

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



ARAÇ CAN BUS VERİLERİ KULLANILARAK EĞİTİCİLİ
SINIFLANDIRMA ALGORİTMALARI İLE SÜRÜŞ STİLİ VE YOL
PROFİLİNİN SAPTANMASI

Berat KARABULUTER

Yüksek Lisans Tezi

YAZILIM MÜHENDİSLİĞİ ANABİLİM DALI

Yazılım Mühendisliği Bilim Dalı

TEMMUZ 2024

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Yazılım Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

**ARAÇ CAN BUS VERİLERİ KULLANILARAK EĞİTİCİLİ
SINIFLANDIRMA ALGORİTMALARI İLE SÜRÜŞ STİLİ VE YOL
PROFİLİNİN SAPTANMASI**

Tez Yazarı

Berat KARABULUTER

Danışman

Dr. Öğr. Üyesi Özgür KARADUMAN

TEMMUZ 2024

ELAZIĞ

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Yazılım Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Başlığı:	Araç CAN Bus Verileri Kullanılarak Eğitici Sınıflandırma Algoritmaları ile Sürüş Stili ve Yol Profilinin Saptanması
Yazarı:	Berat KARABULUTER
İlk Teslim Tarihi:	07.06.2024
Savunma Tarihi:	10.07.2024

TEZ ONAYI

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına göre hazırlanan bu tez aşağıda imzaları bulunan jüri üyeleri tarafından değerlendirilmiş ve akademik dinleyicilere açık yapılan savunma sonucunda OYBİRLİĞİ ile kabul edilmiştir.

Danışman:	Dr. Öğr. Üyesi Özgür KARADUMAN Fırat Üniversitesi, Mühendislik Fakültesi	<i>İmza</i> Onayladım
Başkan:	Doç. Dr. Bihter DAŞ Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım
Üye:	Doç. Dr. Harun BİNGÖL Malatya Turgut Özal Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi	Onayladım

Bu tez, Enstitü Yönetim Kurulunun/...../20..... tarihli toplantısında tescillenmiştir.

İmza

Prof. Dr. Burhan ERGEN
Enstitü Müdürü

BEYAN

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım ‘‘Araç CAN Bus Verileri Kullanılarak Eğitici Sınıflandırma Algoritmaları ile Sürüş Stili ve Yol Profilinin Saptanması’’ Başlıklı Yüksek Lisans Tezimin içindeki bütün bilgilerin doğru olduğunu, bilgilerin üretilmesi ve sunulmasında bilimsel etik kurallarına uygun davrandığımı, kullandığım bütün kaynakları atıf yaparak belirttiğimi, maddi ve manevi desteği olan tüm kurum/kuruluş ve kişileri belirttiğimi, burada sunduğum veri ve bilgileri unvan almak amacıyla daha önce hiçbir şekilde kullanmadığımı beyan ederim.

10.07.2024

Berat KARABULUTER



ÖNSÖZ

Günümüzde araç sayısı arttıkça sürücülerden ve yol profilinden kaynaklanan trafik kazalarındaki artış paralelinde can ve mal kayıplarının da azımsanamaz derecede artmasına neden olmuştur. Bu ana etkenden yola çıkarak literatürdeki pek çok çalışma kişilerin araç sürüş tarzları ve yol profillerinin belirlenmesi ve böylece sürücü ve yol profili kaynaklı oluşabilecek kazaların önüne geçilebilmesi üzerine yoğunlaşmaktadır.

Yapmış olduğum bu çalışmadaki genel amaç, araçlarda bulunan bazı sensör verilerinin CAN Bus yoluyla alınarak sürücülerin sürüş tarzının yanı sıra araçların sürüldüğü yol profilinde belirlemeyi amaçlayan bir sınıflandırma sistemi oluşturmaktır. Elde edilen sonuçlar, trafik güvenliği ve kazaların önlenmesi konusunda karar vericilere ve sürücülere değerli bilgiler sağlayabilir. Ayrıca ileriye dönük olarak sürüş tarzı analizi ve yol profillemesi gibi alanlarda daha fazla araştırma ve geliştirme yapılmasına da katkıda bulunabileceği öngörülmüştür.

Çalışmam süresince bana olan inancını hep sürdürüp hayatımın her alanında yanımda olan sevgili eşim Nuray KARABULUTER'e, kızımız İpek Nil KARABULUTER'e, oğlumuz Vedat KARABULUTER'e ve canım annem Mürüvvet KARABULUTER'e teşekkür ederim.

Ayrıca çalışmam süresince bana vakit ayırıp, fikirleriyle çalışmama yön veren ve yardımlarını esirgemeyen çok değerli danışmanım Dr. Öğr. Üyesi Özgür KARADUMAN'a saygılarımı sunar ve teşekkürü bir borç bilirim.

Berat KARABULUTER

ELAZIĞ, 2024

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	iv
İÇİNDEKİLER	v
ÖZET	vii
ABSTRACT	viii
ŞEKİLLER LİSTESİ	ix
TABLolar LİSTESİ	xi
SİMGELER VE KISALTMALAR	xii
1. GİRİŞ	1
1.1. Problem Tanımı	2
1.2. İlgili Çalışmalar	3
1.3. Tezin Amacı	5
2. SÜRÜŞ STİLLERİ VE YOL PROFİLLERİ.....	6
2.1. Agresif Sürüş Stili	7
2.1.1. Tanımı ve Davranışlar	7
2.1.2. Psikolojik ve Sosyal Etkenler	7
2.1.3. Sonuçları ve Riskleri	7
2.2. Sakin Sürüş Stili	8
2.2.1. Tanımı ve Davranışlar	8
2.2.2. Psikolojik ve Sosyal Etkenler	8
2.2.3. Sonuçları ve Faydaları.....	8
2.3. Düz Yol Profili	9
2.4. Engeli (Bozuk) Yol Profili.....	10
3. ARAÇ BAĞLANTISI VE VERİ OKUMA.....	12
3.1. On-Board Diagnostics (OBD-2).....	12
3.1.1. J1939 Bağlantı Konektörü.....	12
3.1.2. SAE J1962 Bağlantı Konektörü	14
3.2. Controller Area Network (CANBus).....	16
4. MATERYAL VE METOT	19
4.1. Verilerin Toplanması	20
4.2. Verilerin Normalizasyonu	21
4.3. Sınıflandırma Algoritmalarına Sürüş Stilleri ve Yol Profillerinin Öğretilmesi	22
4.3.1. Yapay Sinir Ağları (YSA).....	24
4.3.2. Navie Bayes	28
4.3.3. K En Yakın Komşu (KNN).....	28
4.3.4. Destek Vektör Makineleri (SVM).....	30
4.3.5. C4.5 (J48) Algoritması	32
4.4. Derin Öğrenme Algoritmasıyla Sürüş Stilleri ve Yol Profillerinin Öğretilmesi.....	33
4.4.1. Uzun-Kısa Süreli Hafıza Ağları (LSTM)	34
4.5. Verilerin Doğrulanması	38
5. BULGULAR VE TARTIŞMA	41
5.1. CANBus Üzerinden Verilerin Toplanması.....	41

5.2. Veri Normalizasyonu.....	45
5.3. Yapay Sinir Ağı Üzerinde Öğrenme, Doğrulama ve Test Adımları	46
5.4. Derin Öğrenme	49
5.5. YSA, KNN, Navie Bayes, SVM, C4.5, Derin Öğrenme Modellerinin Başarıları	52
5.6. Doğrulama İşlemleri	54
6. SONUÇLAR.....	58
ÖNERİLER	59
KAYNAKLAR.....	60
ÖZGEÇMİŞ	



ÖZET

Araç CAN Bus Verileri Kullanılarak Eğitici Sınıflandırma Algoritmaları ile Sürüş Stili ve Yol Profiline Saptanması

Berat KARABULUTER

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Yazılım Mühendisliği Anabilim Dalı

Temmuz 2024, Sayfa: xii + 62

Gün geçtikçe kullanılan araçların sayısının artmasıyla araç sürücüleri, sürüş stilleri ve yolların profili kaynaklı kazaları daha fazla yaşamaktadırlar. Artan kazaların paralelinde de kişilerin can ve mal kayıplarında büyük oranda artış yaşandığı resmi kayıtlara geçmektedir. Durum böyle olunca akıllı trafik sistemi üzerine çalışan araştırmacıları, bu oluşabilecek kaza oranlarının azaltılması veya kaza durumuna yol açabilecek nedenlerin sürücülere önceden iletilmesi için yeni çalışmalar yapmaya itmiştir.

Bu tezin amacı, araçların CANBus üzerinden sensör verileri kullanılarak sürücünün sürüş stiline tespiti ile araçların sürüldüğü yolların profiline bulunmasıdır. Bu çalışmada, Yapay Sinir Ağları, En Yakın Komşuluk, Destek Vektör Makineleri, C4.5 Algoritması, Naive Bayes ve Derin Öğrenme algoritması olan LSTM gibi hem geleneksel olarak literatürde kabul edilmiş en çok kullanılan sınıflandırıcı algoritmaların yanında günümüz çağının başarı performansı en yüksek olan ve Derin Öğrenme sınıflandırma algoritması kullanılmıştır. Bu algoritmalar, sürücünün agresif/sakin sürüş stili ile araçların kullanıldığı düz/bozuk yol profillerini tespit etmek için araç OBD II soketinden CANBus yoluyla alınan verileri kullanmıştır.

Çalışmadaki veriler farklı sürücülerin bilinçli bir şekilde literatürde kabul edilen agresif ve sakın sürüş tarzlarında araç sürerek araçtan alınan CAN Bus sensör verileri "agresif" ve "sakın" şeklinde etiketlenmiş eğitim verileri haline getirilmiştir. Ayrıca bozuk ve düz yollarda yapılan çeşitli sürüşlerden toplanan sensör verileri de yol profili için eğitim verileri olarak kullanılmıştır. Bu tezde, derin öğrenme yöntemleri de dahil olmak üzere sınıflandırma algoritmalarının kullanılarak sürüş stili ve yol profili yüksek doğruluk oranında tespit edilebileceği gösterilmiştir. Sonuç olarak, sürüş stili ve yol profili sınıflandırmanın LSTM modelinin diğerlerine göre daha başarılı olduğu tespit edilmiştir.

Elde edilen sonuçlarla oluşabilecek kazaların engellenmesi amacıyla gerek sürücülere ve gerekse yetkililere geri bildirimler sağlayacağı öngörülmektedir.

Anahtar Kelimeler: Sürüş stili, Yol profili, Sınıflandırma Algoritmaları, Derin öğrenme, CAN Bus, OBDII

ABSTRACT

Detection of Driving Style and Road Profile with Trainer Classification Algorithms Using Vehicle CAN Bus Data

Berat KARABULUTER

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences
Department of Software Engineering

July 2024, Pages: xii + 62

As the number of vehicles used increases day by day, drivers are experiencing more accidents due to their driving styles and road profiles. Official records show that in parallel with the increasing number of accidents, there has been a significant increase in the loss of life and property of people. This situation has led researchers working on smart traffic systems to conduct new studies to reduce the accident rates that may occur or to inform drivers in advance of the reasons that may lead to an accident.

The aim of this thesis is to find the profile of the roads on which the vehicles are driven by detecting the driver's driving style by using sensor data via CANBus. In this study, the most used classifier algorithms traditionally accepted in the literature, such as Artificial Neural Networks, Nearest Neighborhood, Support Vector Machines, C4.5 Algorithm, Naive Bayes and LSTM, which is the Deep Learning algorithm, as well as the ones with the highest success performance of today's age and the Deep Learning algorithm, are used. Learning classification algorithm was used. These algorithms used data received from the vehicle OBD II socket via CANBus to detect the driver's aggressive/calm driving style and the smooth/rough road profiles on which the vehicles are driven.

The data in the study were collected by different drivers consciously driving in aggressive and calm driving styles accepted in the literature, and the CAN Bus sensor data taken from the vehicle was converted into training data labeled as "aggressive" and "calm". In addition, sensor data collected from various drives on rough and smooth roads were used as training data for the road profile. In this thesis, it has been shown that driving style and road profile can be determined with high accuracy by using classification algorithms, including deep learning methods. As a result, it has been determined that the LSTM model is more successful in classification of driving style and road profile than others.

It is envisaged that the results obtained will provide feedback to both drivers and authorities in order to prevent accidents that may occur.

Keywords: Driving style, Road profile, Classification Algorithms, Deep learning, CAN Bus, OBDII

ŞEKİLLER LİSTESİ

Sayfa

Şekil 1.1.	2016 DSÖ önde gelen ölüm nedenleri (Her Yaştan).	1
Şekil 3.1.	J1939 bağlantı konnetörlerinin 9 pinli modelinin bağlantı pin uçları	13
Şekil 3.2.	J1939 bağlantı konnetörlerinin 9 pinli gerçek görüntüsü.....	13
Şekil 3.3.	SAE J1962 bağlantı konnetörlerinin 16 pinli modelinin bağlantı pin uçları.....	14
Şekil 3.4.	SAE J1962 bağlantı konnetörlerinin 16 pinli modelinin gerçek görüntüsü	15
Şekil 3.5.	Bus ağ topoloji ile bağlı terminallerin bağlantı şekli	16
Şekil 3.6.	CANBus üzerinde taşınan veri paketi ve paket içeriğindeki verilerin boyutları.....	17
Şekil 4.1.	Sistemin çalışma adımları	19
Şekil 4.2.	Sürüş stili etiketlerini gösterir grafik	23
Şekil 4.3.	Yol profili etiketlerini gösterir grafik.....	23
Şekil 4.4.	Sürüş stili transfer fonksiyonlarının korelasyon değerleri	25
Şekil 4.5.	Yol Profili transfer fonksiyonlarının korelasyon değerleri	26
Şekil 4.6.	Sürüş stili yapay sinir ağı tasarımı	27
Şekil 4.7.	Yol profili yapay sinir ağı tasarımı	27
Şekil 4.8.	KNN sürüş stili sınıflandırması koordinat düzlemi	29
Şekil 4.9.	KNN yol profili sınıflandırması koordinat düzlemi	29
Şekil 4.10.	SVM sürüş stili koordinat düzlemi	31
Şekil 4.11.	SVM yol profili koordinat düzlemi.....	31
Şekil 4.12.	LSTM algoritmasının temel hücre blok diyagram modeli	35
Şekil 4.13.	LSTM modelinin birinci tip çalışma şekli	36
Şekil 4.14.	LSTM modelinin ikinci tip çalışma şekli.....	36
Şekil 4.15.	LSTM modelinin ikinci tip çalışma şekli.....	37
Şekil 4.16.	LSTM modelinin süreç işleyiş diyagramı	38
Şekil 4.17.	Duyarlılık ve keskinlik değerleri arasındaki ilişki	39
Şekil 5.1.	Sürüş stili verisi toplama güzergahı 1	41
Şekil 5.2.	Sürüş stili verisi toplama güzergahı 2	41
Şekil 5.3.	Yol profili verisi toplama güzergahı	42
Şekil 5.4.	Farklı araçlar üzerinden kayıt verilerinin toplanması	44
Şekil 5.5.	Sürüş stili YSA performan değerleri grafikleri	47
Şekil 5.6.	Yol profili YSA performan değerleri grafikleri	48
Şekil 5.7.	Tez için tasarlanan LSTM modelinin iç yapısı ve parametreleri	50

Şekil 5.8.	Sürüş stili eğitim ve doğrulama doğruluğu grafiği	51
Şekil 5.9.	Sürüş stili eğitim ve doğrulama kaybı grafiği	51
Şekil 5.10.	Yol profili eğitim ve doğrulama doğruluğu grafiği.....	52
Şekil 5.11.	Yol profili eğitim ve doğrulama kaybı grafiği	52
Şekil 5.12.	Sürüş stilleri sınıflandırma algoritmaları performans grafiği.....	55
Şekil 5.13.	Yol Profilleri sınıflandırma algoritmaları performans grafiği.....	55
Şekil 5.14.	Sürüş stilleri sınıflandırma algoritmaları test roc eğrisi	56
Şekil 5.15.	Yol Profilleri sınıflandırma algoritmaları test roc eğrisi	57



TABLolar LİSTESİ

Sayfa

Tablo 1.1.	Sürüş stili ve yol profili belirlemede sıkça kullanılan ek donanım araçları karşılaştırması	5
Tablo 2.1.	Agresif ve sakin sürüş stilleri arasındaki karşılaştırmalı ilişki tablosu	9
Tablo 2.2.	Düz ve engebeli yol profilleri arasındaki karşılaştırmalı ilişki tablosu.....	11
Tablo 3.1.	J1939 bağlantı konektörlerinin pin fonksiyonları ve açıklamaları.....	14
Tablo 3.2.	SAE J1962 bağlantı konektörlerinin pin fonksiyonları ve açıklamaları	15
Tablo 3.3.	CANBus üzerinde taşınan veriler ve açıklamaları	17
Tablo 3.4.	CANBus ile çoğu araç üzerinden okunabilecek veriler ve ayrıntıları.....	18
Tablo 4.1.	Seçilen sensör verileri	21
Tablo 4.2.	Normalize edilmiş ve etiketlenmiş veri seti tablosu	22
Tablo 4.3.	Sürüş stili transfer fonksiyonları korelasyon sayısal değerleri.....	25
Tablo 4.4.	Yol Profili transfer fonksiyonları korelasyon sayısal değerleri	26
Tablo 4.5.	MAF Sayısal değerlerin ve One Hot Encoder değer karşılıkları gösterimi.....	37
Tablo 4.6.	Veri sınıflarının sonuç tablosu (Karışıklık Matrisi)	39
Tablo 5.1.	Sürüş stili eğitim veri setlerinin işlenmemiş hali	43
Tablo 5.2.	Yol profili eğitim veri setlerinin işlenmemiş hali	43
Tablo 5.3.	Sürüş stili eğitim veri setlerinin normalize edilmiş hali.....	45
Tablo 5.4.	Yol Profili eğitim veri setlerinin normalize edilmiş hali.....	46
Tablo 5.5.	LTSM'nin çalışma tiplerinin başarılarının karşılaştırılması	49
Tablo 5.6.	Sürüş stili karışıklık değerleri tablosu.....	53
Tablo 5.7.	Yol profili karışıklık değerleri tablosu.....	53
Tablo 5.8.	Sürüş stili model başarı değerleri tablosu	53
Tablo 5.9.	Yol profili model başarı değerleri tablosu	54

SİMGELER VE KISALTMALAR

Simgeler

$H(s)$	: S kaynağının entropisi
h_t	: O andaki gizli durum
r_{xy}	: Korelasyon değeri
X'	: Normalize edilmiş değer

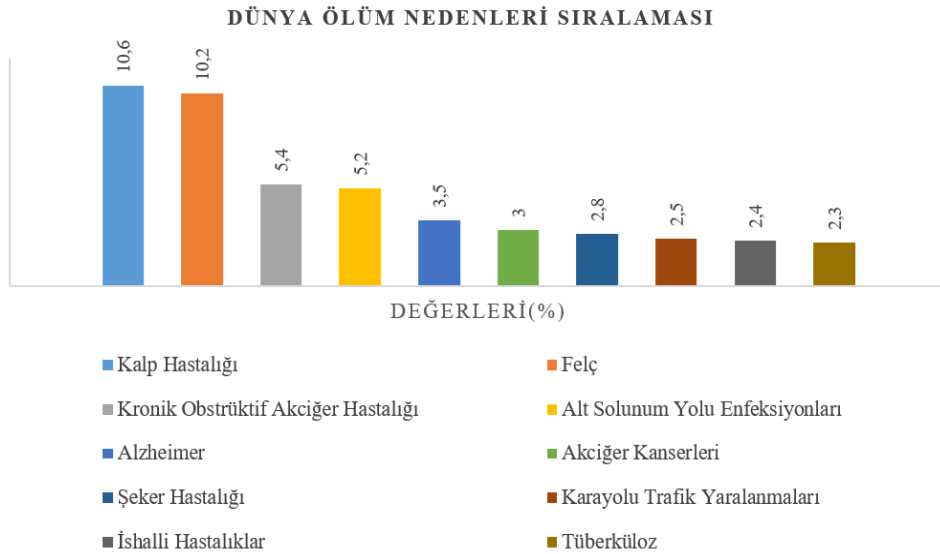
Kısaltmalar

SVM	: Destek Vektör Makineleri
KNN	: K En Yakın Komşu
YSA	: Yapay Sinir Ağı
CanBus	: Controller Area Network
OBD 2	: On-Board Diagnostics
EKU	: Elektronik Kontrol Ünitesi
RNN	: Tekrarlayan Sinir Ağı
LTSM	: Uzun Kısa Süreli Bellek
DSÖ	: Dünya Sağlık Örgütü
PURELİN	: Lineer Transfer Fonksiyonu
TANSİG	: Bipolar Sigmoid Transfer Fonksiyonu
LOGSİG	: Unipolar Sigmoid Transfer Fonksiyonu
MSE	: Mean Squared Error
ROC	: Receiver Operating Characteristic
TP	: True Positive
FN	: False Negative
FP	: False Positive
TN	: True Negative

1. GİRİŞ

Her yıl sürekli olarak yaşanan ölüm, yaralanma ve maddi hasarlara neden olan toplam kaza sayısı yüzbinleri bulmaktadır. Meydana gelen bu kazaların büyük bir çoğunluğu araç sürücülerinin yapmış olduğu hatalar ve araçların sürülmüş oldukları yolların profillerinden kaynaklanmaktadır. Ortaya çıkan bu durum maddi manevi birçok hasara ve kayıplara neden olmakla beraber karşımıza ciddi bir sorun olarak çıkmaktadır. Kara yollarında yaşanan kazaların artışı, trafik güvenliği konusunda daha fazla bilinçlenme, tedbirler alınması ve önceden uyarı sistemleri gibi konularda araştırmacıları çalışmaya teşvik etmiştir.

