacağı TBO konseptinde de yörünge tahmini önemlidir. Bu anlamda geliştirilen yörünge tahmini modelleri, ETA tahmin modelleriyle entegre edilerek veri güdümlü bir AMAN geliştirilebilir. Bu durumda, uçağın TMA içerisinde herhangi bir noktada hangi zamanda olacağı da belirlenebilir. Böylece hava trafik kontrolörlerinin uçuş sıralamalarını dinamik olarak görüntüleyebileceği bir karar destek sistemi geliştirilebileceği gibi eldeki verilerden gecikme tahmini, hava sahası ve havaalanı yoğunluğu gibi tahminler de gerçekleştirilerek geliştirilen uygulamaya entegre edilebilir.

Elde edilen veriler, ilgili dönemde Türkiye’nin en yoğun havalimanı olan İstanbul Atatürk Havalimanı’na ait gerçek veriler olduğundan, sadece ML modellerinin geliştirilmesinde değil, takviyeli öğrenme yöntemleriyle geliştirilecek sistemler aracılığıyla hava trafik kontrolörlerinin veya Hava Trafik Kontrol Bölümü öğrencilerinin gerçek senaryolarla eğitiminde de kullanılabilir.

Bu çalışmanın, EU için stratejik öneme sahip olan AI ve ML yöntemlerinin ATM’de uygulama örneklerinden biri olarak, özellikle ülkemizde havacılık sektörü açısından da öncü çalışmalardan biri olacağı tahmin edilmektedir.

Gelecek çalışmalarda, geliştirilen ETA tahmini ve yörünge tahmini modellerinin, yüksek irtifa rüzgâr ve sıcaklık verileri, kalkış trafiği, TMA ve havaalanı yoğunluğu gibi ilave veriler elde edilerek iyileştirilmesi planlanmaktadır. İyileştirilen bu modeller birbirleriyle entegre edilerek hava trafik kontrolörlerinin kullanabileceği veri güdümlü bir karar destek sistemi geliştirilmesi, geliştirilecek sistemin TMA içerisinde uçakların sıralanmasında kontrolörlere yardımcı olması hedeflenmektedir.

Bunlara ilave olarak, elde edilen gerçek uçuş verileri kullanılarak makine öğrenmesinin farklı bir dalı olan takviyeli öğrenme konularında da çalışmalar yapılması, özellikle hava trafik kontrolörlerinin eğitiminde gerçek senaryolarla kontrolörlerin uygulamalarının kıyaslanabileceği bir değerlendirme ve öneri sisteminin geliştirilmesi düşünülmektedir.

## KAYNAKÇA

Aizenberg, I. N., Aizenberg, N. N. and Vandewalle, J. (2000). *Multiple-Valued Threshold Logic and Multi-Valued Neurons*. Springer.

Alpaydın, E. (2010). *Introduction to Machine Learning* (2<sup>nd</sup> Ed.). Cambridge: *The MIT Press*.

Ayhan, S. and Samet, H. (2016). Aircraft Trajectory Prediction Made Easy with Predictive Analytics. *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD'16)*., California, 21-30.

Ayhan, S., Costas, P. and Samet, H. (2018). Predicting Estimated Time of Arrival for Commercial Flights. *Proc. 24th Int. Conf. Knowledge Discovery & Data Mining (KDD'18)*., London, 33-42.

Basora, L., Morio, J. and Mailhot, C. (2017). A Trajectory Clustering Framework to Analyse Air Traffic Flows. *Seventh SESAR Innovation Days (SIDs 2017)*, Belgrade.

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

Breiman, L. (1996). Bagging Predictors. *Machine Learning*, 24, 123-140.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.

Breiman, L., Friedman, J., Stone, C. J. and Olshen, R. A. (1984). Classification and Regression Trees. *Taylor & Francis*.

Chollet, F. (2018). *Deep Learning with Python*. New York: Manning Publications Co.

Cook, A. (2007). *European Air Traffic Management*. Hampshire: Ashgate Publishing.

Cook, A., Tanner, G. and Lawes, A. (2007). Dynamic Cost Indexing. *6th EUROCONTROL Innovative Research Workshop & Exhibition*, France.

Çetek, F. A. (2015). *Terminal Kontrol Sahalarında Dinamik Kapı Ataması için Simülasyon Modeli Önerisi*. Doktora Tezi. Eskişehir: Anadolu Üniversitesi, Fen Bilimleri Enstitüsü.

Dechter, R. (1986). Learning while Searching in Constraint-Satisfaction-Problems. *Proc. AAAI-86*, 178-183.

DHMI Strateji Geliştirme Daire Başkanlığı (2020). *2019 Havayolu Sektör Raporu*. Ankara.

EUROCONTROL (2018). *Challenges of Growth 2018: European Aviation in 2040*.

EUROCONTROL (2020). *COVID 19 Impact on European Air Traffic: EUROCONTROL Comprehensive Assessment*.