Ülkemizden her yıl Emniyet Genel Müdürlüğü Trafik Başkanlığı tarafından yayınlanan Karayolu Trafik Kaza İstatistikleri' ne göre 2023 yılında oluşan toplam trafik kazası sayısı 1 milyon 232 bin 957 adet olarak istatistiklere geçmiştir. Bu kazaların 1 milyon 35 bin 696 adedi maddi hasarlı, 197 bin 261 adedi ise ölümlü yaralanmalı trafik kazasıdır [1]. Gene aynı istatistik raporu daha detaylı incelendiğinde ise Türkiye'de son yayınlanan 2023 yılı istatistik raporuna göre ölümlü ve yaralanmalı trafik kazalarına neden olan toplam 235 bin 176 kusura bakıldığında ise kusurların %86,8'inin sürücü, %9,5'inin yaya, %2,1'inin taşıt, %1,2'sinin yolcu ve %0,4'ünün yol kaynaklı olduğu görülmektedir [1]. Bu istatistiklerin yıllara göre dağılım incelendiğinde sürücülerin sürüş stillerinden ve araçların kullandıkları yolların profillerinden kaynaklanan kaza oranı ise %40'ın üzerinde olduğu gözükmemektedir. Dünya Sağlık Örgütü'nün (DSÖ) en son olarak 2018 yılında hazırladığı Karayolu Güvenliğine İlişkin Küresel Durum Raporu'nda dünya üzerinde ölüme neden olan ilk 10 faktör içinde 8. sırada trafik kazaları yaralanmaları yer almakta ve trafik kazaları kaynaklı ölümlerin ise 2030 yılında 5. Sıraya yükselmesi ön görülmektedir [2]. Şekil 1.1' de DSÖ önde gelen ölüm nedenleri sıralaması grafiğine yer verilmiştir.



Şekil 1.1. 2016 DSÖ önde gelen ölüm nedenleri (Her Yaştan) [2].

Bu nedenle, kazaların önlenmesi için ilk adım olarak sürücülerin sürüş stilleri ve araçlarını kullandıkları yolların profilleri hakkında bilgi edinmeleri gereklidir. Sürücülerin bilinçli veya bilinçsiz olarak gerçekleştirdikleri agresif sürüş davranışlarının engellenmesi önemlidir. Ayrıca, yol profillerinin kazalara neden olan faktörlerinin en aza indirilmesi ve sürücülerin ani yol profil değişikliklerinden haberdar olması da önem taşımaktadır.

1.1. Problem Tanımı

Bu tez günlük hayatta kullanmış olduğumuz karayolu araçlarını sürerken kişileri maddi, manevi zarar ve hayati risklere neden olabilecek kazaların engellenmesine katkıda bulunabilmek için iki ana problemin üzerinden yoğunlaşmıştır. Bu problemlerden ilki olan araç sürücülerinin sürüş stilleri, ikincisi ise bu araçların kullanılmış oldukları yolların profillerinin belirlenmesi üzerine kurulmuştur. Bu iki problemin tanımlamalarını sırasıyla açıklayacak olursak;

Günümüzde karayolu ulaşımı için kullanmış olduğumuz araçların büyük çoğunluğu hala insanların kontrolü ile sürülen araçlardan oluşur. Bu nedenle, sürücüler anlık yaşadıkları ruh hallerine göre arabalarını sürerken farklı agresif ve sakin sürüş stilleri sergileyebilirler. Örneğin, arabalarını devirsiz kullanabilir, ani fren yapıp hızlı kalkışlar yapabilir, trafiğin olağan akışına ve hız kurallarına uymadan kontrolsüz bir şekilde şerit değiştirebilir ve gereksiz yere aşırı yakıt tüketimi yapabilirler. Bu durum, kullanılan aracın motor ömrünü, sürücü ve yolcuların can güvenliğini ve yakıt tüketimini olumsuz etkileyebilir. Ayrıca, maddi ve manevi birçok zarara yol açabilir.

Araç kullanıcılarından birçoğu, kullandıkları araçları sürerken kendileri bile farkında olmadan agresif stilde araç sürebilirler. Günümüzde, sürücülere sürüş stilleri üzerinden geri dönüt bilgisi veren sistemler çok sayıda aracı kapsayacak şekilde kullanılmamaktadır. Genellikle bu tip geri bildirim veren araçlar pahalı üst segment araçlardır. Bu nedenle, tutarlı ve daha çok aracı kapsayabilecek bir sürüş stili tahmin sisteminin ortaya konması gerektiği açıktır.

Diğer bir problemimiz olan yol profili belirleme ise sürücülerin araçlarını kullanmış oldukları yolların profillerinin bozuk veya sağlam yol olarak belirleme üzerine kurulmuştur. Araçların kullanıldığı yolların profilleri, kişilerin aynı şekilde can ve mal güvenliğinin yanı sıra oluşabilecek maddi ve manevi hasarların ortaya çıkma nedenleri arasında azımsanamaz etkisi olduğu pek çok kamu, tüzel kuruluş ve organizasyonların belirlediği istatistiklerde ve raporlarda ortaya konmuştur [1,2].

Bilindiği üzere araçların kullanılmış oldukları yolların yatış yapısı bize yol profili hakkında bilgi vermektedir. Sürücülerin kontrolündeki araçlar bozuk ya da sağlam yol profillerine göre farklı hız ve güvenlik önlemleri ile sürülmesi gerekmektedir. Sağlam bir yolda ilerlerken sürüş için hız limiti farklı iken bozuk bir yolda ilerlerken farklıdır. Bazen araçlar sağlam bir yolda giderken aniden bozuk bir yol profilinin içine girebiliyor böylece kazaların önü de açılmış oluyor.

Yukarıda tanımlamaları yapılmış olan iki ana problemin çözümüne katkı sağlayabilecek bu tez çalışması araçlarda bulunan OBD 2 aracılığı ile araçlar arasındaki marka, model ve üretim yılını gibi parametrelerin etkisini en aza indirgeyecek şekilde dahili sensörler üzerinden alınan verilerin harici herhangi bir dış donanım veya sensöre ihtiyaç duymadan çalışması prensibine dayanarak oluşturulmuştur.

1.2. İlgili Çalışmalar

Literatür incelendiğinde 2 ana problemimiz üzerine yapılan pek çok çalışma göze batmakta ve araştırmacıların bu problemlere çözüm olabilecek farklı yöntemler geliştirdikleri görülmektedir. Bu problemlerimizden ilki olan sürüş stili belirlemeye yönelik yapılan akademik yayınlardan başlayacak olursak.

Literatür taramasında göze batan çalışmalardan ilki araç dahili sensörleri dışından GPS sensörü gibi donanım kullanan, g kuvveti ölçer donanımları çalışmalarına ekleyen ve son olarak akıllı telefonlarını sürüş stillerini belirlemede kullananlara rastlamaktayız [3-5]. Sürüş stili gibi bir problemin belirlenmesi üzerine çalışan araştırmacılarda en çok kullanılan dış donanımlar içinde kameraları çok sık görmekteyiz [6-9]. Bunun yanında mobil uygulamalarda sürüş stili belirleme çalışmalarında son zamanlarda aktif olarak görülmeye başlanmıştır [3,12-14,23-25]. Kullanılan dış donanımların farklılığı olduğu gibi verilerin işlenip kararlı sonuçların çıkarılması amacıyla sınıflandırma algoritmaları veya bulanık mantık üzerine çalışmaları da uygulayanları görmekteyiz [5,10-11,15-17]. Görüldüğü gibi çalışmaların büyük çoğunluğu ek bir donanım kullanarak sürüş stili belirleme yoluna giden çalışmacılar çok sık rastlanmaktadır. Kimi çalışmalara baktığımızda ise kullanılan araçlar üzerinde bulunan sensör verilerinden herhangi bir değer almadan da sürüş stili belirleme yöntemine giden çalışmalarda bulunmaktadır [18-20]. Göze batan bazı çalışmalar ise simülasyon üzerinden araştırmacıların kendi belirledi oldukları agresif ve sakin sürüş stil kriterlerine göre tespitler yapmışlardır [26-28]. Gene literatür taramalarında sürüş stili tespitinde emisyon değerleri üzerindeki farklılıklardan yola çıkarak çıkarımlar yapan çalışmalarda dikkat çekmektedir [21-22].

Günümüzde derin öğrenme algoritmalarının ortaya çıkması ve bu algoritmaların doğru yanıtlar verme oranının gün geçtikçe artması sürüş stili belirlemede kullanılmasının önünü açmıştır. Derin öğrenme algoritmalarının sürüş stillerine uyarlanması özellikle çok sayıda GPS tabanlı yapılan çalışmalar göze batmaktadır. Düzenli okunan GPS koordinat verileri yoğun veri tabanları oluşturması ve bu dataların işlenebilmesi için veri performansı yüksek algoritmalara ihtiyaç duyurmuştur. Derin öğrenme algoritmaları da burada devreye girerek toplana büyük dataların anlamlandırılmasında kullanılmıştır. Kimi çalışmalarda, GPS verileri kullanılarak sürüş stilleri karakterize edilmektedir. GPS verileri istatistiksel özellik matrislerine dönüştürülerek, zaman serisi verileri ile sürücü davranışları öğrenilmektedir. Bu yöntem, sürücülerin agresif hızlanma gibi ince

sürüş davranışlarını tespit etmek için kullanılmaktadır [29]. Başka bir derin öğrenme algoritmaları kullanılan GPS tabanlı yapılan sürüş stili belirleme çalışması ise; sürücülerin GPS verilerine dayalı sürüş davranışlarını analiz etmek için derin öğrenme yöntemlerini kullanarak tehlikeli sürüş davranışlarını sınıflandırmak için etkili ve sağlam sınıflandırıcılar geliştirilmiştir [30]. Derin öğrenme kullanılarak yapılan ve farklı veriler üzerinden çalışan bir değer çalışmada da; araç sensörlerinde ivme ve yakıt tüketimi verilerinden elde edilen sürekli veri akışları üzerinde makine ve derin öğrenme yöntemlerini kullanarak sürüş davranışlarını analiz etmektedir. Çalışmada, çeşitli algoritmaların performansları doğruluk, F1 skoru ve ROC eğrisi gibi metriklerle değerlendirilmiş başarıları kabul edilir sonuçlar elde edilmiştir [31].

Bir diğer problemimiz olan yol profili belirleme, son yıllarda üzerine çalışmaların artarak devam ettiği alanlardan biridir. Sürücülerin araçlarını sürdükleri yollarda oluşmuş olan bozukluklar, oluşturabileceği maddi ve manevi hasarlar düşünüldüğünde çalışmalarında bu alana yönelmesine ön ayak olmuştur. Literatür taramasın yapıldığında yol profili üzerine yapılan çalışmalardan kamera kullanarak yapılan eğitici sınıflandırma algoritmalarını sık sık rastlanmaktadır. Kimi bilişim dünyası dışındaki inşaat mühendisliği alanında da yol profili belirleme üzerine yapılan çalışmalarda araçlarda kamera kullanılarak SVM ve diğer eğitici öğrenme algoritmalarını kullanarak yol yüzeyi bozulmalarını tahmin etme üzerine odaklanmaktadır [32]. Yol profilinde oluşan bozuklukların belirlenmesi için derin öğrenme algoritmaları kullanarak yol görüntüleme ve çukur tespitini ele alan ve özellikle 3D yol veri modelleme ve segmentasyon teknikleri üzerine odaklanan çalışmalar da bulunmaktadır [33]. Bazı derin öğrenme çalışmalarında ise yol profilinin ve araç süspansiyon sistemine ait durum değişkenlerinin tahmin edilmesi için derin öğrenme yöntemlerini incelenmektedir. Geleneksel model tabanlı yöntemlerin aksine, veriye dayalı derin öğrenme yaklaşımları kullanılarak yüksek doğruluk ve hızlı hesaplama ile encoder - decoder yapısına sahip, iki aşamalı dikkat mekanizması içeren derin tekrarlayan sinir ağı (RNN) kullanılmıştır [34].

İncelediğimiz çalışmalarda, araç CANBus verilerinin yanında ek sensörler veya donanımlar kullanılarak sürüş stili ve yol profillerinin tespit edildiği görülmektedir. Ayrıca, sadece harici bir donanım üzerinden sürüş stili ve yol profillerinin tespitini belirleme yöntemlerin de mevcut olduğunu görmekteyiz. Bu çalışmaların içinde geleneksel kabul görmüş eğitici öğrenme algoritmalarının yanı sıra derin öğrenme algoritmalarının da son zamanlarda çalışmalarda bolca yer aldığı ortaya çıkmaktadır.

Sürüş stili veya yol profili belirlemede kullanılan ek donanım araçlarının işlevleri ile kullanım alanları arasındaki ilişkileri gösterilmiş hali Tablo 1.1' de verilmiştir.

Tablo 1.1. Sürüş stili ve yol profili belirlemede sıkça kullanılan ek donanım araçları karşılaştırması

Donanım	İşlevi	Sürüş Stili	Yol Profili
Kamera	Görüntü hareket algılama	✓	✓
Gyroskop	Dönme hareketi algılama	✓	
İvmeölçer	İvme değişimi algılama	✓	✓
Gps	Konum belirleme	✓	✓
Ses izleme sistemleri	Ses ve konuşma analizi	✓	
Lidar	Çevresel nesne algılama		✓
Biometrik Sensörler	Fizyolojik durum Algılama	✓	
Termal Kameralar	Sıcaklık değişimi algılama		✓

1.3. Tezin Amacı

Bu tez çalışmasında, araçtan okunan CANBus verileri kullanılarak agresif ve sakin sürüş stillerinin yanı sıra bozuk veya sağlam yol profillerinin belirlenmesi hedeflenmiştir. Bu sistemi tanımlarken, ek kamera veya sensör gibi harici donanımlara ihtiyaç duyulmadan, sadece aracın dahili sensörlerinden elde edilen en az sayıda ilişkili veriler kullanılarak hızlı ve etkili bir analiz yapılması amaçlanmıştır. Analiz sürecinde, literatürde kabul görmüş ve başarı oranı en yüksek eğitici sınıflandırma algoritmaları ve derin öğrenme modeli kullanılmıştır. Bu algoritmalar aracılığıyla yapılan testler sonucunda, yüksek doğruluğa sahip sınıflandırmalar elde edilmiştir. Sonuç olarak, bu çalışmada, ek donanım gereksinimi olmaksızın, sadece CANBus verileri ile sürüş stilleri ve yol profillerinin başarılı bir şekilde belirlenebileceği gösterilmiştir. Sistemin doğruluğunu artırmak için çeşitli testler gerçekleştirilmiş ve elde edilen sınıflandırmaların doğruluğu yüksek seviyelerde çıkmıştır.

2. SÜRÜŞ STİLLERİ VE YOL PROFİLLERİ

Sürüş stili, literatürde ilk kez 1993 yılında Elander, West ve French tarafından tanımlanmıştır. Bu tanım, bir sürücünün araç kullanma alışkanlıklarını ve yol üzerindeki davranışlarını kapsamaktadır. Araştırmacılar, sürüş stili'nin hızlanma, frenleme, dönüş yapma ve genel yol tutumu gibi çeşitli unsurları içerdiğini belirtmişlerdir. Bu çalışmalar, sürüş stili'nin bireysel sürücü alışkanlıkları, araç dinamikleri ve çevresel faktörlerin bir kombinasyonu olarak nasıl ortaya çıktığını ve bu farklılıkların trafik kazası riskleri üzerindeki etkilerini incelemiştir. Elander, West ve French'in çalışması, sürüş stili'ni anlamak için psikolojik ve davranışsal faktörleri birleştirmiştir. Bu yaklaşım, sürüş davranışlarının yalnızca araç teknik özellikleri veya yol koşullarıyla değil, aynı zamanda sürücünün kişilik özellikleri ve psikolojik durumlarıyla da şekillendiğini ortaya koymuştur. Çalışmalarında, farklı sürüş stillerinin yakıt tüketimi, araç güvenliği ve yolculuk süresi gibi önemli etkilere sahip olduğunu vurgulamışlardır [35].

Sürüş stili'nin analizi, sürücülerin risk alma eğilimleri, dikkat düzeyleri ve stres yönetimi gibi psikolojik faktörlerle de ilişkilendirilmiştir. Örneğin, agresif sürücüler genellikle yüksek hızlarda seyretme, ani frenlemeler yapma ve diğer sürücülerle daha fazla etkileşime girme eğilimindedir. Buna karşılık, daha sakin sürücüler yavaş ve istikrarlı bir hızda seyretmeyi, ani hareketlerden kaçınmayı ve trafik kurallarına daha fazla uymayı tercih ederler. Bu farklılıklar, sürüş stili'nin güvenlik ve verimlilik üzerindeki etkilerini anlamak için kritik öneme sahiptir.

Ayrıca bu araştırmalar, sürüş stili'nin bireysel farklılıkları nasıl yansıttığını ve bu farklılıkların sürücü eğitimi ve trafik güvenliği politikaları üzerinde nasıl etkili olabileceğini de incelemiştir. Sürücü eğitimi programları, farklı sürüş stillerini dikkate alarak bireylerin daha güvenli ve verimli sürüş alışkanlıkları geliştirmelerine yardımcı olmayı amaçlamaktadır. Bu bağlamda, sürüş stili'nin literatürdeki tanımlanması ve incelenmesi, trafik güvenliğini artırmaya yönelik stratejilerin geliştirilmesine katkı sağlamaktadır [35].

Yol profili kavramı, modern literatürde yol yüzeyinin pürüzlülüğünü ve düzgünlüğünü analiz etmek amacıyla ortaya çıkmıştır. Yol profili çalışmaları, özellikle mühendislik ve ulaşım bilimlerinde, yol güvenliği ve konforunu artırmak amacıyla önemli bir yer tutar. Yol profili analizinin kökenleri, yol yapımı ve bakımının hayati önem taşıdığı Roma İmparatorluğu dönemine kadar uzanabilir. Roma yolları, taş döşeme ve derin yol yatakları gibi yeniliklerle inşa edilmiştir, bu da yolların dayanıklılığını ve kullanım ömrünü artırmıştır [36].

Modern yol profili analizleri, 20. yüzyılın ortalarından itibaren daha sistematik hale gelmiştir. 1980'lerde yol profili ölçümleri ve analizleri üzerine yapılan çalışmalar, teknolojinin gelişimi ile birlikte hız kazanmıştır. Yol Profili Kullanıcıları Grubu, bu alandaki bilgi alışverişini ve yenilikleri desteklemek amacıyla kurulmuş ve yol pürüzlülüğü, düzgünlük, yüzey dokuları ve lastik-yol etkileşimi gibi konular üzerinde çalışmalara odaklanmıştır [37].

Yol profili analizleri, günümüzde gelişmiş sensörler ve veri toplama teknikleri kullanılarak gerçekleştirilmekte ve bu analizler yol güvenliği, sürüş konforu ve araç performansının iyileştirilmesine önemli katkılar sağlamaktadır. Yol profili literatürünün gelişimi, hem tarihsel kökleri hem de modern teknolojik yenilikleri içeren geniş bir yelpazeyi kapsamaktadır.

2.1. Agresif Sürüş Stili

Agresif sürüş, bir sürücünün trafik kurallarına aykırı ve diğer yol kullanıcıları için tehlike yaratacak şekilde araç kullanma davranışlarını tanımlayan bir kavramdır. Bu tür sürüş, genellikle aşırı hız, ani frenlemeler, sık sık şerit değiştirme, diğer araçları yakın takip etme, sinyal vermeden dönüş yapma ve trafik işaretlerine uymama gibi davranışları içerir. Agresif sürüşün temel özellikleri ve etkileri aşağıda detaylandırılmaktadır;

2.1.1. Tanımı ve Davranışlar

Aşırı Hız: Trafik limitlerini aşarak yüksek hızda araç kullanma.

Ani Frenleme: Gerekli olmayan ani ve sert frenler yapma.

Yakın Takip: Önündeki araçla güvenli mesafeyi korumadan, çok yakın mesafede sürüş yapma.

Agresif Şerit Değiştirme: Sürekli olarak ve sinyal vermeden ani şerit değiştirme.

Diğer Sürücülere Yönelik Düşmanca Davranışlar: Kornaya aşırı şekilde basma, el hareketleri yapma veya diğer sürücülerini taciz etme.

2.1.2. Psikolojik ve Sosyal Etkenler

Stres ve Öfke: Sürücüler genellikle stresli veya öfkeli olduklarında daha agresif davranışlar sergileyebilirler.

Rekabetçilik: Bazı sürücüler, trafikte diğer sürücülerle rekabet etme eğiliminde olabilirler.

Zaman Baskısı: Geç kalma korkusu veya hızlı bir şekilde varış noktasına ulaşma isteği agresif sürüşü tetikleyebilir.

2.1.3. Sonuçları ve Riskleri

Kaza Riski: Agresif sürüş, trafik kazalarına neden olma olasılığını önemli ölçüde artırır.

Yüksek Yakıt Tüketimi: Sürekli hızlanma ve yavaşlama, aracın yakıt verimliliğini düşürür.

Yasal Sonuçlar: Trafik kurallarını ihlal eden sürücüler, para cezaları, ehliyetin askıya alınması gibi yasal yaptırımlarla karşılaşabilirler.

Agresif sürüşün anlaşılması ve bu tür davranışların önlenmesi, trafik güvenliğini artırmak, kazaları azaltıp maddi ve manevi hasarları minimuma indirmek için önemlidir. Tez çalışmasının

en büyük amaçlarından biride sürücülere sürüşleri sırasında elde edilen araçlarındaki sensör verileriyle sürüş stilleri hakkında geri dönütler vermektir.

2.2. Sakin Sürüş Stili

Sakin sürüş, bir sürücünün araç kullanma esnasında güvenli, sorumlu ve öngörülebilir davranışlar sergilemesi olarak tanımlanır. Bu tür sürüş, trafik kurallarına tam uyum, yol ve hava koşullarına uygun hızda araç kullanma, diğer yol kullanıcılarına saygılı davranma ve genel olarak trafik ortamında dikkatli ve sabırlı olma gibi özellikler içerir. Sakin sürüşün özellikleri ve faydaları aşağıda detaylandırılmaktadır.

2.2.1. Tanımı ve Davranışlar

Hız Sınırlarına Uyum: Trafik işaretleri ve hız limitlerine uymak, hız sınırlarını aşmamak.
Yumuşak Hızlanma ve Frenleme: Ani hızlanma ve frenlemeden kaçınmak, yumuşak ve kontrollü bir şekilde hızlanıp yavaşlamak.
Güvenli Takip Mesafesi: Önündeki araçla güvenli bir takip mesafesi bırakmak, ani frenlemelere karşı yeterli reaksiyon süresi bırakmak.
Trafik Kurallarına Saygı: Trafik ışıkları, dur işaretleri ve diğer trafik kurallarına tam uyum göstermek.
Dikkatli ve Öngörülebilir Sürüş: Yola ve diğer sürücülere karşı dikkatli olmak, sinyal kullanarak öngörülebilir sürüş manevraları yapmak.

2.2.2. Psikolojik ve Sosyal Etkenler

Duygusal Kontrol: Stresli veya öfkeli olduğunda araç kullanmaktan kaçınmak, duygularını kontrol altında tutmak.
Sabır ve Hoşgörü: Trafikte diğer sürücülere karşı sabırlı ve hoşgörülü olmak, agresif davranışlardan kaçınmak.
Odaklanma ve Dikkat: Yol boyunca dikkatini yola ve trafik koşullarına vermek, dikkat dağınıcı unsurlardan kaçınmak.

2.2.3. Sonuçları ve Faydaları

Daha Az Kaza Riski: Sakin sürüş, trafik kazası riskini önemli ölçüde azaltır. Güvenli sürüş davranışları, sürücünün ve yolcuların güvenliğini artırır.
Yakıt Verimliliği: Yumuşak hızlanma ve frenleme, yakıt tüketimini optimize eder ve aracın verimliliğini artırır.

Uzun Araç Ömrü: Araç üzerinde daha az mekanik stres oluşur, bu da aracın bakım masraflarını azaltır ve ömrünü uzatır.

Daha Az Stres ve Rahat Yolculuk: Sakin sürüş, sürücü için daha az stresli ve daha rahat bir sürüş deneyimi sağlar.

Aşağıda gösterilen karşılaştırmalı Tablo 2.1’ de, agresif ve sakın sürüş tarzlarının farklı yönlerini vurgulamakta ve her iki sürüş tarzının güvenlik, yakıt verimliliği, araç aşınması ve genel sürüş deneyimi üzerindeki etkilerini karşılaştırmalı olarak ortaya koymaktadır [35,38].

Tablo 2.1. Agresif ve sakın sürüş stilleri arasındaki karşılaştırmalı ilişki tablosu [35,38]

Özellik	Agresif Sürüş	Sakin Sürüş
Hız	Yüksek hız, hız sınırlarını aşma eğilimi	Trafik kurallarına uygun hız
Frenleme	Ani ve sert frenlemeler	Yumuşak ve kontrollü frenleme
Şerit Değiştirme	Sık ve sinyal vermeden şerit değiştirme	Gerektiğinde, sinyal vererek şerit değiştirme
Takip Mesafesi	Diğer araçlara yakın takip etme	Güvenli takip mesafesini koruma
Davranışsal Eğilimler	Stres, öfke, rekabetçilik	Sakinlik, sabır, hoşgörü
Yakıt Tüketimi	Yüksek yakıt tüketimi (ani hızlanma ve frenlemeler nedeniyle)	Daha düşük yakıt tüketimi (düzenli hızlanma ve frenleme)
Araç Aşınması	Artan mekanik aşınma ve bakım gereksinimi	Azalan aşınma, daha düşük bakım maliyetleri
Güvenlik	Artan kaza riski	Azalan kaza riski
Yasal Sonuçlar	Trafik cezaları, ehliyetin askıya alınması	Trafik kurallarına uyum, yasal sorunların azalması
Sürüş Konforu	Azalmış sürüş konforu, artan yolcu şikayetleri	Artmış sürüş konforu, daha az stresli yolculuk
Çevresel Etkiler	Yüksek emisyonlar	Daha düşük emisyonlar

2.3. Düz Yol Profili

Düz yol profili, yüzeyinin büyük ölçüde pürüzsüz ve düzgün olduğu, minimal eğim ve yüzey dalgalanmalarının bulunduğu yol türlerini ifade eder. Bu tür yollar, genellikle yüksek hızda seyahat eden araçlar için güvenli ve konforlu bir sürüş deneyimi sunar. Düz yol profilleri, genellikle yoğun trafik akışının olduğu ve yüksek hız limitlerinin bulunduğu alanlarda tercih edilir. Şehir içi ana yollar, otopanlar ve otoyollar bu tür yol profillerine örnek olarak verilebilir.

Düz yol profilinin özelliklerini aşağıdaki gibi sıralayabilir;

Pürüzsüz Yüzey: Düz yollar, minimum yüzey pürüzlülüğüne sahiptir, bu da araçların daha az titreşim ve sarsıntı ile hareket etmesini sağlar.

Düşük Eğim: Bu yollar genellikle düz veya çok az eğimli olup, araçların sabit hızlarda güvenle seyretmesine imkan tanır.

Yüksek Sürüş Konforu: Düz yol profilleri, sürüş konforunu artırır ve sürücü yorgunluğunu azaltır. Bu özellikler, uzun mesafe seyahatlerinde önemli bir avantaj sağlar.

Bu yolların sürücülere sağlamış olduğu avantajları ise aşağıdaki gibi sıralayabiliriz;

Güvenlik: Düz yollar, ani manevraları ve kazaları önlemek için daha güvenlidir.

Yakıt Verimliliği: Düzgün bir yol yüzeyi, araçların daha az yakıt tüketmesine yardımcı olur.

Düşük Bakım Maliyeti: Pürüzsüz yüzeyler, araç bileşenlerinin daha az aşınmasına neden olur, bu da bakım maliyetlerini düşürür.

Dünya Sağlık Örgütü (DSÖ), düz yol profillerinin trafik kazalarını azaltmada önemli bir rol oynadığını ve modern yolların pürüzsüz ve düzgün olması, yol güvenliği ve sürüş konforu açısından büyük önem taşımaktadır [37].

2.4. Engebeli (Bozuk) Yol Profili

Engebeli yol profili, yol yüzeyinde çeşitli pürüzler, çukurlar, tümsekler ve düzensizliklerin bulunduğu yol türlerini ifade eder. Bu tür yollar, genellikle düşük bakım standartlarına sahip bölgelerde veya doğal olarak engebeli arazilerde bulunur. Engebeli yol profilleri sürüş konforunu azaltır ve araç üzerinde ekstra mekanik stres yaratırlar.

Engebeli yol profilinin özelliklerini aşağıdaki gibi sıralayabilir;

Yüzey Pürüzlülüğü: Engebeli yollar, çeşitli düzensizlikler ve yüzey bozuklukları içerir. Bu pürüzler, sürüş sırasında titreşim ve sarsıntıya neden olur.

Çukurlar ve Tümsekler: Bu yollar, sıkça karşılaşılan çukurlar ve tümseklerle karakterizedir. Bu da sürüş güvenliğini tehlikeye atabilir ve aracın süspansiyon sistemine zarar verebilir.

Düzensiz Yüzey: Yol yüzeyindeki düzensizlikler, sürüş konforunu ve kontrolünü zorlaştırır.

Bu yolların sağlamış olduğu avantajları ise aşağıdaki gibi sıralayabiliriz;

Doğal Estetik: Kırsal ve doğal alanlarda, engebeli yollar doğal estetiği koruyabilir.

Düşük Maliyetli Yapım: Bu tür yollar, daha az altyapı yatırımı gerektirir.

Bu yolların oluşturduğu dezavantajları ise aşağıdaki gibi sıralayabiliriz;

Düşük Sürüş Konforu: Yüzey pürüzlülüğü nedeniyle, araç içinde yüksek düzeyde titreşim ve sarsıntı hissedilir.

Araç Hasarı: Sürekli düzensiz yüzey koşulları, aracın süspansiyon ve diğer mekanik bileşenlerinde hızla aşınmaya neden olabilir.

Güvenlik Riskleri: Çukurlar ve tümsekler, sürücünün kontrolü kaybetmesine ve potansiyel kazalara yol açabilir.

Aşağıdaki karşılaştırmalı Tablo 2.2' de, düz ve bozuk yol profillerinin farklı yönlerini vurgulamakta ve her iki yol profilinin sürüş konforu, güvenlik, yakıt verimliliği ve araç aşınması üzerindeki etkilerini ortaya koymaktadır [36,38].

Tablo 2.2. Düz ve bozuk yol profilleri arasındaki karşılaştırmalı ilişki tablosu [36,38]

Özellik	Düz Yol Profili	Bozuk Yol Profili
Yüzey Pürüzlülüğü	Pürüzsüz ve düzgün yüzey	Pürüzlü, düzensiz yüzey, çukurlar ve tümsekler
Sürtüş Konforu	Yüksek sürtüş konforu, az sarsıntı ve titreşim	Düşük sürtüş konforu, yüksek sarsıntı ve titreşim
Araç Aşınması	Daha az mekanik aşınma, düşük bakım maliyetleri	Artan mekanik aşınma, yüksek bakım maliyetleri
Yakıt Verimliliği	Yüksek yakıt verimliliği	Düşük yakıt verimliliği
Güvenlik	Daha yüksek güvenlik, düşük kaza riski	Daha düşük güvenlik, yüksek kaza riski
Hız ve Performans	Yüksek hız ve performans	Düşük hız ve performans
Bakım Gereksinimleri	Düşük bakım gereksinimleri	Yüksek bakım gereksinimleri
Kullanım Alanları	Şehir içi yollar, otobanlar, otoyollar	Kırsal yollar, orman yolları, dağlık bölgeler
Emisyonlar	Daha düşük emisyonlar	Daha yüksek emisyonlar
Yolcuların Rahatlığı	Yolcular için rahat ve konforlu	Yolcular için rahatsız edici
Yüzey Pürüzlülüğü	Pürüzsüz ve düzgün yüzey	Pürüzlü, düzensiz yüzey, çukurlar ve tümsekler

3. ARAÇ BAĞLANTISI VE VERİ OKUMA

Araçlar üzerindeki anlık verileri alabilmek için gerek bağlantı yapısı gerekse de veri okuma protokolü birlikte değerlendirilmelidir. Günümüzdeki modern araçlar, güvenlik, performans ve emisyon kontrolünü optimize etmek için birçok farklı sensörle donatılmıştır. Bu sensörlerden elde edilen veriler, aracın Elektronik Kontrol Ünitesi (EKU) tarafından işlenir ve değerlendirilir. Sensör verilerinin doğru bir şekilde okunması, aracın durumunu izlemek, arızaları teşhis etmek ve bakım gereksinimlerini belirlemek açısından kritik öneme sahiptir.

Günlük hayatta pek çoğumuz araçlar üzerindeki sensörlerin sadece arıza durumlarında bizlere arızaları bildirmek için olduğunu düşünsede aslında durum bu düşüncenin çok daha ilerisindedir. Araçlarımızda bulunan onlarca sensör araç sürüşlerimiz sırasında bizlere fark etsek de etmesek de sürekli olarak belirli saykılarla verileri vermektedir. Bu veriler aslında araçlarımızın nasıl bir işleyiş halinde olduğu hakkında da bizlere büyük datalar oluşturacak şekilde verilere sunmaktadır.

Bu verilere ulaşmak aslından iki aşama üzerinden kurulu bir yapıdır. Bunlardan ilki On-Board Diagnostics II (OBD-2) denilen veri bağlantı soketleri ile araç ile fiziksel katman üzerinden bağlantıyı sağlamak diğeri ise Controller Area Network (CanBus) denilen endüstriyel uygulamalarda kullanılan bir veri iletişim protokolü üzerinden sensör verilerinin araç üzerinden alınmasıdır.

3.1. On-Board Diagnostics (OBD-2)

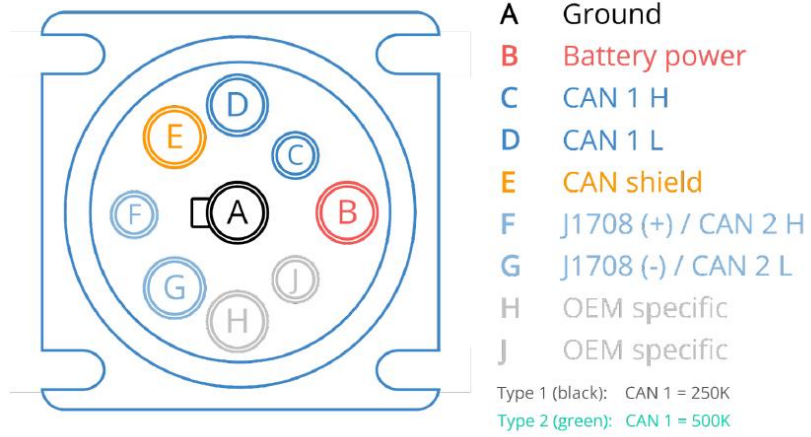
Günümüzde modern araçların hayatımıza girmesiyle birlikte araçlar üzerindeki motor, emisyon, arıza ve diğer sensörlerin verilerine ulaşmak için bir gereklilik ortaya çıkmıştır. Bu gereklilikten doğarak 1996 tarihinde ilk önce Amerika Birleşik Devletleri'nde tüm araçlarda zorunlu hale getirildi ardından 2001 yılında Avrupa Birliği ülkelerinde benzinli araçlarda ve son olarak 2003 yılında ise tüm dizel araçlarda bulunması zorunluluğu getirildi.

Günümüzde kullanmış olduğumuz araçlarımızın hemen hemen tüm motor parçalarının üzerinde bulunan ve o parçalarda oluşabilecek arızaları kullanıcılara bildirmesi, servislerinde onarım işlerinin kolaylaştırması için onlarca sensör bulunmaktadır. Bu sensörler, motor devrinden, akü voltajına, motora giren oksijen miktarından, yakıt sıcaklığına kadar pek çok bilgi bu sensörler sayesinde okunup anlık olarak OBS-2 üzerinden toplanabilmektedir.

3.1.1. J1939 Bağlantı Konektörü

OBD-2 konektörlerinin yaygın bulunan iki çeşiti araçlar üzerinde bulunmaktadır. Bunlardan ilki daha çok ağır vasıta, iş ve tarım araçları gibi araçların üzerinde bulunan çeşidi olan

J1939 konnetörlerin 9 pinli modelidir. Bu modelin pin yapılarını aşağıdaki Şekil 3.1’ de görebilirsiniz.



Şekil 3.1. J1939 bağlantı konnektörlerinin 9 pinli modelinin bağlantı pin uçları [39]

Yukarıda pin yapıları verilmiş olan J1939 modelinin bir bilgisayara ya da el terminaline bağlanması için diğer uçları uygun bir sokete dönüştürülmesi gereklidir. Bu soket günümüzde USB olabileceği gibi bir seri port olabilmektedir. Aşağıda J1939 modeline ait en çok kullanılan tarzların dan olan bağlantı kablosunun gerçek resmi Şekil 3.2’ de verilmiştir.



Şekil 3.2. J1939 bağlantı konnektörlerinin 9 pinli gerçek görüntüsü [39]

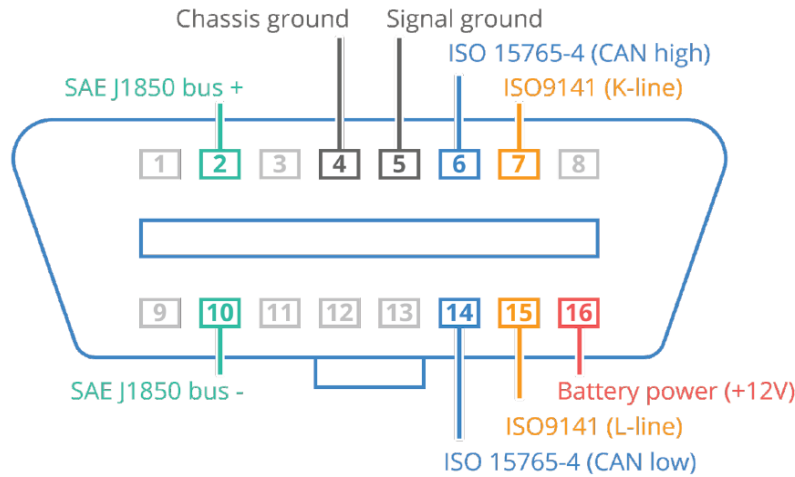
Yukarıda verilen J1939 bağlantı konektörünün üzerinde bulunan dokuz adet pinin fonksiyonu ve açıklamaları aşağıdaki Tablo 3.1’ de verilmiştir.

Tablo 3.1. J1939 bağlantı konektörlerinin pin fonksiyonları ve açıklamaları

Pin Numarası	Fonksiyon	Açıklama
1	Şasi Topraklaması	Araç şasisi ile bağlantılı topraklama
2	Güç (Kalıcı 12V)	Kalıcı 12V güç kaynağı
3	J1939 CAN Yüksek (CAN High)	J1939 veri yolu yüksek sinyal hattı (CAN H)
4	Batarya Gücü (Geçici 12V)	Geçici 12V güç kaynağı (kontak açıkken)
5	Zemin (Ground)	Araç topraklaması
6	J1939 CAN Düşük (CAN Low)	J1939 veri yolu düşük sinyal hattı (CAN L)
7	J1708 A+	J1708 protokolü için A+ sinyal hattı
8	J1708 B-	J1708 protokolü için B- sinyal hattı
9	Şasi Topraklaması veya OEM Özeli	Üreticiye özel veya ek şasi topraklaması

3.1.2. SAE J1962 Bağlantı Konektörü

Günlük araçlarımızda kullandığımız bir diğer OBD-2 soketimizi ise SAE J1962 standart numaralı 16 pinli olan modelidir. Bu modeli en çok göreceğimiz araçlar otomobillerdir. Konektörlerin OBD-2 bağlantı soketi genellikle araçların direksiyona yakın olan bölümlerinde bulunmaktadır. Bu modelin pin yapılarını aşağıdaki Şekil 3.3’ de görebilirsiniz.



Şekil 3.3. SAE J1962 bağlantı konektörlerinin 16 pinli modelinin bağlantı pin uçları [39]

Yukarıda pin yapıları verilmiş olan SAE J1962 modelinin bir bilgisayara ya da el terminaline bağlanması için diğer uçları uygun bir sokete dönüştürülmesi gereklidir. Bu soket günümüzde USB olabileceği gibi bir seri port olabilmektedir. Aşağıda SAE J1962 modeline ait en çok kullanılan tarzların dan olan bağlantı kablosunun gerçek resmi Şekil 3.4’ de verilmiştir.



Şekil 3.4. SAE J1962 bağlantı konektörlerinin 16 pinli modelinin gerçek görüntüsü [39]

Yukarıda verilen bağlantı konektörünün üzerinde bulunan on altı adet pinin fonksiyonu ve açıklamaları aşağıdaki Tablo 3.2’ de verilmiştir.

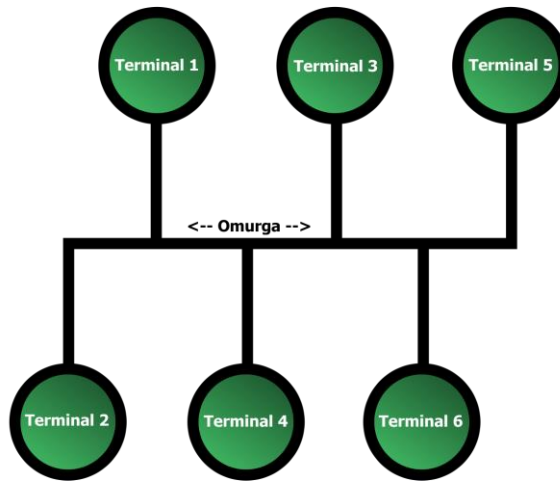
Tablo 3.2. SAE J1962 bağlantı konektörlerinin pin fonksiyonları ve açıklamaları

Pin Numarası	Fonksiyon	Açıklama
1	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
2	SAE J1850 Bus+	SAE J1850 PWM/VPW veri iletim hattı
3	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
4	Şasi Topraklaması	Araç şasi topraklaması bağlantısı
5	Sinyal Topraklaması	Sinyal topraklaması bağlantısı
6	ISO 15765-4 (CAN high)	CAN yüksek veri iletim hattı
7	ISO 9141-2	Veri iletim hattı
8	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
9	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
10	SAE J1850 Bus-	SAE J1850 PWM/VPW veri iletim hattı
11	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
12	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
13	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
14	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır
15	SAE J1850 Bus+	SAE J1850 PWM/VPW veri iletim hattı
16	Üreticiye Özel	Üreticiye özel uygulamalar için ayrılmıştır

3.2. Controller Area Network (CANBus)

CANBus, tam açılımı ile Control Area Network Bus (Kontrol Alanı Ağı Veri Yolu) ilk olarak 1983 yılında içinde Robert BOSCH tarafından geliştirildi [40]. Otomotiv sektöründe zamanla ortaya çıkan kablo yükü problemini ortadan kaldırmak için düşünülüp geliştirilmiş bir çözümdür. O yıllarda araçlar üzerinde donanım parçalarının artması bunlarında sürücüler tarafından kontrol edilmesinin gerekliliği ve sensör teknolojisinin ilerleyişi ile bu sensörlerinde gerek sürücü gerekse de araç üzerindeki donanımlarla haberleşmesi gerekliliği bu teknolojinin ortaya çıkmasını sağlamıştı. Zamanla ortaya çıkan bu donanım ve sensör yoğunluğu araçların üzerindeki kablo yükünün artmasına neden olmasıyla araçların ağırlığında ekstradan +70 kg'ye kadar yük oluşturmuştur. Böylece CANBus veri iletişim protokolünün ortaya çıkmasıyla bu kablo yükünden kurtulunmuş ve daha hızlı ve güvenli bir alt yapıya kavuşturulmuştur.

CANBus haberleşme protokölü adından da anlaşılacağı gibi BUS tabanlı bir ağ topolojisinin temelleri üzerine oturtulmuştur. Bu yapı diğer adres temelliler gibi değilde, ID temelli bir yapı üzerinde çalışmaktadır. Burada mesajlar birer frame yapısı üzerinde gönderilmektedir. Bus topolojisi adına backbone (Omurga) adı verilen bir ana hat (kablo) üzerinde bulunan ve bu omurga hatta bağlı olan terminallerden oluşan bir sistemdir. Burada mesaj iletiminde mesaj paketleri ana omurga üzerinden herkesin duyabileceği şekilde yayılır ama mesajı bu bağlı terminallerden ilgili olanı alıp yorumlar. Mesajlaşma bu topolojide çift yönlü olarak ilerletilebilmektedir. Bus topolojinin de en büyük avantajlarından biriside kablo miktarını azaltarak tek bir hat üzerinden verilerin iletilmesi ve cevaplarının aynı hat üzerinden alınması temeline dayanmaktadır. Bu yüzden CAN haberleşme protokölü Bus topolojilerin üzerine kurulmasına sebep oluşturmuştur. Aşağıdaki Şekil 3.5' de Bus ağ topolojide çalışan terminallerin bağlantı şeklini görebilirsiniz.



Şekil 3.5. Bus ağ topoloji ile bağlı terminallerin bağlantı şekli

CANBus haberleşme protokolü üzerinden veri paketleri iletilirken üç ana temel parametre ile paketlerin içleri doldurulur. Bunlardan ilki olan ID parametresi bu iletilecek veri mesajının hangi kontrol ünitesi ile ilgili olduğunun bilgisi tutulan bir parametredir. Diğer paket parametresi ise verinin uzunluk bilgisinin taşındığı parametredir, burada verilebilecek en yüksek veri uzunluk değeri ise 8 byte'dır. Verinin uzunluk bilgisi 0 ile 8 byte arasında olup ilgili elektronik kontrol ünitesinin alacağı veri boyutuna göre oluşturulur. Son parametremiz ise verinin taşıyacağı data bilgisini tutmaktadır. Bu data bilgisi ilgili elektronik kontrol ünitesine yapması gereken işlemi veya ondan gelecek olan bilginin verilerini taşımaktadır. Aşağıda verilmiş olan Şekil 3.6' da bir CANBus üzerinde taşınan veri paketini ve paket içeriğindeki verilerin boyutları gösterilmektedir. Tablo 3.3' de ise bu veri paketinde bulunan birimlerin açıklamaları bulunmaktadır.