European Aviation Artificial Intelligence High Level Group (2020). *The FLY AI Report: Demystifying and Accelerating AI in Aviation/ATM*.

Fernandez, E.C., Cordero, J.M., Vouros, G., Pelekis, N., Kravaris, T., Georgiou, H., Fuchs, G., Andrienko, N., Andrienko, G., Casado, E., Scarlatti, D. and Costas, P. (2017). DART: A Machine-Learning Approach to Trajectory Prediction and Demand-Capacity Balancing. *Seventh SESAR Innovation Days*, Belgrade.

Ghatak, A. (2017). *Machine Learning with R*. Singapore: Springer.

Glina, Y., Jordan, R. and Ishutkina, M. (2012). A Tree-based Ensemble Method for Prediction and Uncertainty Quantification of Aircraft Landing Times. *Proc. 10th Conf. Artificial and Computational Intelligence and its Applications to Environmental Sciences*, New Orleans.

Hinton, G. *et al.* (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition. *IEEE Signal Processing Magazine*, 29, 82–97.

Horst, A. and Schindler, M. (2012). *Project E.02.12 COMPASS, D2.2 Data Analysis*.

IATA, EUROCONTROL and CANSO (2008). *Flight Efficiency Plan: Fuel and Emissions Savings*.

ICAO (2005). *Annex 2: Rules of Air* (10<sup>th</sup> Ed.). ICAO Publications.

ICAO (2018a). *Annex 11: Air Traffic Services*. ICAO Publications.

ICAO (2018b). *Annex 14: Aerodromes-Volume I-Aerodromes Design and Operations* (8<sup>th</sup> Ed.). ICAO Publications.

ICAO (2020). *Effects of Novel Coronavirus (COVID-19) on Civil Aviation: Economic Impact Analysis*. Canada.

II Beysolow, T. (2017). *Introduction to Deep Learning Using R*. California: Appress.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2017). *An Introduction to Statistical Learning with Applications in R* (8<sup>th</sup> Ed.). London: Springer.

Kern, C.S., Medeiros, I.P. and Yoneyama, T. (2015). Data-Driven Aircraft Estimated Time of Arrival Prediction. *Proc. IEEE Systems Conference (SysCon)*, Vancouver, 727-733.

Krizhevsky, A., Sutskever, I. and Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Proc. Advances in Neural Information Processing Systems*, 25, 1090–1098.

LeCunn, Y., Bengio, Y. and Hinton, G. (2015). Deep Learning. *Nature*, 521, 436-444.

Leege, A.M.P., Paassen, M.M. and Mulder, M. (2013). A Machine Learning Approach to Trajectory Prediction. *AIAA Guidance, Navigation, and Control (GNC) Conf.*, Boston.

Ma, L., and Tian, S. (2020). A Hybrid CNN-LSTM Model for Aircraft 4D Trajectory Prediction. *IEEE Access*, 8, 134668-134680.

Meinshausen, N. (2006). Quantile Regression Forests. *J. Machine Learning Research*, 7(35), 983-999.

Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill.

Mondoloni, S. (2006). Aircraft Trajectory Prediction Errors: Including a Summary of Error Sources and Data. *FAA/Eurocontrol Action Plan 16*.

Murphy, J. R., Clayton, J., Reisman, R. J. and Wright, R. (2003). Physics-Based and Parametric Trajectory Prediction Performance Comparison for Traffic Flow Management. *AIAA Guidance, Navigation, and Control Conf. and Exhibition*. Texas.

Roe, A. W., Pallas, S. L., Kwon, Y. H. and Sur, M. (1992). Visual Projections Routed to the Auditory Pathway in Ferrets: Receptive Fields of Visual Neurons in Primary Auditory Cortex. *J. Neuroscience*, 12(9), 3651-3664.

SESAR (2012). *The Roadmap for Sustainable Air Traffic Management: European ATM Master Plan*.

SESAR Joint Undertaking (2020). *European ATM Master Plan Executive View*. Luxembourg: Publications Office of the European Union.

Shi, Z., Xu, M., Pan, Q., Yan, B. and Zhang, H. (2018). LSTM-based Flight Trajectory Prediction. *Int. Joint Conf. Neural Networks (IJCNN)*, Rio de Janeiro.

Slattery, R. and Zhao, Y. (1997). Trajectory Synthesis for Air Traffic Automation. *J. Guidance, Control and Dynamics*, 20(2), 232-238.

Sutskever, I. Vinyals, O. and Le. Q. V. (2014). Sequence to Sequence Learning with Neural Networks. *Proc. Advances in Neural Information Processing Systems*, 27, 3104–3112.

Sutton, R. S. and Barto, A. G. (2015). *Reinforcement Learning: An Introduction* (2<sup>nd</sup> Edt.). Cambridge: The MIT Press.