SOF	Mesage Identifier (SID)	RTR	IDE	R0	0-8 Bytes Veri	CRC	ACK	EOF	IFS
1 Bit	11 Bit	1 Bit	1 Bit	1 B	0 ~ 64 Bit	15 Bit	1 Bit	7 Bit	3 Bit

Şekil 3.6. CANBus üzerinde taşınan veri paketi ve paket içeriğindeki verilerin boyutları

Tablo 3.3. CANBus üzerinde taşınan veriler ve açıklamaları

Veri	Açıklama
SOF - Çerçeve Başlangıcı	Çerçevenin başlangıcını işaret eden ve düğümler arasındaki senkronizasyon için kullanılan 1 bit.
SID - Standart Tanımlayıcı	Mesajın kimliğini ve önceliğini temsil eden 11 bit.
RTR - Uzak Transfer İsteği	Bilgi başka bir düğümden gerektiğinde dominant olan 1 bit.
IDE - Tanımlayıcı Etkinleştirildi	Bu bit, genişletilmiş tanımlayıcının gönderilip gönderilmeyeceğini belirtir. 1 bit
R0	Gelecekte kullanılmak üzere ayrılmış bitler. 1 bit
Data -Veri	Taşınan veri bilgisidir. 0 ~ 64 bit
CRC - Döngüsel Artıklık Kontrolü	Verinin doğruluğunu kontrol etmek için kullanılır. 15 bit
ACK - Onay Biti	Mesajı alan cihaz, hatasız mesajın alındığını belirtmek için bu biti değiştirir. Eğer hiçbir cihaz biti değiştirmezse, gönderici mesajı yeniden gönderir. 1 bit
EOF - Çerçeve Sonu	Mesajın sonunu işaret eden 7 bit.
IFS - Çerçeve Arası Boşluk	Alicının alınan çerçeveyi doğru şekilde işleyip uygun tampon alanına taşınması için yeterli zamanı temsil eden 3 bit.

CANBus'ın araçlarda standartlaşmaya başlamasıyla birlikte araçlardaki arıza ve sensör okuma işlemleri daha farklı boyuta ulaşmaya başladı. Günümüzde servisler hata ve arıza kontrollerini CANBus yoluyla okuyor var olan akıllı sistemleri üzerinden bu hata ve arıza durumlarının düzeltilmesi için gerekli adımlar servis görevlisine adım adım bildiriliyor. 2002 yılı sonrası araçlar üzerinde CANBus üzerinden okunacak sensör verilerinde yoğun bir artış yaşanmaya

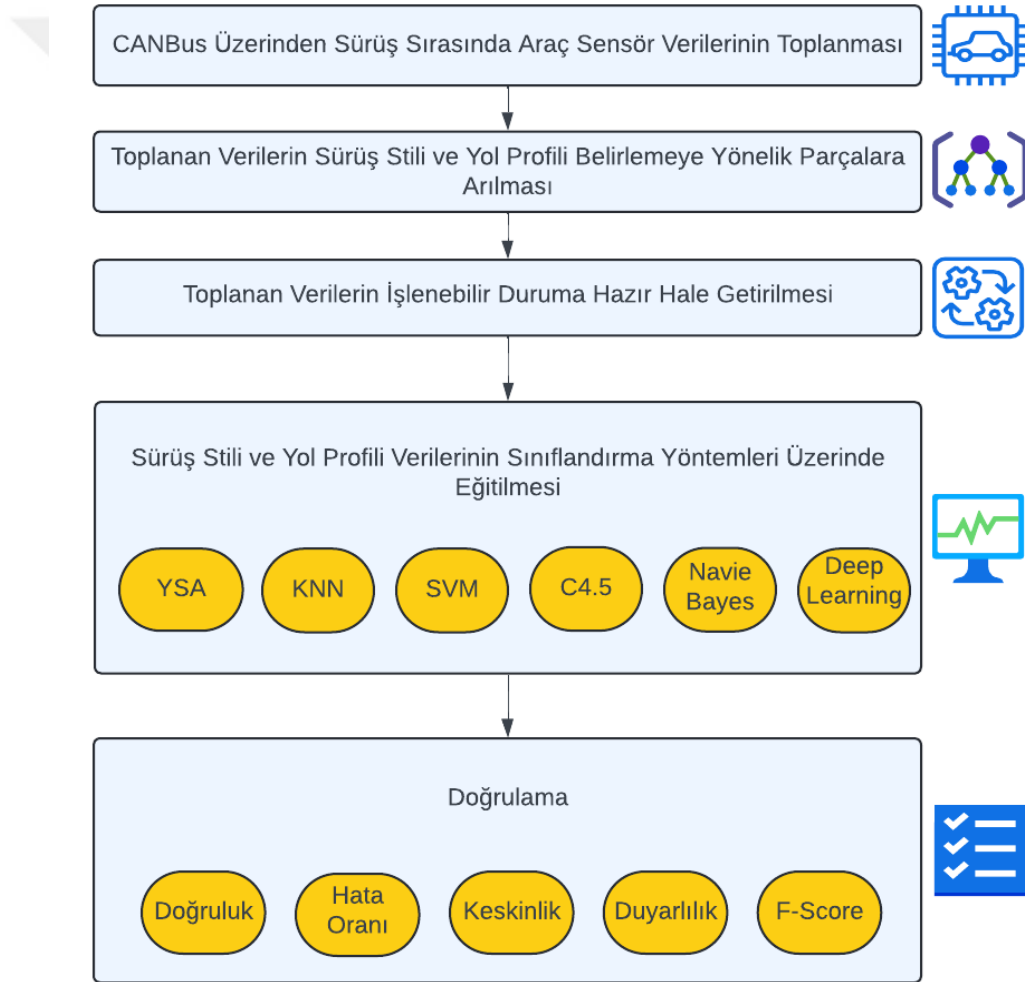
başlandı. Araçların markası ve modeline göre okunabilecek veriler arasında büyük çeşitlilik farkları oluşmaya başladı. Özellikle ortak olarak okunabilecek veriler bir standart hal almaya başladı. Aşağıdaki Tablo 3.4’ de günümüzde kullanılan çoğu araçtan CANBus yoluyla okuyabileceğimiz veriler, birimleri, maksimum ve minimum ölçüm aralıkları vb. yer verilmiştir.

Tablo 3.4. CANBus ile çoğu araç üzerinden okunabilecek veriler ve ayrıntıları [41]

PIDs (hex)	PIDs (dec)	Veri Alanı	Açıklama	Min. Değ.	Maks. Değ.	Birim
4	4	1	Hesaplanan Motor Yüğü	0	100	%
5	5	1	Motor Soğutma Sıcaklığı	-40	215	C
0A	10	1	Yakıt Basıncı	0	765	kPa
0B	11	1	Manifold Basıncı	0	255	kPa
0C	12	2	Motor RPM	0	16383,75	rpm
0D	13	1	Araç Hızı	0	255	km/h
0F	15	1	Emme Hava Sıcaklığı	-40	215	C
10	16	2	Hava Akış Hızı	0	65535	g/s
11	17	1	Gaz Pedalı Konumu	0	100	%
46	70	1	Ortam Hava Sıcaklığı	-40	215	C
5C	92	1	Motor Yağı Sıcaklığı	-40	210	C

4. MATERYAL VE METOT

Bu tez çalışmasında tasarlanan sistemin teori süreci, beş ana basamakta ilerlemektedir. Bu beş basamak, sistemin işleyişini ve amacına ulaşmasını sağlamak için dikkatle yapılandırılmıştır. Teori sürecinin ilerleyişi Şekil 4.1' de görsel olarak ayrıntılı bir şekilde gösterilmektedir. Bu adımların her biri, sistemin bütünlüğünü ve etkinliğini sağlamak için belirli bir sırayla ve dikkatle takip edilmesi gereken aşamalardır. Teori basamaklarının ayrıntılı açıklamaları, sistemin nasıl çalıştığını ve her bir basamağın hangi görevleri üstlendiğini net bir şekilde ortaya koymak için sonraki başlıklarda tek tek detaylı olarak ele alınmıştır. Bu sayede, sistemin tasarımı ve işleyişi hakkında kapsamlı bir anlayış sağlanmaktadır.



Şekil 4.1. Sistemin çalışma adımları

4.1. Verilerin Toplanması

Yapmış olduğumuz bu çalışmanın temel amacı, aracın OBD 2 portu üzerinden elde edilen araç sensör verilerini kullanarak, sürücülerin agresif ve sakin sürüş stilleri ile araçları sürdükleri yolların bozuk veya düz profillerinin belirlenmesi ve bu sürüş stilleri ile yol profillerinin eğitici sınıflandırma algoritmaları üzerinde belirleyip tahmin performanslarının analiz edilmesidir. Bu çalışmada, veri toplama işlemi belirlenen birden çok güzergah üzerinde gerçekleştirilmiş olup, agresif ve sakin sürüş stilleri ile bozuk ve düz yol profilleri için ayrı ayrı yapılmıştır. Bu süreçte, her sürüş tarzları ile yol profilleri için iki adet veri tablosu toplanmıştır.

Bu sürüşler sırasında, araç OBD2 portu üzerinden kaydedilen tüm sensör verileri, bizim yapacağımız tahmin işlemi için gerekli olan tüm verileri kapsamamaktadır. Dolayısıyla, toplanan sensör verileri arasından tahmin sonucunu doğru bir şekilde verebilecek olan en az sayıda ve en ilişkili değişkenlerin seçimi yapılmıştır. Veri toplama işlemi sırasında, MAF (hava akış oranı), hesaplanan yük, motor devri, anlık ekonomi ve araç hızı gibi değişkenler seçilmiştir. Bu değişkenlerin agresif ve sakin sürüşlerin yanı sıra bozuk veya düz olan yol profillerinde sürüş sırasında sürekli olarak değişkenlik gösterdiği belirlenmiştir.

Diğer sensör verileri ise, zaman, sürülen mesafe, ortalama ekonomi gibi, agresif ve sakin sürüşler ve bozuk ve düz yol profilleri hakkında doğrudan yorum yapılamayacak ve doğrusal artış gösteren değerler olduğundan dolayı analiz sürecine dahil edilmemiştir. Bu çalışmada, sürüş davranışlarını ve yol profillerini doğru bir şekilde analiz edebilmek için seçilen sensör verileri, agresif ve sakin sürüşler sırasında ve bozuk, düz yol sürüşleri sırasında farklı dinamikler sergileyen önemli parametreler olarak öne çıkmaktadır.

Sonuç olarak, belirli bir güzergah üzerinde toplanan bu veriler, sürüş tarzlarının doğru bir şekilde sınıflandırılması ve tahmin edilmesi amacıyla kullanılmıştır. Tablo 4.1.' de ise bu güzergah üzerinde toplanan ve seçilmiş olan sensör verilerini detaylı bir şekilde gösterilmektedir. Bu tablo, hangi sensör verilerinin analiz sürecinde kullanıldığını ve bu verilerin nasıl bir değişkenlik gösterdiğini detaylandırarak çalışmanın veri analizi bölümünde önemli bir referans noktası oluşturmaktadır. Bu şekilde, sürüş tarzlarının ve yol profillerinin belirlenmesi ve sınıflandırılması için kullanılan verilerin kapsamı ve niteliği açık bir şekilde ortaya konmuş olmaktadır.

Bu çalışmada kullanılan yöntem ve süreçler, sürüş stilleri ve yol profilleri belirlenmesi ve sınıflandırılması için geliştirilen algoritmaların etkinliğini değerlendirmekte önemli bir rol oynamaktadır. Sürüş verilerinin doğru ve etkili bir şekilde toplanması, bu verilerin anlamlı değişkenlere dönüştürülmesi ve ardından sınıflandırma algoritmalarının uygulanması, çalışmanın başarısı için kritik öneme sahiptir. Bu kapsamda, veri toplama, ön işleme, özellik seçimi ve sınıflandırma süreçleri dikkatle yapılandırılmış ve yürütülmüştür.

Tablo 4.1. Seçilen sensör verileri

Adet	Veri Türleri				
	A	B	C	D	E
X1	A_i	B_i	C_i	D_i	E_i
X2	A_{i+1}	B_{i+1}	C_{i+1}	D_{i+1}	E_{i+1}
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
Xn-1	.	.	.	.	.
Xn	A_{i+n}	B_{i+n}	C_{i+n}	D_{i+n}	E_{i+n}

A : MAF' dan geçen hava akış oranı (Gr/Sn)

B : Hesaplan yük (%)

C : Motor devri (RPM)

D : Anlık ekonomi (100 Km/Lt)

E : Hız (Km/h)

4.2. Verilerin Normalizasyonu

Normalizasyon işlemi, sınıflandırma algoritmalarında tahmin için kullanılacak veri setlerinin değerlerinin dağılımını düzenli hale getirmek amacıyla kullanılmaktadır. Bu işlem, değerler arasındaki ölçüt farkından doğabilecek baskınlıkların etkisini azaltmakla kalmaz, aynı zamanda sınıflandırma algoritmalarının öğrenme ve sınıflandırma aşamalarındaki performans kayıplarının önüne geçmek için de kritik bir rol oynar. Normalizasyon ile veri setindeki değerlerin belirli bir aralıkta kalmasını sağlar, böylece büyük ve küçük değerler arasındaki farkın sınıflandırma algoritmasının performansı üzerindeki olumsuz etkileri en aza indirgenir. Örneğin, bazı sensör verileri çok büyük değerler alırken, diğerleri çok küçük değerlerde kalabilir. Bu durumda, büyük değerler küçük değerlerin üzerindeki etkisini baskılayarak sınıflandırma algoritmasının dengesiz ve hatalı sonuçlar üretmesine neden olabilir. Normalizasyon, bu tür dengesizliklerin giderilmesi için uygulanır ve tüm verilerin aynı ölçekte değerlendirilmesini sağlar. Ayrıca, normalizasyon işlemi, sınıflandırma algoritmalarının daha hızlı ve daha verimli çalışmasına olanak tanır. Algoritmalar, normalizasyon sayesinde verileri daha hızlı işler ve bu da genel işlem süresini kısaltır. Bu da, sınıflandırma algoritmalarının öğrenme sürecini hızlandırarak daha kısa sürede daha doğru sonuçlar elde edilmesine katkıda bulunur.

Sonuç olarak normalizasyon işlemi, veri setlerindeki değerlerin dağılımını düzenleyerek sınıflandırma algoritmalarının tahmin ve öğrenme aşamalarındaki performans kayıplarını önlemek için vazgeçilmez bir adım olarak karşımıza çıkmaktadır. Bu işlem, veri setinin homojenliğini artırarak, sınıflandırma sürecinin daha sağlıklı ve dengeli bir şekilde ilerlemesini sağlar. Verilerin normalizasyonu, sınıflandırma algoritmalarının daha doğru ve güvenilir sonuçlar üretmesine yardımcı olur ve bu nedenle her veri analiz sürecinde dikkatle uygulanması gereken bir aşamadır.

CANBus üzerinden elde edilen beş adet veri setinde de her veri sütunu $XS=[X_1, X_2, \dots, X_{n-2}, X_{n-1}, X_n]$ dizisi olarak kabul edilmektedir. Bu veri setlerinin normalizasyon işlemleri aşağıdaki Denklem 4.1 nolu Min-Max formülüyle hesaplanmaktadır.

$$X' = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (4.1)$$

Burada, X' = Değerin normalize edilmiş hali, X_i = Normalize edilecek değer, $X_{\min}$ = Kullanılan veri setindeki minimum değer, $X_{\max}$ = Kullanılan veri setindeki maksimum değeri göstermektedir.

Aşağıda verilen Tablo 4.2' de CANBus aracılığı ile toplanan beş adet sensör verisinin normalize edilmiş ve etiketlenilmiş oluşacak tablonun yapısı gösterilmiştir.

Tablo 4.2. Normalize edilmiş ve etiketlenmiş veri seti tablosu

Adet	Veri Türleri				
	A	B	C	D	E
1	NormA _i	NormB _i	NormC _i	NormD _i	NormE _i
1	NormA _{i+1}	NormB _{i+1}	NormC _{i+1}	NormD _{i+1}	NormE _{i+1}
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
0	.	.	.	.	.
0	NormA _{i+n}	NormB _{i+n}	NormC _{i+n}	NormD _{i+n}	NormE _{i+n}

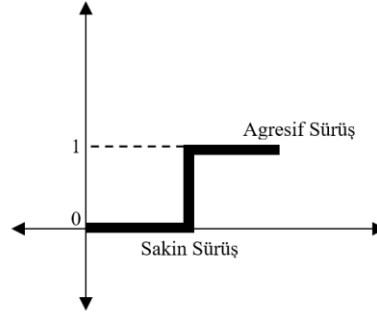
4.3. Sınıflandırma Algoritmalarına Sürüş Stilleri ve Yol Profillerinin Öğretilmesi

Sınıflandırma algoritmalarında öğrenme işleminin etkili bir şekilde gerçekleştirilebilmesi için, CANBus'tan alınan verilerin hangi sürüş tarzlarında kayıt altına alındıysa doğru bir şekilde etiketlenmesi büyük önem taşımaktadır [42]. Bu etiketleme işlemi, algoritmaların gelecekte karşılaşacağı ve sürüş tarzı ve yol profili hakkında bilgi içermeyen veri setlerini doğru bir şekilde sınıflandırabilmesine olanak tanır.

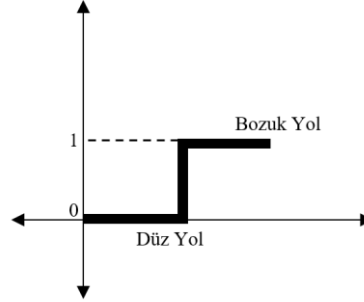
Yapmış olduğumuz bu çalışmada, agresif sürüş tarzını belirtmek için 1, sakin sürüş tarzını belirtmek için ise 0 değeri kullanılmıştır. Gene yol profili belirleme de ise düz yol 0 ile bozuk yol ise 1 ile etiketlenmiştir. Bu etiketleme sistemi, algoritmaların eğitim sürecinde hangi verilerin hangi sürüş tarzını yol profilini temsil ettiğini net bir şekilde anlamasını sağlar. Böylece, algoritmalar eğitim sırasında edindikleri bilgiler doğrultusunda, yeni gelen veriler üzerinde doğru sınıflandırmalar yapabilir.

Etiketleme işlemi tamamlandıktan sonra, bu veriler normalize edilerek sınıflandırma algoritmalarının performansını artırmak için kullanılır. Normalizasyon işlemi, verilerin belirli bir aralıkta ve düzenli bir dağılımda olmasını sağlayarak algoritmaların daha hızlı ve verimli çalışmasını mümkün kılar. Bu tez çalışmasında, normalize edilmiş agresif ve sakin sürüş veri seti ile düz ve bozuk yol profili bilgileri Tablo 4.2' de gösterildiği gibi ayrı ayrı bir şekilde öğrenme veri setleri olarak sunulmuştur.

Buna ek olarak, Şekil 4.2 ile Şekil 4.3' de sırasıyla agresif ve sakin sürüş ile düz ve bozuk yol profilini nitelendiren etiket değerleri grafik üzerindeki görsel olarak gösterilmiştir. Bu grafikler, farklı sürüş tarzlarının veri setleri üzerindeki etkilerini net bir şekilde ortaya koyar ve algoritmaların hangi aralıklarda hangi sürüş tarzlarını tespit ettiğini görsel olarak anlamamıza yardımcı olur.



Şekil 4.2. Sürüş stili etiketlerini gösterir grafik



Şekil 4.3. Yol profili etiketlerini gösterir grafik

CANBus verilerinin etiketlenmesi ve normalizasyon işlemleri, sınıflandırma algoritmalarının öğrenme sürecinde ve gelecekteki tahminlerinde doğruluk ve etkinlik sağlamak için kritik adımlardır. Bu adımlar sayesinde, sürüş tarzlarının doğru bir şekilde belirlenmesi ve sınıflandırılması daha mümkün hale gelir. Yapmış olduğumuz çalışmada kullanılan etiketleme ve normalizasyon yöntemleri, sınıflandırma algoritmalarının performansını optimize etmekte ve sürüş tarzları ve yol profillerinin doğru bir şekilde sınıflandırılmasına büyük katkı sağlamaktadır.

4.3.1. Yapay Sinir Ağları (YSA)

Aşağıdaki gösterilen Denklem 4.2' de, tezde kullanmış olduğumuz eğitici sınıflandırma algoritmalarından biri olan yapay sinir ağında kullanılacak olan transfer fonksiyonunun seçimi için kullanmış olduğumuz korelasyon hesaplamasını detaylı bir şekilde sunulmaktadır. Korelasyon değeri, gerçek sonuç ile tahmin edilen sonuç arasındaki ilişkinin yönünü belirlememize olanak tanıyan önemli bir metrik olarak öne çıkmaktadır [42]. Bu hesaplama sayesinde, sistemimizin elde ettiği sonuçlar ile eğitim verilerinde vermiş olduğumuz sonuç değerleri arasındaki ilişkiyi net bir şekilde analiz edebilmekte ve yorumlanabilmektedir.

Korelasyon değeri, -1 ile +1 arasında değişen değerler alabilmektedir. Bu değer -1'e yakın olması, iki değer arasında negatif bir ilişkinin bulunduğunu ve bir değer artması durumunda diğerinin azaldığını göstermektedir. Diğer yandan, korelasyon değerinin +1'e yakın olması, iki değer arasında pozitif bir ilişki olduğunu ve bir değer artması durumunda diğerinin de arttığını ifade etmektedir. Korelasyon değerinin 0 olması durumunda ise, iki değer arasında herhangi bir ilişkinin bulunmadığı ve bağımsız oldukları sonucuna varılmaktadır.

Yapay sinir ağında kullanılacak olan transfer fonksiyonunun seçimi, korelasyon hesaplaması ile doğrudan ilişkilidir. Elde edilen korelasyon değeri, transfer fonksiyonunun ne kadar etkili olduğunu ve tahmin edilen sonuçların gerçek sonuçlara ne kadar yakın olduğunu belirlenmesine yardımcı olur. Bu sayede, yapay sinir ağının performansını artırmak için en uygun transfer fonksiyonunu seçilmelidir. Korelasyon hesaplaması, yapay sinir ağında kullanılacak olan transfer fonksiyonunun seçiminde kritik bir rol oynamaktadır. Gerçek sonuçlar ile tahmin edilen sonuçlar arasındaki ilişkinin yönünü ve gücünü belirleyerek, sistemimizin doğruluğunu ve güvenilirliğini artırmak amacıyla hangi transfer fonksiyonunun daha uygun olduğunu tespit edebilmemizi sağlar. Bu yöntem, yapay sinir ağlarının eğitim sürecinde ve performans değerlendirmelerinde önemli bir adım olarak karşımıza çıkmaktadır.

$$r_{xy} = \frac{\sum xy - \frac{(\sum x)(\sum y)}{N}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{N}\right] \left[\sum y^2 - \frac{(\sum y)^2}{N}\right]}} \quad (4.2)$$

x : Bağımsız değişken (Tahmin için kullanılan)

y : Bağımlı değişken (Tahmin edilmeye çalışılan)

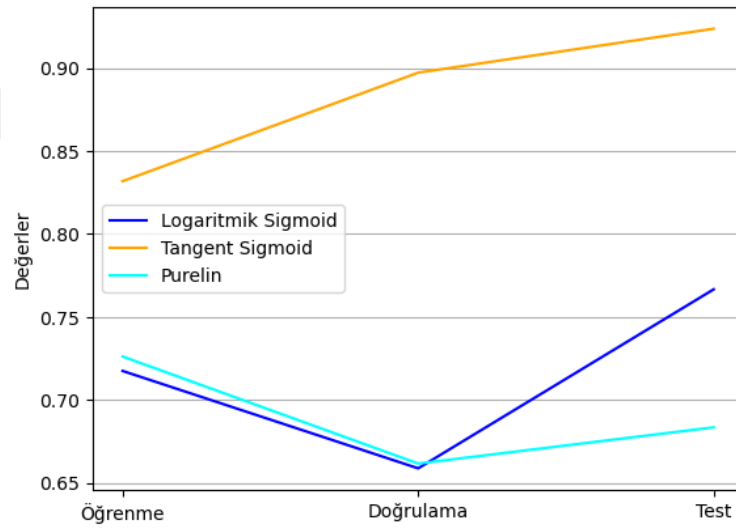
N : Yapılan gözlemlerin sayısı

Yapılan test uygulamaları sonucunda, çeşitli transfer fonksiyonlarının performansını değerlendirmek ve en başarılı olanı seçmek amacıyla Lineer Transfer Fonksiyonu (PURELIN), Unipolar Sigmoid Transfer Fonksiyonu (LOGSIG) ve Bipolar Sigmoid Transfer Fonksiyonu (TANSIG) kullanılmış ve bu fonksiyonların korelasyon değerleri hem öğrenme, hem doğrulama,

hem de test verileri için ayrı ayrı hesaplanmıştır. Elde edilen sayısal değerler, her bir transfer fonksiyonu için detaylı bir şekilde Şekil 4.4’ de sürüş stilleri, Şekil 4.5’ de ise yol profili dataların için görsel olarak sunulmuştur.

Şekil 4.4’ e ve Şekil 4.5’ e bakıldığında, korelasyon değerlerinin analiz edildiği görülmektedir. Bu analiz sonucunda her iki problemimiz içinde, Bipolar Sigmoid Transfer Fonksiyonu (TANSIG) korelasyon değeri açısından en iyi performansı gösteren fonksiyon olarak öne çıkmıştır. Korelasyon değeri +1’e en yakın olan transfer fonksiyonunun Bipolar Sigmoid Transfer Fonksiyonu (TANSIG) olduğu net bir şekilde görülmektedir. Bu bulgu, TANSIG fonksiyonunun, gerçek sonuçlarla tahmin edilen sonuçlar arasındaki ilişkiyi en doğru şekilde yansıttığını göstermektedir.

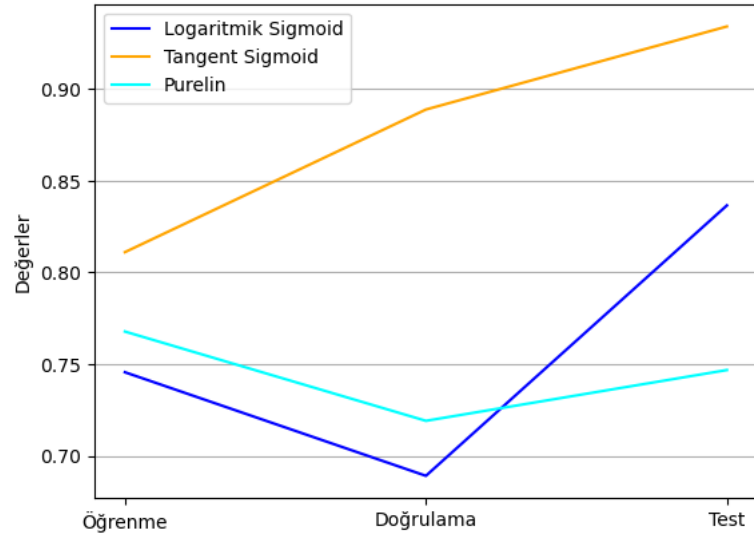
Ayrıca, bu transfer fonksiyonlarına ait korelasyon değerlerinin sayısal olarak yer aldığı Tablo 4.3 ve Tablo 4.4’ de dikkatle incelenmiştir. Tablo incelendiğinde, tüm transfer fonksiyonlarının korelasyon değerlerinin detaylı bir karşılaştırması yapılabilmekte ve bu karşılaştırma sonucunda en tutarlı ve güvenilir tahminlerin Bipolar Sigmoid Transfer Fonksiyonu (TANSIG) kullanılarak elde edildiği görülmektedir.



Şekil 4.4. Sürüş stili transfer fonksiyonlarının korelasyon değerleri

Tablo 4.3. Sürüş stili transfer fonksiyonları korelasyon sayısal değerleri

	Öğrenme	Dorulama	Test
	Korelasyon	Korelasyon	Korelasyon
LogSig	0,7174	0,6587	0,7666
TanSig	0,8318	0,8972	0,9237
Purelin	0,7261	0,6616	0,6834



Şekil 4.5. Yol Profili transfer fonksiyonlarının korelasyon değerleri

Tablo 4.4. Yol Profili transfer fonksiyonları korelasyon sayısal değerleri

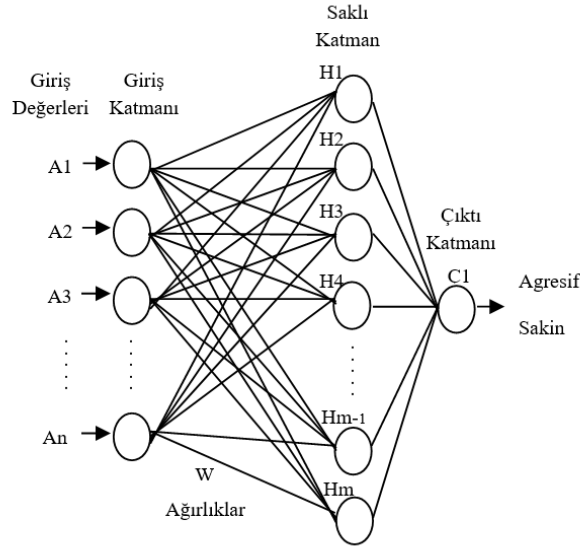
	Öğrenme	Doğrulama	Test
	Korelasyon	Korelasyon	Korelasyon
LogSig	0,7456	0,6891	0,8366
TanSig	0,8111	0,8889	0,9342
Purelin	0,7678	0,7191	0,7468

Tablo 4.3’ deki sürüş stili verileri ile Tablo 4.4’ deki yol profil verileri, her bir transfer fonksiyonunun performansını sayısal olarak ortaya koymakta ve bu veriler ışığında TANSIG fonksiyonunun üstünlüğünü vurgulamaktadır. En yüksek korelasyon değerine sahip olan TANSIG fonksiyonu, hem öğrenme sürecinde, hem doğrulama aşamasında, hem de test verileri üzerinde en doğru sonuçları üreten transfer fonksiyonu olarak belirlenmiştir. Bu test uygulamaları ve analizler, transfer fonksiyonlarının performansını değerlendirerek, hangi fonksiyonun daha etkin ve doğru sonuçlar verdiğini belirlememize yardımcı olmuştur. Korelasyon değeri açısından en iyi performansı gösteren Bipolar Sigmoid Transfer Fonksiyonu (TANSIG), yapay sinir ağlarının eğitim ve tahmin süreçlerinde en tutarlı sonuçları sağlamaktadır. Bu nedenle, bu çalışmada en uygun transfer fonksiyonu olarak TANSIG tercih edilmiş ve diğer transfer fonksiyonlarına kıyasla daha üstün bir performans sergilediği tespit edilmiştir.

Transfer fonksiyonu olarak Bipolar Sigmoid Transfer Fonksiyonu (TANSIG) seçilmesiyle oluşturulan YSA’ nın transfer hesabı aşağıdaki Denklem 4.3’ e göre hesaplanmaktadır.

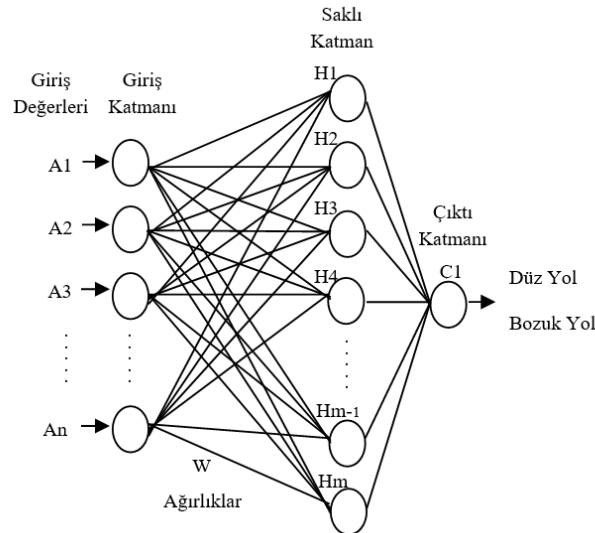
$$F(NET) = \frac{(e^{NET} + e^{-NET})}{(e^{NET} - e^{-NET})} \quad (4.3)$$

Yapay sinir ağı için yukarıdaki karar işlemlerinden sonra sürüş stilleri için Şekil 4.6 ve yol profilleri için Şekil 4.7’ deki gibi yapay sinir ağı modellerini oluşturulması için gerekli tüm teorik veriler toplanmıştır.



A1: MAF’ dan hava akış oranı, **A2:** Hesaplanan yük, **A3:** Motor devri, **A4:** Anlık ekonomi, **A5:** Araç hızı

Şekil 4.6. Sürüş stili yapay sinir ağı tasarımı



A1: MAF’ dan hava akış oranı, **A2:** Hesaplanan yük, **A3:** Motor devri, **A4:** Anlık ekonomi, **A5:** Araç hızı

Şekil 4.7. Yol profili yapay sinir ağı tasarımı

4.3.2. Naive Bayes

Tez için kullanmış olduğumuz diğer bir eğitici sınıflandırma algoritması da Naive Bayes algoritmasıdır. Bu algoritma, her bir özelliğin birbirinden koşulsal bağımsız olduğunu varsayar ve öğrenilmek istenen kavramın tüm bu özelliklere koşulsal olarak bağlı olduğu bir Bayes ağı olarak değerlendirilebilir. [42]. Naive Bayes algoritması, istatistiksel hesaplamaları yaparken ilk olarak değişken tanımlama işlemi ile başlar. Bu adımda, kullanılacak olan veri kümeleri belirlenir ve bu kümeler üzerinde işlem yapılır. Bizim çalışmamızda kullanılan değişkenler, Tablo 4.1' de görüldüğü üzere $X = [A_i, B_i, C_i, D_i, E_i]$ şeklinde tanımlanan veri kümeleri olarak ifade edilmektedir. Bu veri kümeleri, her biri belirli sürüş parametrelerini temsil eden öznitelikler içermektedir. Bu özniteliklere dayalı olarak, Naive Bayes algoritması ile sınıflandırma işlemleri gerçekleştirilir. Ayrıca, bu veri kümelerinde m adet sınıf olduğunu varsayarsak, bu sınıfları $C_1, C_2, \dots, C_m$ şeklinde tanımlayabiliriz.

Naive Bayes algoritmasında, bir örneğin belirli bir sınıfa ait olma olasılığı, yani agresif ya da sakin sürüş olup olmadığını veya düz ya da bozuk yol olup olmadığını belirlemek için kullanılan hesaplamalar Denklem 4.4 ile ifade edilir. Bu formül, belirli bir sınıf i'den bir örneğin agresif yada sakin veya düz yada bozuk yol olma olasılığını hesaplamak için kullanılır. Algoritmanın temel varsayımlarından biri olan bağımsız özniteliklerin hesabı ise Denklem 4.5 ile yapılır.

$$P(X|C_i) = P(X|C_i)P(C_i) \quad (4.4)$$

$$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i) = P(x_1|C_i) \times P(x_2|C_i) \times \dots \times P(x_n|C_i) \quad (4.5)$$

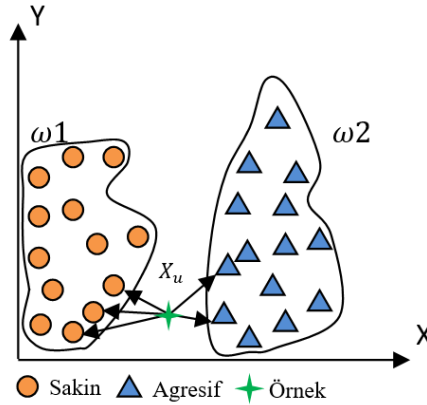
Bu yöntem, her bir özneneliğin diğer özniteliklerden bağımsız olarak değerlendirildiği ve bu bağımsızlık varsayımı altında olasılık hesaplamalarının yapıldığı bir süreçtir. Naive Bayes algoritması, bu şekilde her bir özneneliğin katkısını ayrı ayrı değerlendirerek toplam olasılığı hesaplar ve sınıflandırma kararını verir.

4.3.3. K En Yakın Komşu (KNN)

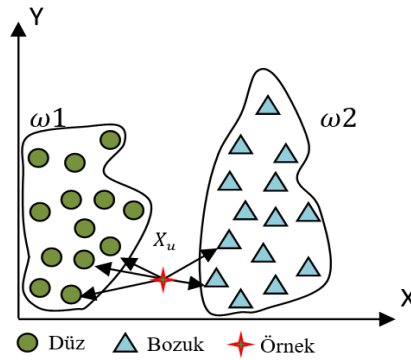
Bu eğitici sınıflandırma algoritması, yapı olarak basit, anlaşılır ve uygulanabilir bir algoritma türüdür. Algoritma, temel prensipleri ve işleyişi bakımından oldukça yalın olup, kullanıcılar tarafından kolaylıkla anlaşılabilir ve uygulanabilir bir yapıya sahiptir. Bu algoritma, bellek tabanlı bir yapıda çalışır, bu nedenle kendisine gelen sınıflandırılacak veriler her seferinde tekrar tekrar hesaplanarak değerlendirilir [42]. Bu yapı, algoritmanın dinamik olarak her veri seti için yeniden hesaplama yapmasına ve her seferinde en güncel verilerle çalışmasına olanak tanır.

Algoritmanın temel mantığı, Tablo 4.1' de gördüğümüz $X = [A_i, B_i, C_i, D_i, E_i]$ veri dizilerindeki değişkenlerin, sistemimizin sınıflandırmak istediği agresif ve sakin sürüş tarzları ile düz ve bozuk yol tiplerine göre n boyutlu bir uzayda yerleştirilmesine dayanır. Bu değişken grupları, belirlenen n boyutlu uzayda hangi sınıfa daha yakınsa, o sınıfa dahil edilir. Bu şekilde veriler algoritmada sınıflandırılmış olur. Sınıflandırma işlemi sırasında, grupların kendilerine gelen değişkenlere olan uzaklıkları belirlenir. Bu uzaklıklar hesaplanarak, en kısa olan grup tespit edilir ve bu değişken seti o sınıfa dahil edilir. Böylece, her bir veri kümesi, en yakın olduğu sınıfa atanmış olur. Bu yöntem, her veri kümesinin hangi sürüş tarzına ait olduğunu belirlemek için kullanılan etkili bir tekniktir.

Özellikle K En Yakın Komşu (KNN) algoritması, bu sınıflandırma yönteminin temelini oluşturur. KNN algoritması, veri noktalarının n boyutlu uzayda hangi sınıfa dahil edileceğini belirlemek için kullanılır. Bu algoritma, veri noktalarının komşuluk ilişkilerini değerlendirerek, her bir veri noktasının en yakın olduğu sınıfı belirler. Şekil 4.8' de sürüş stilleri için, Şekil 4.9' da ise yol profilleri için KNN algoritmasının n boyutlu bir uzayda hangi sınıfa dahil edileceğini gösteren bir koordinat düzlemi sunulmuştur. Bu şekil, algoritmanın işleyişini görsel olarak anlamamıza yardımcı olur ve veri noktalarının nasıl sınıflandırıldığını açıkça gösterir.



Şekil 4.8. KNN sürüş stili sınıflandırması koordinat düzlemi



Şekil 4.9. KNN yol profili sınıflandırması koordinat düzlemi

Bu uzaklık ölçme yöntemleri için çeşitli hesaplama formülleri bulunmaktadır. Sınıflandırma işlemlerinde en yaygın olarak kullanılan yöntemlerden biri Öklit mesafe ölçümüdür. Biz de bu çalışmamızda, en çok tercih edilen ve etkin sonuçlar veren Öklit mesafe ölçümünü kullanarak sınıflandırma yöntemini uygulamaktayız. Yukarıdaki Şekil 4.8 ve Şekil 4.9’ da gösterilen Xu verisi, sınıflandırılmak istenen veriyi temsil etmektedir. Bu verinin ait olduğu sınıfı belirlemek için kullanılan Öklit mesafe ölçüm yöntemi, iki nokta arasındaki düz çizgi mesafesini hesaplayan bir formüle dayanır. Bu formül, iki nokta arasındaki en kısa mesafeyi, yani doğrudan mesafeyi ölçer.

Öklit mesafe ölçümü, matematiksel olarak, iki nokta arasındaki farkların karelerinin toplamının karekökü olarak tanımlanır. Bu formül, n boyutlu bir uzayda iki nokta arasındaki uzaklığı hesaplamak için kullanılabilir. Denklem 4.6’ da, bu hesaplama yönteminin matematiksel ifadesini göstermektedir.

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4.6)$$

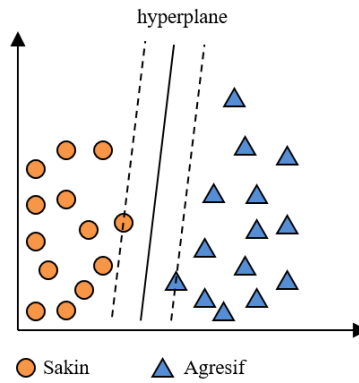
Formülde belirtilen x_i ve y_i değerleri, bilinen Öklit formülünde yatay ve dikey eksenlerdeki uzaklık miktarlarını temsil etmektedir. Bu değerler, iki veri noktası arasındaki her bir eksendeki farkları ifade eder. Bu farkların karelerinin toplamının karekökü alınarak iki nokta arasındaki Öklit mesafesi bulunur. Bu yöntem, sınıflandırma algoritmaları içerisinde özellikle K En Yakın Komşu (KNN) algoritmasında yaygın olarak kullanılır. KNN algoritmasında, sınıflandırılmak istenen veri noktası ile eğitim veri setindeki diğer noktalar arasındaki Öklit mesafeleri hesaplanır. Bu mesafeler kullanılarak, en yakın komşular belirlenir ve sınıflandırılacak veri noktasının ait olduğu sınıf tespit edilir.

Öklit mesafe ölçümü, basitliği ve doğruluğu nedeniyle tercih edilen bir yöntemdir. Verilerin uzaydaki dağılımını ve birbirlerine olan yakınlıklarını etkili bir şekilde belirleyerek, doğru ve güvenilir sınıflandırma sonuçları elde edilmesini sağlar. Bu nedenle, tez çalışmasında bu yöntemi kullanarak, sürüş verilerinin sınıflandırılmasında başarılı sonuçlar elde etmekteyiz.

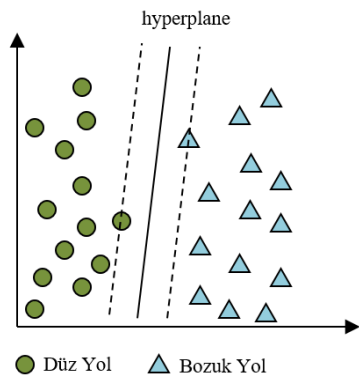
4.3.4. Destek Vektör Makineleri (SVM)

Destek Vektör Makineleri (SVM), bu tez çalışmasında sürüş stilleri ve yol profilleri belirlemede sınıflandırma işlemlerinde kullanılmış olup, prensip olarak bu iki sınıfı birbirinden en uzak mesafede ve eşit mesafede ayıran bir hiper düzlemin (hyperplane) bulunması temeline dayalı olarak çalışmaktadır. SVM algoritması, iki sınıf arasındaki en geniş boşluğu (marjin) sağlayacak şekilde verileri ayıran bir hiper düzlem bulmayı amaçlar. Bu hiper düzlem, verilerin yüksek boyutlu bir uzayda ayrılmasını sağlar ve bu ayırım işlemi, verilerin sınıflandırılmasında büyük bir doğruluk sunar.