Takacs, G. (2014). Predicting Flight Arrival Times with a Multistage Model. *IEEE Int. Conf. Big Data*, Washington, 78-84.

Wang, L., Li, X. and Mao, J. (2020). Integrating ARIMA and Bidirectional LSTM to Predict ETA in Multi-Airport Systems. *Integrated Communications Navigation and Surveillance Conf. (ICNS)*, Herndon, USA.

Wang, Z., Liang, M. and Delahaye, D. (2017). Short-term 4D Trajectory Prediction Using Machine Learning Methods. *Seventh SESAR Innovation Days (SIDs 2017)*, Belgrade.

Wang, Z., Liang, M. and Delahaye, D. (2020). Automated Data-Driven Prediction on Aircraft Estimated Time of Arrival. *J. Air Transport Management*, 88.

Zeng, W., Quan, Z., Zhao, Z., Xie, C. and Lu, X. (2020). A Deep Learning Approach for Aircraft Trajectory Prediction in Terminal Airspace. *IEEE Access*, 8, 151250-151266.

Zheng, A. and Casari, A. (2018). *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. California: O'Reilly.

**http-1:** <https://www.airlines.org/dataset/world-airlines-traffic-and-capacity/#> (Eriřim tarihi: 20.12.2020)

**http-2:** <https://www.airlines.org/dataset/world-airlines-traffic-and-capacity/#> (Eriřim tarihi: 20.12.2020)

**http-3:** <https://www.dhmi.gov.tr/Sayfalar/Istatistikler.aspx> (Eriřim tarihi: 20.12.2020)

**http-4:** <https://www.flightradar24.com> (Eriřim tarihi: 20.12.2020)

**http-5:** <https://ext.eurocontrol.int/ddr/> (Eriřim tarihi: 20.12.2020)

**http-6:** <https://mesonet.agron.iastate.edu/> (Eriřim tarihi: 20.12.2020)

**http-7:** <https://developers.google.com/kml/> (Eriřim tarihi: 20.12.2020)

**http-8:** <https://hezarfen.mgm.gov.tr/KodA/BasMetar.aspx> (Eriřim tarihi: 20.12.2020)

**http-9:** <https://app.pluralsight.com/library/courses/r-understanding-machine-learning/table-of-contents> (Eriřim tarihi: 20.12.2020)

**http-10:** <https://www.coursera.org/learn/machine-learning/home/welcome> (Erişim Tarihi: 20.12.2020)

**http-11:** [https://tr.wikipedia.org/wiki/%C4%B0nsan\\_ekolokasyonu](https://tr.wikipedia.org/wiki/%C4%B0nsan_ekolokasyonu) (Erişim Tarihi: 20.12.2020)

**http-12:** <https://www.youtube.com/watch?v=WHYCs8xtzUI> (Erişim Tarihi: 20.12.2020)

**http-13:** <https://www.coursera.org/learn/neural-networks-deep-learning/home/welcome> (Erişim Tarihi: 20.12.2020)

**http-14:** <https://www.h2o.ai/> (Erişim Tarihi: 20.12.2020)

**http-15:** <https://deeplearning4j.org/> (Erişim Tarihi: 20.12.2020)

## ÖZGEÇMİŞ

Adı Soyadı : Onur BAŞTÜRK

Yabancı Dil : İngilizce (KPDS: 92)

### Eğitim ve Mesleki Geçmişi:

- 2021 - Şimdi, Bilişim Personeli, Adalet Bakanlığı Bilgi İşlem Genel Müdürlüğü, Araştırma Geliştirme Şubesi
- 2010 – 2021, Öğretim Görevlisi, Anadolu Üniversitesi, Bilgisayar Araştırma ve Uygulama Merkezi
- 2006 – 2010, Araştırma Görevlisi, Cumhuriyet Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü

### Yayınları ve/veya Bilimsel/Sanatsal Faaliyetleri:

- Akbaba, H., Baştürk, O. ve Kantarcı, A. G. (2017). “Development of a Device for Regulating Magnetic Field with Electromagnets for the in-vitro Delivery of Nucleic Acid-Associated Magnetic Nanoparticles”, *2017 CRS Annual Meeting and Exposition*
- Namazov, M. ve Baştürk, O. (2010). “Stable and Optimal Controller Design for Takagi-Sugeno Fuzzy Model Based Control Systems via Linear Matrix Inequalities”, *Int. Symp. Innovations in Int. Sys. and App. (INISTA 2010)*, Kayseri.
- Namazov, M. ve Baştürk, O. (2009). “Fuzzy Proportional Derivative Controller Design for Position Control of a DC Motor”, *1st Int. Fuzzy Sys. Symp. (FUZZYSS'09)*, Ankara.

### Mesleki Birlik/Dernek/Kuruluş Üyelikleri:

- 2012 – Şimdi, Institute of Electrical and Electronics Engineers (IEEE)
- 2006 – 2010, Elektrik Mühendisleri Odası (EMO)