Şekil 4.10’ da sürüş stilleri için ve Şekil 4.11’ de ise yol profilleri için gösterilen kesikli çizgiler, her iki sınıfa da en yakın olan verilerin sınır çizgisini belirlemektedir. Bu kesikli çizgiler, sınıflandırılacak olan verileri arasındaki mesafenin belirlenmesinde kritik bir rol oynamaktadır. Bu sınır çizgileri, verilerin hangi sınıfa ait olduğunu belirlemek için kullanılan karar sınırlarını temsil eder. Hiper düzlem, bu sınır çizgileri arasında yer almakta ve iki grup arasındaki mesafeyi ikiye bölmektedir. Eğitim verileri kullanılarak bulunan bu sınır çizgisi, her iki sınıfa da eşit mesafede olacak şekilde konumlandırılmıştır. Bu sayede, daha sonra gelecek olan test verilerinin hangi sınıfa ait olduğunu belirlemek mümkündür. Test verileri, aynı koordinat düzlemi üzerinde yerleştirilip, bu sınır çizgisinin hangi tarafında kaldığına göre sınıflandırılır. Eğer bir test verisi hiper düzlemin agresif sürüş tarafında kalıyorsa, bu veri agresif sürüş sınıfına ait olarak kabul edilir. Aynı şekilde, eğer test verisi hiper düzlemin sakin sürüş tarafında kalıyorsa, bu veri sakin sürüş sınıfına ait olarak değerlendirilir. Gene bir test verisi hiper düzlemin düz yol tarafında kalıyorsa, bu veri düz yol sınıfına ait olarak kabul edilir. Aynı şekilde, eğer test verisi hiper düzlemin bozuk yol tarafında kalıyorsa, bu veri bozuk yol sınıfına ait olarak değerlendirilir. Bu yöntemle, sınıflandırma işlemi etkin bir şekilde gerçekleştirilir.



Şekil 4.10. SVM sürüş stili koordinat düzlemi



Şekil 4.11. SVM yol profili koordinat düzlemi

4.3.5. C4.5 (J48) Algoritması

Tez çalışmasında kullanmış olduğumuz bir diğer sınıflandırma algoritması, C4.5 olarak da bilinen J48 algoritmasıdır. Bu algoritma, sahip olduğumuz veri kümelerine uygun olması nedeniyle seçilmiş olup, sınıflandırma işlemlerinde etkin ve güvenilir sonuçlar sağlamaktadır. C4.5 algoritması, sayısal değerler kullanarak karar ağaçları oluşturma yeteneğine sahip bir algoritmadır ve bizim araç CANBus üzerinden elde etmiş olduğumuz verilerin de sayısal veriler olmasından dolayı bu algoritma ideal bir seçim olmuştur [42].

C4.5 algoritmasının seçilmesinin bir diğer nedeni ise, olmayan değerler bulunmasına rağmen sınıflandırma işlemlerine kesmeden devam edebilme yeteneğidir. Bu özellik, verilerin tam olmadığı durumlarda bile algoritmanın etkili bir şekilde çalışabilmesini sağlamakta ve sınıflandırma işlemlerinin doğruluğunu artırmaktadır. Bu nedenle C4.5 algoritması, projede karşılaşılabilecek eksik veri durumlarında dahi güvenilir sonuçlar elde edebilmek amacıyla tercih edilmiştir. Sınıflandırma işleminin doğru bir şekilde gerçekleştirilebilmesi için her bir değişken kümesinin eşik değerinin belirlenmesi gerekmektedir. Bu eşik değerinin hesaplanabilmesi için, değişken kümesindeki değerlerin sıralı bir biçimde yazılması ve bu sıralı değerler arasından uygun bir eşik değerinin belirlenmesi gerekir. Tablo 4.1' de görünen değerler, her bir veri türü için $\{A_i, A(i+1), \dots, A(i+n)\}$ şeklinde sıralı bir biçimde kümelenmelidir.

Eşik değer hesaplaması, sıralı olan bu veri kümelerinin orta değerinin bulunması ile gerçekleştirilir. Orta değer, veri kümesinin ortasında yer alan değeri temsil eder ve bu değerinin bulunması ile eşik değeri hesaplanmış olur. Bu işlem, verilerin doğru bir şekilde sınıflandırılması için kritik öneme sahiptir. Denklem 4.7' de gösterildiği gibi orta değerinin bulunmasıyla eşik değeri belirlenir ve bu değer sınıflandırma işlemlerinde referans noktası olarak kullanılır.

$$t_k = \frac{A_k + A_{k+1}}{2} \quad (4.7)$$

C4.5 algoritmasının bir diğer önemli avantajı, karmaşık ve büyük veri setlerinde dahi yüksek performans gösterebilmesidir. Algoritmanın yapısı gereği, büyük veri setlerinde bile hızlı ve etkili bir şekilde çalışarak doğru sonuçlar elde edebilmesi projede kullanılması için önemli bir neden olmuştur. Ayrıca bu algoritma, verilerin analiz edilmesi ve sınıflandırılması sürecinde kullanıcı dostu bir yapı sunar.

Eşik değerleri belirlenmiş olan sayısal veri gruplarının karar ağacı üzerindeki dallanmalarını doğru bir şekilde yerleştirebilmek için, her bir veri grubunun entropilerinin hesaplanması gerekmektedir. Entropi, bilgi teorisinde bir sistemin belirsizlik derecesini ölçen bir metriktir ve karar ağacında kullanılacak parametrelerin belirsizlik ölçüsünü belirlemek için kritik bir rol oynar. Bu sayede, karar ağacının dallanma noktalarında hangi parametrelerin kullanılacağı ve bu parametrelerin hangi sırayla yerleştirileceği daha net bir şekilde ortaya konulabilir. Her bir

parametrenin entropi deęerinin hesaplanması bu parametrelerin sınıflandırma işlemlerindeki belirleyiciliklerini ölçmek için gereklidir. Entropi deęeri, bir parametrenin ne kadar belirsiz veya bilgi açısından ne kadar zengin olduğunu gösterir. Düşük entropi deęeri, parametrenin yüksek belirleyicilięe sahip olduğunu ve karar ağacında kritik bir rol oynayabileceğini gösterir. Yüksek entropi deęeri ise, parametrenin daha fazla belirsizlik içerdığını ve karar ağacında daha az belirleyici bir rol oynayabileceğini işaret eder. Entropi deęerleri Denklem 4.8’ de gösterildiğı gibi hesaplanmaktadır.

$$H(S) = - \sum_{i=1}^n p_i \log_2(p_i) \quad (4.8)$$

Yukarıdaki denklemde; $H(S)$: S kaynağının entropi deęerini, n üretilen mesaj sayısı, p_i mesajların üretilme olasılığıdır.

4.4. Derin Öğrenme Algoritmasıyla Sürüş Stilleri ve Yol Profillerinin Öğretilmesi

Tez çalışmasının son sınıflandırma algoritması olan Derin Öğrenme(Deep Learning) son yıllarda üzerine hemen hemen tüm disiplinler üzerinde çalışmalar yapılan bir algoritma haline geldi. Son yıllarda GPU teknolojisinin oluşup gelişmesi bir devrim gibi derin öğrenme kavramı herkesin hayatına soktu. Gelişen teknoloji ile birlikte verilerin boyutu git gide artmakta haliyle var olan eski teknoloji algoritmalar bu verilerin işleyip yorumlayamaz hale gelmesine neden olmaktadır. Derin öğrenme algoritmaları temelinde yapay sinir ağları üzerinden gelişen bir algoritma olmakla beraber yapay sinir ağlarına farkla çok daha fazla gizli katman barındırabilmesiyle ve çok büyük veriler üstünde işlemler yapabilmesi onun en ön plana çıkmasının sebeplerindendir.

Tez çalışması sırasında test sürüşlerinde araç üzerinden CANBus aracılığıyla eş zamanlı olarak veri toplanırken, hem aracın veri üretme hızı hemde kullanmış olduğumuz yazılım veri kaydetme hızına bağılı olarak milisaniyeler aralığında sensör verisi toplanabilmektedir. Bu durum uzun araç sürüşlerinde haliyle elliye yakın sensör ve parametrenin on binlerce hatta yüz binlerce satır halinde veri oluşturmaya olanak sağlamaktadır. Bu büyük veri kümesi içinde derin öğrenme algoritmalarının da başarısını görme gerekliliğı kaçınılmaz olmuştur.

Derin öğrenme yapıları doğası gereğı resim, video, dil işleme vb. durumlar üzerinde kullanılır olsa da kimi modeller ise büyük sayısal veri topluluklarında da kullanılmaktadır. Tez çalışması sırasında sayısal veriler üzerinde çalışmaları yapılmış en çok tercih edilen Uzun Kısa Süreli Hafıza Ağları (LSTM, Long-Short Term Memory) üzerinden testler yapılmıştır. LSTM modelinin başarı deęerlendirme aşamasında kullanılan performans kıyaslama ölçütü olarak MAE (Ortalama Mutlak Hatası), MSE (Ortalama Kare Hatası), RMSE (Kök Ortalama Kare Hatası), MAPE (Ortalama Mutlak Yüzde Hatası) hata deęerleri alınarak yapılmıştır. Aşağıdaki

denklemlerde sırasıyla MAE Denklem 4.9' da, RMSE Denklem 4.10' da, MAPE Denklem 4.11' de, MAE Denklem 4.12' de verilmiştir.

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (4.9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad (4.10)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (4.11)$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right| \quad (4.12)$$

4.4.1. Uzun-Kısa Süreli Hafıza Ağları (LSTM)

LSTM algoritması, uzun süreli bağımlılıkları ve öğrenme yeteneği olan bir Tekrarlayan Sinir Ağı (RNN) türü sinir ağıdır. LSTM hücreleri, bir bilgiyi uzun süreli olarak belleğinde saklayabilen ve gerektiğinde bu bilgiyi tekrar geri çağırabilen özel bir yapıya sahiptir. Bu özellikleri sayesinde, LSTM'ler zaman içindeki bağımlılıkları ve ilişkileri daha etkili bir şekilde modelleyebilir.

LSTM algoritmaları kullanım alanları olarak; Doğal dil işlemede, konuşma tanıma işlemlerinde, zaman serisi tahmininde, video analizi ve müzik, ses işleme alanlarında sıkça kullanılan bir yapıdır. LSTM yapısı RNN' e göre daha karmaşık bir şekilde çalışmaktadır.

Çalışma yapısı dört ana adımdan oluşmaktadır bunlardan ilki unutma kapısı olarak adlandırılan hücre durumunda hangi bilgilerin tutulup hangi bilgilerin atılacağına belirlendiği kısımdır. Unutma kapısındaki hangi değerin atılıp hangi değerin saklanmaya devam edeceği konusu kararı ise, bir önceki durumdan zaman adımından gelen verinin hesaplanmış sayısal değeri 0'a yakın ise atılması gerektiği 1'e yakın ise tutulması gerektiğini gösterir. Unutma kapısında yapılan bu işlem için kullanılan Denklem 4.13' de verilmiştir.

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (4.13)$$

İkinci kısım olan girdi kapısı hücre durumuna hangi yeni bilgi ekleneceğini belirleyen kısımdır. Çalışma yapısı olarak mevcut zaman içinde giren bilginin önceki zamandan üretilerek gelen gizli durumla birleştirilerek Şekil 4.12' de gösterildiği gibi sigmoid aktivasyon fonksiyonu ile işleme sokulup bu işlemle sonucunda gene ilk katmanda olduğu gibi bize 0 ile 1 arasında bir sayısal değer oluşturur. Oluşan bu ön değer hiperbolik tanjant aktivasyon fonksiyonu ile oluşan değerlerin çarpılmasıyla çıkan sonuca göre var olan birim duruma eklenip eklenmeyeceği karar verilir [43]. Girdi kapısında yapılan bu işlem için kullanılan Denklem 4.14' de verilmiştir.

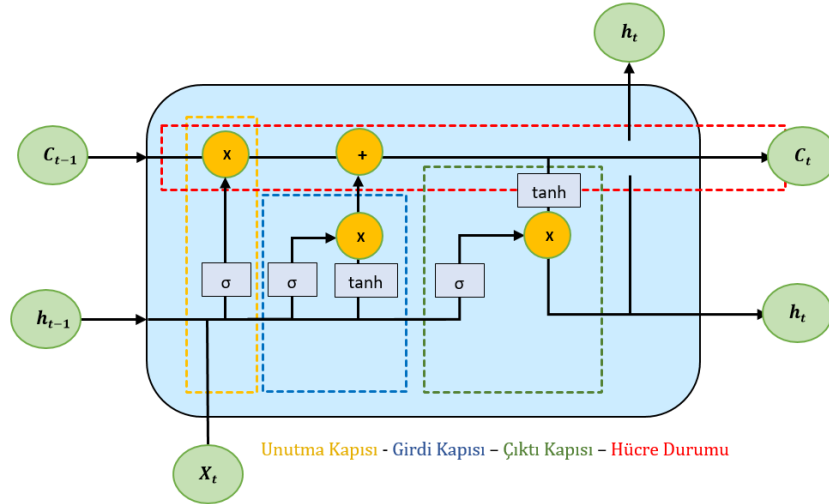
$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (4.14)$$

Üçüncü kısım olan çıkış kapısı unutmaya kapısı ile girdi kapısının durumlarına göre hücre durumunu güncelleyen kısımdır. Çalışma prensibi olarak çıkış kapısı var olan zamandaki giriş bilgisi ile önceki zamandan gelen gizli durum bilgisinin bir sigmoid fonksiyonla işleyip bir sayısal değer oluşturur. Bu oluşan sayısal değer 0 ile 1 arasında olup 0 'a yakın durumlar çıktı için önemli bir değer olmadığını gösterirken 1' yakın olan değerler ise çıktı için önem arz eden bir değer olduğunu gösterir. Çıktı kapısında yapılan bu işlem için kullanılan Denklem 4.15' de verilmiştir.

$$Q_t = \sigma(W_O \cdot [h_{t-1}, X_t] + b_O) \quad (4.15)$$

Son kısım olan hücre durumu ise tahmin yapabilmek için diğer kapılar boyunca ortaya çıkan anlamlı bilgilerin taşındığı bir hat olarak değerlendirilebiliriz. Hücre durumu sayesinde kısa süreli bellek problemi ortadan kaldırılır ve önceki veriler ağ zincirleri boyunca ileriki zaman dilimlerine taşınır.

Aşağıda gösterilen Şekil 4.12' de LSTM algoritmasının yapısı bakımından temel hücre blok diyagram modeli gösterilmiştir.



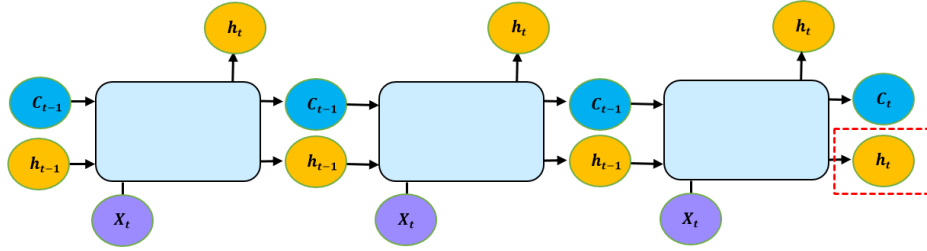
Şekil 4.12. LSTM algoritmasının temel hücre blok diyagram modeli [44]

Tez çalışmasındaki verilerin sürüş sırasından belirli zaman aralıklarında araç üzerindeki OBD2 soketinden CANBus veriyolu ile toplanırken, sıralı gelmesi ve bir önceki adımda toplanan veri satırının bir sonraki adımdaki veri satırının olan ilişkisinin bulunması çalışmada LSTM çalışma yapısına uygun olacağını hakkında öngörüler vermiştir. Yukarıda verilen Tablo 4.1' de görüldüğü gibi araç üzerinden sensör verileri, sürüş sırasında düzenli olarak belirli zaman periyotlarında sürücünün ve çevre şartlarının tepkilerine göre kayıt altına alınmaktadır.

Tez çalışması sırasında LSTM modeli ile ilgili testler yapılırken LSTM'ni kendi içinden bulunan üç adet alt modeli üzerinde gerekli araştırmalar yapıp bu alt modellerden veri setine

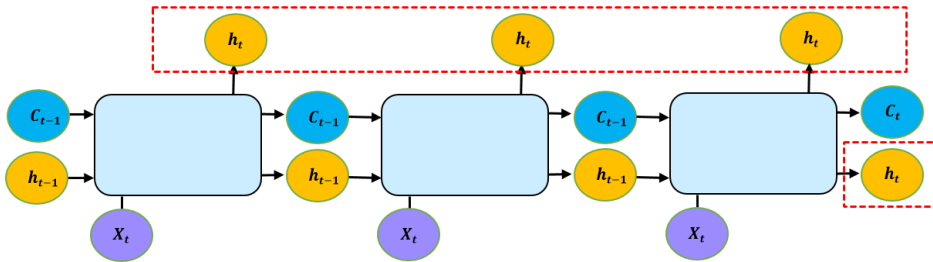
uygun olan iki modelin testler yapılmıştır. LSTM'nin kendi içinde bulunan üç alt modellerini sırasıyla inceleyecek olursak;

LSTM'nin birinci kullanım şeklinde hücre var olan kendini tekrarlamalar sonucunda elde ettiği en son gizli durumun (Hidden State) son gizli durum (Hidden Last State) adı verilen son çıkışının kullanılmasıyla oluşan bir çalışma biçimidir. Bu çalışma biçimin bir adet sayısal bir veri sınıflandırıcı tarafından çıkış olarak verilir. Aşağıdaki Şekil 4.13' de LSTM' in birinci çalışma modelinin temsili ve verdiği son veri gösterilmiştir.



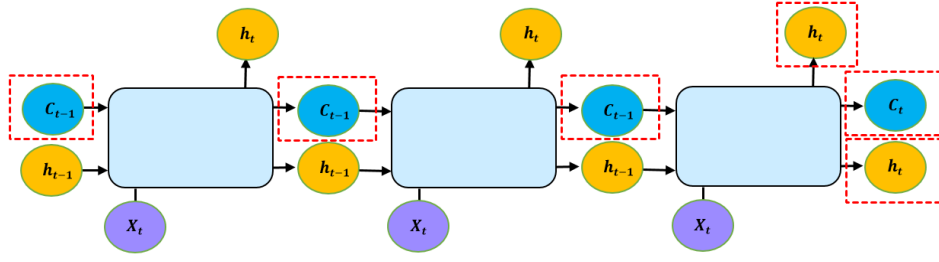
Şekil 4.13. LSTM modelinin birinci tip çalışma şekli [44]

İkinci kullanım şeklinde ise birinci kullanıma göre çok daha fazla çıkışı bize üretmesidir. Çalışma mantığına baktığımızda her zaman diliminde hücreye gelen yeni verinin model üzerinde işlenmesiyle oluşan tüm gizli durumların hafıza saklanması ve ileriki durumlarda bu veriler kullanılarak sınıflandırma işleminin ilgi veri setinin yapısına göre daha iyi performanslar elde etmesi sağlanabilmektedir. Bu çalışma şeklinde Şekil 4.14' de verilen örnek gösterimde bize dört adet çıkış verisi verilir. Aşağıdaki Şekil 4.14' de LSTM' in ikinci çalışma modelinin temsili ve verdiği son veri gösterilmiştir.



Şekil 4.14. LSTM modelinin ikinci tip çalışma şekli [44]

Son çalışma şekli olan üçüncü tip çalışmada LSTM' de, her zaman diliminde hücrenin oluşturduğu tüm hücre durumlarını (Cell State) ve son üretilen gizli durumu (Hidden Last State) in hepsi sınıflandırıcı tarafından çıkış olarak verilmektedir. Bu çalışma şekli daha çok decoder-encoder gerektiren problemlerin çözümünde kullanılmaktadır. Aşağıdaki Şekil 4.15' de LSTM' in ikinci çalışma modelinin temsili ve verdiği son veri gösterilmiştir.



Şekil 4.15. LSTM modelinin üçüncü tip çalışma şekli [44]

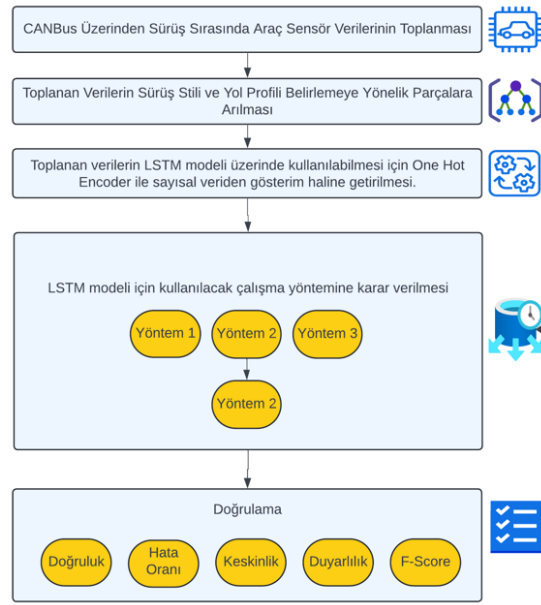
Tez çalışmasında kullanılan çalışma modeli için bu üç modelden ikinci olan çalışma tipi seçilmiştir. Üçüncü model çalışma yapısı olarak decoder-encoder problemleri üzerinde uygulanmakta olduğundan birinci ve ikinci çalışma tipi var olan veri setimiz için en uygun olan çalışma biçimi olarak gözükmemektedir. Değerlendirmeye alınan birinci ve ikinci çalışma tipleri var olan veri setleri üzerinde testler yapılmasıyla çıkan sonuçlara göre ikinci modelin tez çalışmasına daha uygun olduğuna karar verilmiştir. Yapılan çalışma modeline ilişkin çıkan değerlendirme sonuçları Bölüm 5’ te ayrıntılı olarak verilmiştir.

LSTM’ de aynı yapay sinir ağları gibi sayısal verilerin direk ağ üzerinde işlenmesi bazı problemlere neden olmaktadır. Bu problemlerin başında parametre olarak yollanan verilerin değer skalaları arasındaki farklılıkların olması bir değerın başka bir değer baskın gelmesine neden olabilmektedir. Yapay sinir ağları üzerinde nasıl normalizasyon kullanılıyorsa LSTM’de de bun durumun önüne geçmek için en çok tercih edilen yapı olan One Hot Encoder kullanılır. Böylece değerler direk değil o değerlerin birer gösterimi şeklinde verilir modele. Aşağıdaki tabloda elimizdeki değerlerden bazılarının hem sayısal hem de One Hot Encoder edilmiş hali aşağıdaki Tablo 4.5’ te gösterilmiştir.

Tablo 4.5. MAF Sayısal değerlerin ve One Hot Encoder değer karşılıkları gösterimi

MAF Sayısal	One Hot Encoder
2	[010000000000]
3	[000100000000]
11	[000000000001]
6	[000000100000]
8	[000000001000]
9	[000000000100]
10	[000000000010]

Aşağıdaki Şekil 4.16’ da LSTM’nin tez çalışması süresince sınıflandırma sürecini gösterir süreç diyagramı verilmiştir.



Şekil 4.16. LSTM modelinin süreç işleyiş diyagramı

4.5. Verilerin Doğrulanması

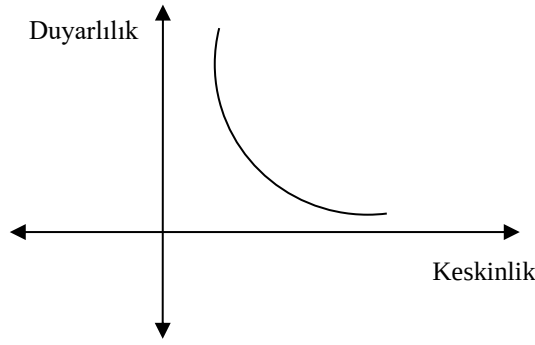
Veri seti olarak seçilen değişkenler arasındaki ilişki kolayca belirlenemediğinden ve veri sayısının fazlalığından dolayı, veriler arasındaki ilişkiyi analiz etmek ve sınıflandırmak için çeşitli sınıflandırma algoritmaları kullanılmıştır. Bu algoritmalar, mevcut veri setleri üzerinde eğitilerek, daha sonra gelen yeni verilere göre sürüş stili ve yol profili çeşitlerinin sınıflandırmalarını gerçekleştirmektedir [42].

Yapay sinir ağları kullanılarak yapılan tahmin performansı, Denklem 4.16' da gösterilen MSE (Hataların Karelerinin Ortalaması) metriğine göre değerlendirilmektedir. MSE, tahmin edilen değerler ile gerçek değerler arasındaki farkların karelerinin ortalamasını temsil eder. Formül sonucu elde edilen değer ne kadar küçük olursa, sistemin o kadar doğru tahmin işlemi gerçekleştirdiği kabul edilir. Bu metrik, modelin performansını ölçmek ve iyileştirmek için kritik bir göstergedir.

$$MSE = \frac{1}{k} \sum_{i=1}^k (y_i - \tilde{y}_i)^2 \quad (4.16)$$

Yukarıdaki denklemde; MSE: Hataların karelerinin ortalaması, k toplam veri sayısı, y_i i. andaki değer, $\tilde{y}_i$ i. andaki tahmin edilen değer

Tüm eğitici sınıflandırma algoritmaları olan Naive Bayes, SVM, KNN, C4.5, Derin Öğrenme için de başarıların doğrulamasını ise duyarlılık ve keskinlik değerlerine göre hesaplanabilmektedir. Duyarlılık ve Keskinlik değerleri arasında ilişki Şekil 4.17' de gösterildiği gibi bir parabolik bir ters ilişki bulunmaktadır.



Şekil 4.17. Duyarlılık ve keskinlik değerleri arasındaki ilişki [42]

Şekil 4.17’ de görüldüğü gibi, bir değer duyarlılık sonucu ile aynı değerin keskinlik sonucu ters orantılı olmasından dolayı yüksek bir duyarlılık sonucu düşük bir keskinlik sonucunu ortaya çıkabilmektedir. Bu durum bize başarı için kesin bir kıyaslama yapabileceğimiz bir sonuç vermemektedir. Bu nedenle, hesaplanan bu değerlerin harmonik ortalaması alınarak F-Score değeri hesaplanmakta ve bu değere göre başarı ölçülmektedir.

Tablo 4.6’ de, veri sınıflandırma sonuçlarının (Karışıklık Matrisi) tablo görünümü sunulmaktadır. Ayrıca, (4.17) numaralı denklem doğruluk oranını, (4.18) numaralı denklem hata oranını, (4.19) numaralı denklem keskinlik değerini, (4.20) numaralı denklem duyarlılığı ve (4.21) numaralı denklem ise F-Score değerini hesaplamaktadır. Bu denklemler, sınıflandırma performansının farklı açılardan değerlendirilmesine olanak tanır ve elde edilen sonuçların daha kapsamlı bir şekilde analiz edilmesini sağlar.

	Öngörülen (Hayır)	Öngörülen (Evet)	
Gerçek (Hayır)	TN	FP	TN + FP
Gerçek (Evet)	FN	TP	FN + TP
	TN + FN	FP + TP	

Tablo 4.6. Veri sınıflarının sonuç tablosu (Karışıklık Matrisi)

Yukarıdaki Tablo 4.6’ deki terimlerden TP True Positive, FN False Negative, FP False Positive, TN True Negative ‘yi göstermektedir.

$$\text{Doğruluk Oranı: } \frac{TP+TN}{TP+FP+FN+TN} \quad (4.17)$$

$$\text{Hata Oranı: } \frac{FP+FN}{TP+FP+FN+TN} \quad (4.18)$$

$$\text{Keskinlik: } \frac{TP}{TP+FP} \quad (4.19)$$

$$\text{Duyarlılık: } \frac{TP}{TP+FN} \quad (4.20)$$

$$\text{F-Score: } \frac{2 \cdot \text{Duyarlılık} \cdot \text{Keskinlik}}{\text{Duyarlılık} + \text{Keskinlik}} \quad (4.21)$$

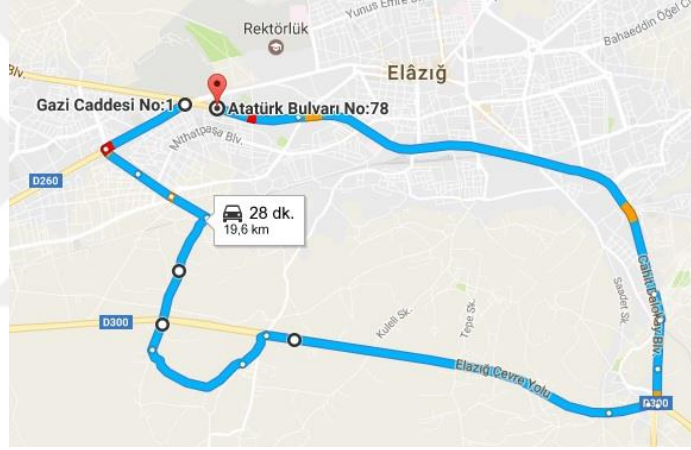


5. BULGULAR VE TARTIŞMA

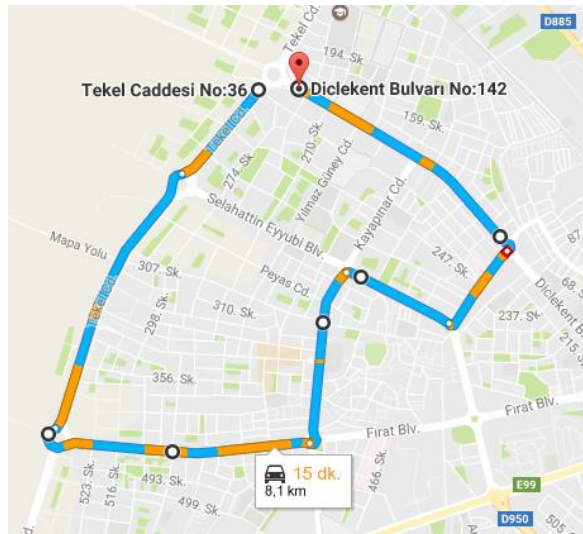
5.1. CANBus Üzerinden Verilerin Toplanması

Tez çalışması için veri toplama süreci, Şekil 5.1 ve Şekil 5.2’ de verilen güzergah üzerinde, 2 farklı araç sürücüsünün ikişer tur attırılmasıyla gerçekleştirilmiştir. Sürücülere bu güzergah üzerinde literatürde tanımlanan şekilde hem agresif hem de sakin sürüş tarzlarına göre her seferinde birini kapsayacak şekilde sürdürülmüş ve bu şekilde her sürücü için agresif ve sakin sürüş verileri CANBus üzerinden ayrı ayrı toplanmıştır.

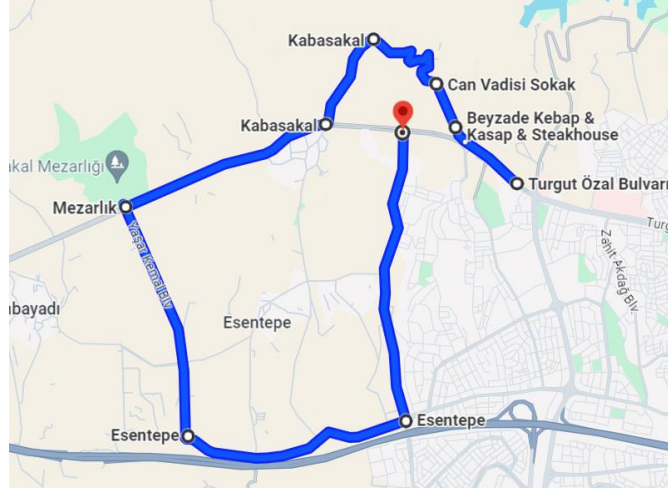
Diğer problemimiz olan yol profili belirlemeye yönelik olarak Şekil 5.3’ de verilen güzergah üzerinde 2 farklı araç sürücüsünün ikişer tur attırılmasıyla gerçekleştirilmiştir. Yine sürücüler yol profili belirlemede düz yol ve bozuk yol güzergahlarını barındıran rotada sürüşlerini gerçekleştirirken CANBus üzerinden anlık olarak veriler toplanmıştır.



Şekil 5.1. Sürüş stili verisi toplama güzergahı 1



Şekil 5.2. Sürüş stili verisi toplama güzergahı 2



Şekil 5.3. Yol profili verisi toplama güzergahı

Bu Tez çalışmanın amacı, aracın OBD 2 portu üzerinden elde edilen seçili araç sensör verilerini toplayarak sürücülerin sürüş stillerinin ve araçlarını sürdükleri yolların profillerinin eğitici sınıflandırma algoritmalarıyla belirlemek ve bu sürüş tarzlarının sınıflandırma algoritmaları üzerindeki tahmin performanslarını değerlendirmektir. Veri toplama işlemi, belirlenen bir güzergahlar üzerinde agresif ve sakin sürüşlerin ayrı ayrı yapılmasıyla gerçekleştirilmiş ve iki ayrı veri tablosu oluşturulmuştur. Gene yol profilleri içinde tek seferde bozuk ve düz yol profillerini barındıran güzergahlarda veri toplama işlemi gerçekleştirilmiştir.

Bu sürüşler sırasında araç OBD2 portundan kayıtları toplanan tüm sensör verileri, sınıflandırma işlemi için gerekli olan verileri tam olarak karşılamamaktadır. Bu nedenle, gelen tüm sensör verileri içinden sınıflandırma sonucunu en yüksek doğrulukta verecek en az sayıda ve en ilişkili parametreler seçilmiştir. Veri toplama işlemi sırasında sensör verileri olarak MAF (hava akış oranı), hesaplanan yük, motor devri, anlık ekonomi ve araç hızı gibi veriler seçilmiştir. Bu değişkenlerin, agresif ve sakin sürüşlerin yanı sıra düz ve bozuk yol profili belirleme sürüşleri sırasında sürekli olarak değişkenlik gösterdiği belirlenmiştir. Diğer sensör verileri, sürülen zaman, sürülen mesafe, akü voltajı, toplam tüketim, ortalama ekonomi vb., sürüş stilleri ve yol profilleri hakkında doğrudan sınıflandırma yapmaya elverişli olmadığından ve lineer artış gösteren değerler olduklarından dolayı hesaplama işlemlerine dahil edilmemiştir. Tablo 5.1’ de yukarıdaki şekillerde verilen üç güzergah üzerinde sürüş sırasında toplanan sürüş stili belirleme sensör verileri aynı şekilde Tablo 5.2’ de ise yol profili belirleme sensör verileri detaylı bir şekilde sunulmaktadır. Tablolarda 1 ile ifade edilenler agresif sürüş ile Bozuk yolları, 0 ile ifade edilenler sakin sürüş ile düz yolları temsil etmektedir. Buradaki veriler herhangi bir normalizasyon işlemine tabi tutulmamış sensörler üzerinden gelen ham verilerdir.

Tablo 5.1. Sürüş stili eğitim veri setlerinin işlenmemiş hali

Toplanan Veri Türleri						
Nu	Sürüş Tipi	MAF Verisi	Hesaplanan Yük	Motor Devri	Anlık Ekonomi	Hızı
1	1	6,73	86,8	2156	11,5	41,8
2	1	7,14	99,6	2311	11,9	48,9
.	1	.	.	.	.	.
.	1	.	.	.	.	.
.	1	.	.	.	.	.
471	1	5,54	77,8	2370	11,5	48,5
472	1	1,87	15,9	861	1	0,6
473	0	0,94	19,7	905	1,7	1,6
.	0	.	.	.	.	.
.	0	.	.	.	.	.
.	0	.	.	.	.	.
1102	0	0,95	15,9	865	2	4,2
1103	0	0,67	14,9	961	1	0,6

Tablo 5.2. Yol profili eğitim veri setlerinin işlenmemiş hali

Toplanan Veri Türleri						
Nu	Yol Tipi	MAF Verisi	Hesaplanan Yük	Motor Devri	Anlık Ekonomi	Hızı
1	1	5,83	76,4	2366	11,5	46,1
2	1	6,44	89,2	2531	12,4	53,5
.	1	.	.	.	.	.
.	1	.	.	.	.	.
.	1	.	.	.	.	.
366	1	6,51	86,8	2370	11,5	48,5
367	1	3,82	35,2	861	1	0,6
368	0	0,64	91,7	1105	4,7	32,6
.	0	.	.	.	.	.
.	0	.	.	.	.	.
.	0	.	.	.	.	.
841	0	1,75	14,9	765	2	3,2
842	0	0,98	12,6	962	3,1	2,6

Yukarıdaki tablolarda görülen verilerin toplanması işlemini araçlarımız üzerinde bulunan OBD2 portu ile CANBus veri yolundan sağlamaktayız. Her araçta CANBus verileri standart değildir. Araçların marka, model ve hatta paket farklılıklarına göre bile değişebilmektedir. Yukarıdaki tabloda verilen beş adet sensör çeşidi tüm araç modellerinde rastlanabilecek en çok kullanılan sensörlerin verisidir. Bu veriler araçlarımızda bulunan ve genellikle sürücülerin rahatlıkla ulaşabileceği yerlerde bulunan OBD2 soketi yardımıyla toplanmaktadır. Yukarıdaki tablolarda bulunan sensörlerin veri toplama aşamalarına ait resimler farklı araç modellerinde Şekil 5.4’ de gösterilmiştir.



Şekil 5.4. Farklı araçlar üzerinden kayıt verilerinin toplanması

5.2. Veri Normalizasyonu

Tezde kullanılan sınıflandırma algoritmalarımız arasında yer alan yapay sinir ağıları, sınıflandırmada verilerin eğitim, doğrulama ve test aşamalarına tabi tutulmadan önce normalize edilmesi gereken bir algoritma türüdür. Bu normalizasyon işlemi, veri seti içinde bulunan değerler arasındaki büyüklüklerinden kaynaklı birbirine olan baskınlıkları ortadan kaldırmakta ve sistemin daha dengeli çalışmasını sağlamaktadır. Ayrıca, normalizasyon, hatalı veri değerlerinin sistemin performansı üzerindeki olumsuz etkilerini minimize etmek amacıyla da büyük önem taşımaktadır. Yukarıda Denklem 4.1 ile gösterilen Min-Max normalizasyon yöntemi, ham verilerin değerlerini +1 ile 0 aralığına dağıtmaktadır. Bu yöntem, verilerin daha homojen bir dağılım göstermesini sağlarken, aynı zamanda algoritmanın daha doğru ve güvenilir sonuçlar üretmesine olanak tanımaktadır. Min-Max normalizasyonu veri setindeki en yüksek ve en düşük değerler arasındaki farkı kullanarak tüm verileri belirli bir aralığa sıkıştırmakta ve bu sayede veri setindeki tüm değerler eşit bir öneme sahip olmaktadır.

Aşağıdaki Tablo 5.3' de, hem agresif hem de sakin sürüş verilerinin normalize edilmiş halleri ile Tablo 5.4' de hem düz hem de bozuk yol verilerinin normalize edilmiş halleri karışık bir şekilde sunulmuştur.

Tablo 5.3. Sürüş stili eğitim veri setlerinin normalize edilmiş hali

Toplanan Veri Türleri						
Nu	Sürüş Tipi	MAF Verisi	Hesaplanan Yük	Motor Devri	Anlık Ekonomi	Hızı
1	1	0.94	0.85	0.86	0.93	0.85
2	1	0.96	0.89	0.93	0.95	0.91
.	1	.	.	.	.	.
.	1	.	.	.	.	.
.	1	.	.	.	.	.
471	1	0.75	0.74	0.95	0.96	0.99
472	1	0.19	0.14	0.2	0.4	0.4
473	0	0.04	0.06	0.03	0.06	0.05
.	0	.	.	.	.	.
.	0	.	.	.	.	.
.	0	.	.	.	.	.
1102	0	0.04	0.06	0.1	0.09	0.07
1103	0	0.2	0.4	0.07	0.2	0.3

Tablo 5.4. Yol Profili eğitim veri setlerinin normalize edilmiş hali

Nu	Yol Tipi	Toplanan Veri Türleri				
		MAF Verisi	Hesaplanan Yük	Motor Devri	Anlık Ekonomi	Hızı
1	1	0.91	0.82	0.84	0.91	0.84
2	1	0.92	0.87	0.90	0.94	0.92
.	1	.	.	.	.	.
.	1	.	.	.	.	.
.	1	.	.	.	.	.
471	1	0.85	0.81	0.89	0.96	0.99
472	1	0.27	0.24	0.19	0.5	0.4
473	0	0.10	0.16	0.08	0.08	0.09
.	0	.	.	.	.	.
.	0	.	.	.	.	.
.	0	.	.	.	.	.
1102	0	0.19	0.1	0.13	0.14	0.12
1103	0	0.7	0.4	0.09	0.22	0.36

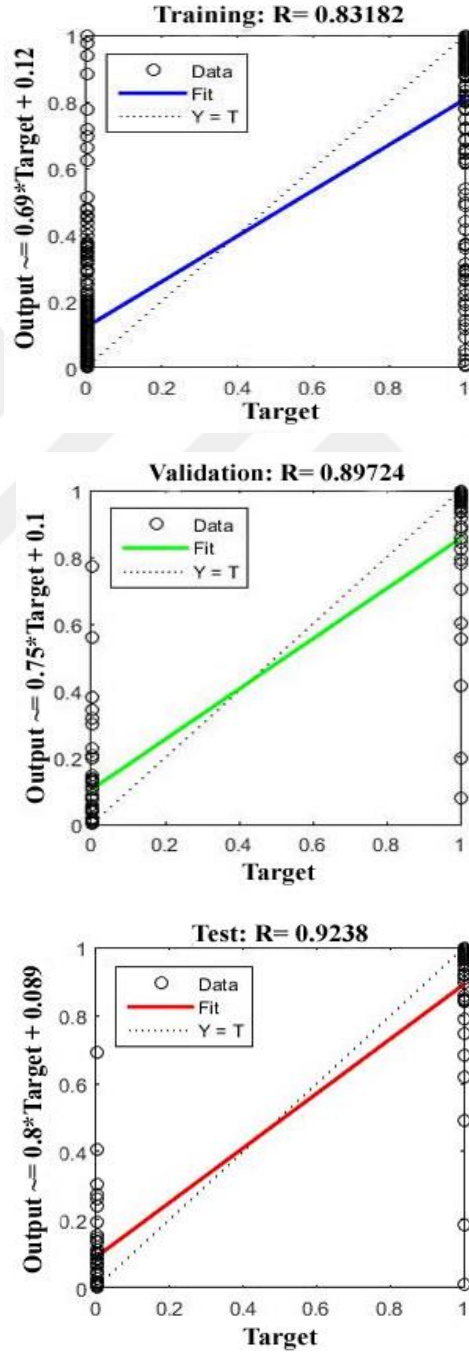
Yukarıdaki tablolarda, "Sürüş Tipi" sütunu, veri seti toplanırken kayıt altına alınan sürüş türlerini göstermektedir. Agresif sürüş verileri "1" ile sakin sürüş verileri ise "0" ile temsil edilmiştir. Aynı şekilde diğer tablodaki "Yol Tipi" sütunu, veri seti toplanırken kayıt altına alınan yol türlerini göstermektedir. Bozuk yol verileri "1" ile, düz yol verileri ise "0" ile temsil edilmiştir. Böylece, modelin eğitimi sırasında hangi verilerin agresif, hangilerinin ise sakin sürüşe veya verilerin bozuk, hangilerinin ise düz yol profiline ait olduğu açıkça anlaşılabilmektedir ve bu da modelin daha doğru tahminler yapmasını sağlamaktadır.

5.3. Yapay Sinir Ağı Üzerinde Öğrenme, Doğrulama ve Test Adımları

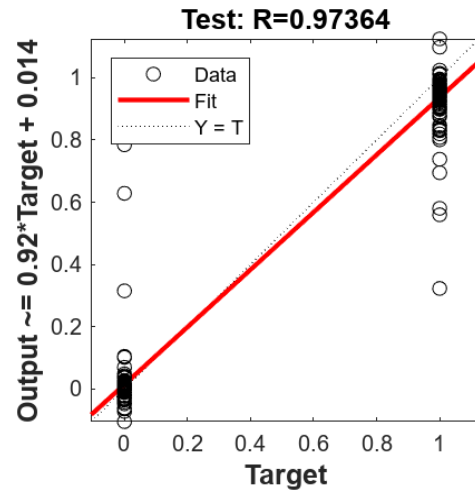
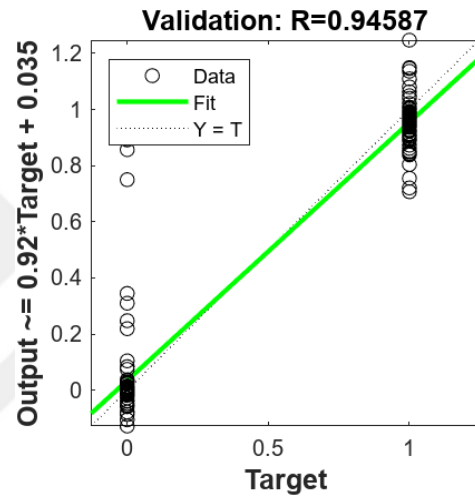
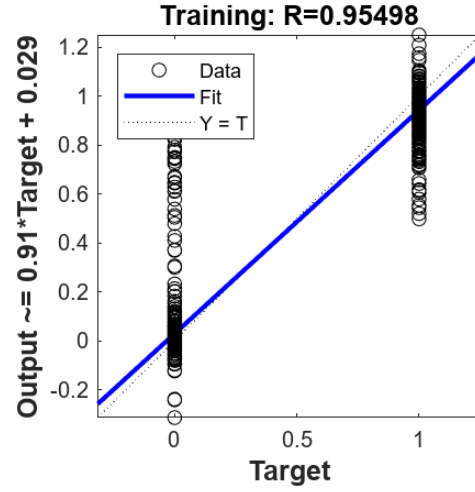
Araçtan elde edilen verilerin okunması, seçilmesi ve normalizasyon işlemlerinden geçirilmesinin ardından, yapay sinir ağının tasarımına karar verilmiştir. Korelasyon hesaplamaları, Denklem 4.2 kullanılarak yapılmış ve bu hesaplamalar sonucunda yapay sinir ağlarında kullanılacak transfer fonksiyonunun seçimi aşamasına geçilip. Şekil 4.4 ve Şekil 4.5’ de gösterildiği gibi, korelasyon performanslarına dayanarak en uygun olan Bipolar Sigmoid Transfer Fonksiyonu (TANSIG) seçilmiştir.

Yapay sinir ağıımızda gizli katman sayısının belirlenmesinde, ezberleme ve performans hızı gibi faktörler dikkate alınmıştır. Katman sayısı, fazla sayıda katman içermeyen ve hızlı sonuçlar alabileceğimiz bir şekilde belirlenmiştir. Yukarıda, Şekil 4.6 ve Şekil 4.7’ de tasarlanan yapay sinir

ağı modelinin şekli verilmiştir. Aşağıda yer alan Şekil 5.5’ te sürüş stilleri için Şekil 5.6’ da yol profilleri için tasarladığımız yapay sinir ağı modelinin öğrenme, doğrulama ve test değerlerinin dağılımı gösterilmektedir. Şekil 5.5’ te ve Şekil 5.6’ da “Data” olarak belirtilen şekiller, verilerimizi temsil etmektedir. “Y=T” olarak gösterilen kesikli çizgi, en başarılı sonuçları veren köşegen eksenini temsil ederken, “Fit” olarak belirtilen doğru ise, çalışmamız sonucunda elde ettiğimiz başarıyı göstermektedir. Fit doğrusu köşegene ne kadar yakın ve paralelse, elde edilen sonuçların o kadar başarılı olduğunu ifade etmektedir.



Şekil 5.5. Sürüş stili YSA performan değerleri grafikleri



Şekil 5.6. Yol profili YSA performan değerleri grafikleri

5.4. Derin Öğrenme

Derin öğrenme modelleri üzerinde tez için uygun modelin seçiminde var olan veri setinin için en uygun olan model olarak karar verme sürecinde LSTM'nin en uygun olan karar verme mekanizması olduğuna kesinleşmişti. LSTM çalışma yapısı ve sınıflandırma sırasında önceki zaman birimlerinde gelen verilerden elde ettiği sonuçları saklayabilmesi ve sonraki zaman birimlerinden sınıflandırma sürecine dahil edebilmesi var olan başarı oranının artmasına olumlu katkıda bulunmuştur.

LSTM'nin çalışma şekilleri olarak üç türde çalıştığı konusuna değinilmişti ve bu çalışma biçimlerinden üçüncü tipin var olan veri topluluklarını ile sınıflandırma amacımıza uygun olmadığından birinci ve ikinci tip çalışma modelleri üzerinden gidilmesine karar verilmiştir. Tez çalışması sırasında elde var olan veri topluluklarında hem agresif/sakin sürüş stili belirlemede hem de bozuk/düz yol profili belirlemede her iki örnek için var olan verilerle test sınıflandırmaları yapılmış bunların içinde başarı oranları kendi içinde değerlendirilmiştir. Aşağıdaki Tablo 5.5' te LSTM modelinin hem birinci tip hemde ikinci tip çalışma şekillerine göre performanslarının değerlendirmesi verilmiştir.

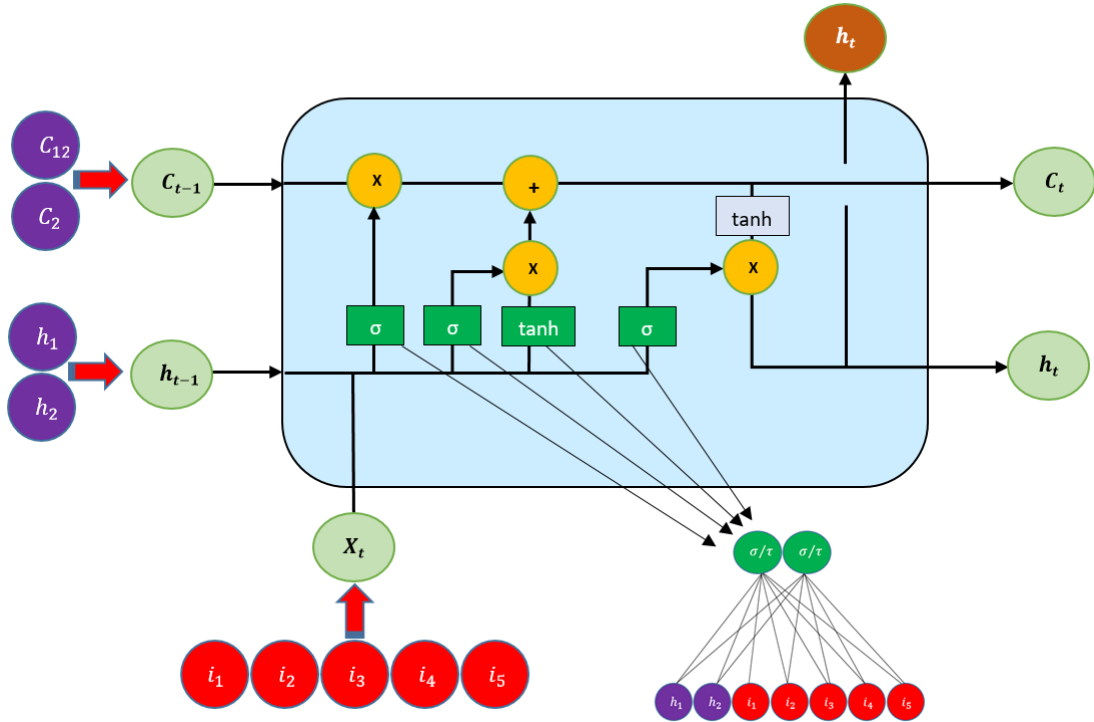
Tablo 5.5. LSTM'nin çalışma tiplerinin başarılarının karşılaştırılması

Veri Seti	LSTM Çalışma Tipi	MAE	MSE	RMSE	MAPE
Yol Profil	Tip 1	0,0593	0,0593	0,2435	nan
Yol Profil	Tip 2	0,0247	0,0247	0,1571	nan
Sürüş Stili	Tip 1	0,0652	0,0042	0,0635	nan
Sürüş Stili	Tip 2	0,0477	0,0476	0,2185	nan

Bu testler ile en iyi performansı veren epoch, activation, optimizer ve metrics içerinden elde ettikleri sonuçları MAE (Ortalama Mutlak Hatası), MSE (Ortalama Kare Hatası), RMSE (Kök Ortalama Kare Hatası), MAPE (Ortalama Mutlak Yüzde Hatası) değerleri alınarak karşılaştırma yapılmıştır. Hesaplama ile ilgili denklemler sırasıyla MSE Denklem 4.9' da, RMSE Denklem 4.10' da, MAE Denklem 4.11' de, MAPE Denklem 4.12' de verilmiş ve karşılaştırma hesaplamaları bu sonuçlara göre yapılmıştır. Yukarıdaki Tablo 5.5' de LSTM'nin çalışma tiplerine göre performans değerlendirmelerindeki çıkan sonuca göre her iki problemimiz içinde çalışma tipi ikinin en yüksek performans verdiği gözükmemektedir.

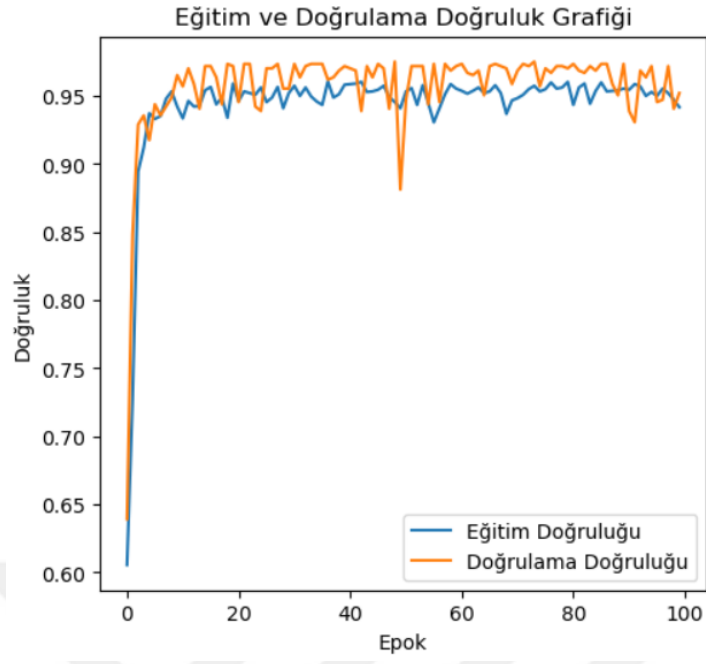
Tez çalışmasında tasarlanmış olan LSTM modelinin iç tasarımı Şekil 5.7' de verilmiştir. Tasarımı inceleyecek olduğumuzda her zaman diliminde LSTM hücresi parametreleri olan Cell State(C_{t-1}) ve Hidden State(h_{t-1}) değerleri iki boyutlu vektörel değerler kullanılmıştır. Aynı şekilde Input(X_t) değeri veri seti grubumuzda olan beş adet parametrenin girişinde kaynaklı olarak beş boyutlu olarak kullanılmıştır. LSTM'nin iç kısmına girdiğimizde ise üç adet Sigmoid, bir adet

Tangenth Network'ün olduğu görülmektedir. Bu networkler Hidden State(h_{t-1}) ve Input(X_t) verilerinin birleşmesiyle oluşmaktadır. Oluşan networkler şeklide Şekil 5.7 üzerinde belirtilmiştir. Tez için en başarılı olan çalışma tipi olarak seçilen ikinci tip çalışma biçimi gereği her zaman diliminde oluşan Hidden State(h_t) değerleri hafızada saklanmakta sonraki işlemlerde geriye dönük olarak kullanılabilir. Burada öğrenme işlemi aynı sürecin var olan tüm Timesteps'lerinde yaşanmaktadır. Hücre içindeki tüm işlemler bitirildikten sonra oluşan 2 boyutlu Cell State(C_t) ve Hidden State(h_t) değerleri bir sonraki zaman adımının Cell State(C_{t-1}) ve Hidden State(h_{t-1}) giriş değerleri olarak kullanılmaktadır.

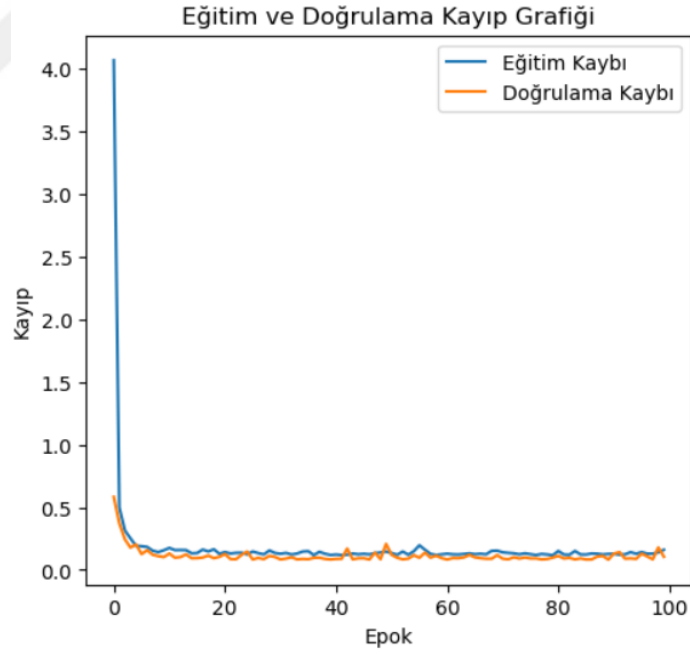


Şekil 5.7. Tez için tasarlanan LSTM modelinin iç yapısı ve parametreleri [44]

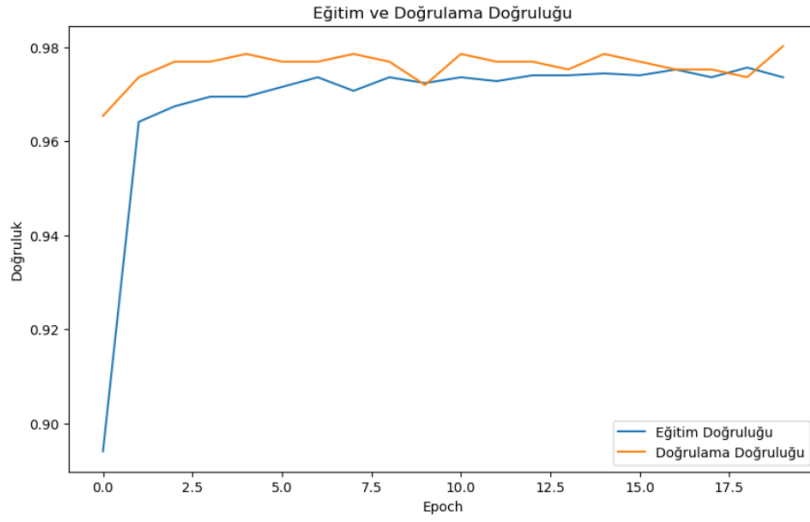
Tez çalışmasında derin öğrenme algoritması olarak seçtiğimiz LSTM modelinin sınıflandırma performans değerlemesi için yapılmış olan testler sonucunda sürüş stili sınıflandırmasındaki eğitim-doğrulama doğruluğu Şekil 5.8' de, eğitim-doğrulama kaybı grafiği de Şekil 5.9' da verilmiştir. Gene aynı şekilde yol profili belirleme sınıflandırmasındaki eğitim doğrulama doğruluğu Şekil 5.10' da, eğitim doğrulama kaybı grafiği de Şekil 5.11' de verilmiştir.



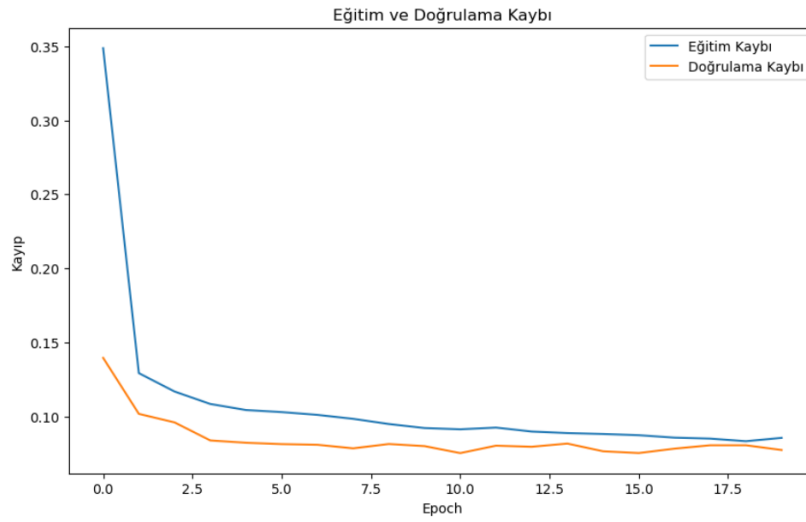
Şekil 5.8. Sürüş stili eğitim ve doğrulama doğruluğu grafiği



Şekil 5.9. Sürüş stili eğitim ve doğrulama kaybı grafiği



Şekil 5.10. Yol profili eğitim ve doğrulama doğruluğu grafiği



Şekil 5.11. Yol profili eğitim ve doğrulama kaybı grafiği

5.5. YSA, KNN, Navie Bayes, SVM, C4.5, Derin Öğrenme Modellerinin Başarıları

Yukarıda Tablo 4.6' de verilen Veri Sınıflandırma Sonuç Tablosu (Karışıklık Matrisi) gösterilmekte olup bu tablo üzerinden verilen formatta verileri yerleştirdiğimizde çıkan değerleri sırasıyla Denklem 4.17, 4.18, 4.19, 4.20 ve 4.21 uyguladığımızda sonuçların başarımları hakkında yorum yapabiliriz. Aşağıdaki Tablo 5.6' da sürüş stilleri için Tablo 5.7' de ise yol profilleri için verdiğimiz karışıklık matrisi değerlerinin sayısal karşılıkları verilmiştir. Aşağıdaki bu tablolarda veriler oluşturulurken mevcut veri setinin %90'sı öğrenme verilerine ayrılırken %10' test verilerine ayrılmasıyla oluşan değerlere göre hazırlanmıştır.

Tablo 5.6. Sürüş stili karışıklık değerleri tablosu

Algoritmalar	Değerler			
	TN	FN	FP	TP
YSA	287	34	17	337
KNN	253	51	51	320
Navie Bayes	287	68	17	303
SVM	253	51	51	320
C4.5	270	51	51	320
LSTM	444	13	19	196

TN : True Negative

FN : False Negative

FP : False Positive

TP : True Positive

Tablo 5.7. Yol profili karışıklık değerleri tablosu

Algoritmalar	Değerler			
	TN	FN	FP	TP
YSA	1998	69	22	944
KNN	2009	58	230	736
Navie Bayes	1995	72	68	898
SVM	2002	65	30	936
C4.5	1999	68	15	951
LSTM	1857	52	43	1081

TN : True Negative

FN : False Negative

FP : False Positive

TP : True Positive

Yukarıdaki Tablo 5.7’de ve Tablo 5.8’ de karışıklık matris değerlerinin Denklem 4.17, 4.18, 4.19, 4.20 ve 4.21 uygulanması sonucu oluşan ölçüt değerleri gösteren sürüş stilleri için Tablo 5.8’ de ve yol profilleri için Tablo 5.9’ daki değerlere göre algoritmaların başarılarının sonuçlarını ve kendi içlerindeki kıyaslamalarını yapabiliriz.

Tablo 5.8. Sürüş stili model başarı değerleri tablosu

Algoritmalar	Ölçütler				
	Doğruluk	Hata Oranı	Keskinlik	Duyarlılık	F-Score
YSA	0,924	0,076	0,952	0,908	0,929
KNN	0,849	0,151	0,863	0,863	0,863
Navie Bayes	0,924	0,076	0,952	0,908	0,929
SVM	0,874	0,126	0,947	0,817	0,877
C4.5	0,849	0,151	0,863	0,863	0,863
LSTM	0,952	0,048	0,912	0,938	0,952

Tablo 5.9. Yol profili model başarı değerleri tablosu

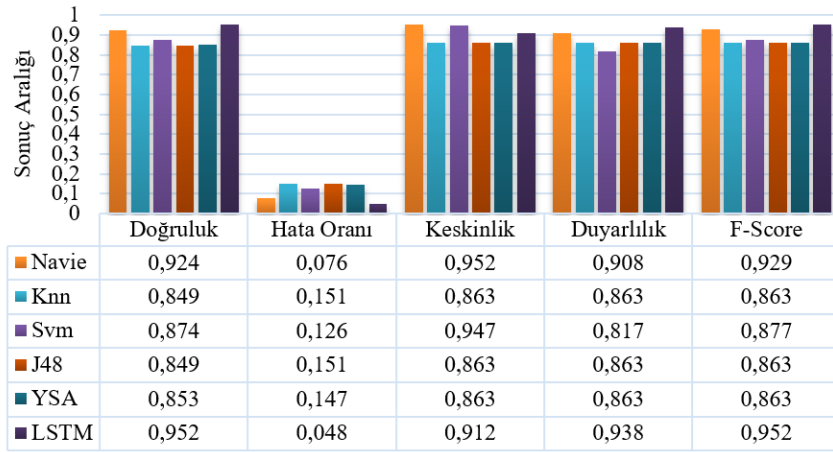
Algoritmalar	Ölçütler				
	Doğruluk	Hata Oranı	Keskinlik	Duyarlılık	F-Score
YSA	0,97	0,030	0,977	0,932	0,954
KNN	0,905	0,095	0,762	0,927	0,836
Navie Bayes	0,954	0,046	0,93	0,926	0,928
SVM	0,969	0,031	0,969	0,935	0,952
C4.5	0,973	0,027	0,984	0,933	0,958
LSTM	0,977	0,030	0,963	0,957	0,964

5.6. Doğrulama İşlemleri

Tez çalışmasında tasarlanan sistemin doğrulanması prosedürü, eğitimler için kullanılan verilerden farklı olarak elde edilen agresif, sakin sürüş ve düz, bozuk yol verileri kullanılarak gerçekleştirilmiştir. Bu veri grupları sürüş stili için agresif ve sakin sürüş olarak, yol profili içinde düz yol ve bozuk yol olarak ayrı ayrı kaydedilmiş ve gerekli normalizasyon işlemlerinden geçirildikten sonra sınıflandırma algoritmaları üzerinde test edilmiştir. Test işlemlerinin sonuçları Tablo 5.8 ve Tablo 5.9’ da sunulmuştur.

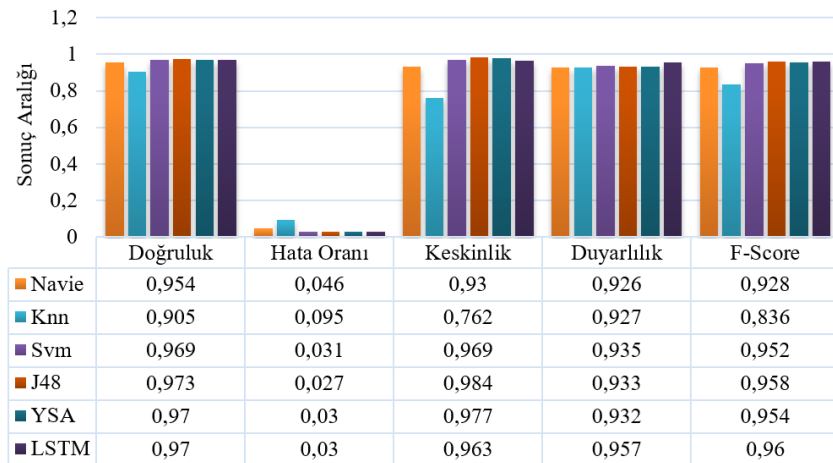
Keskinlik ve duyarlılık değerlerine göre yapılan yorumlar, matematiksel açıdan kesin bir yargıya varmamızı sağlamaktadır [42]. Şekil 4.17’ de görüldüğü gibi, bu iki değer arasındaki ters parabolik ilişki nedeniyle bir değer artması diğer değer değişken bir oranda azalmasına yol açmaktadır [42]. Bu nedenle, kesin bir karar verilmesi gerektiğinde, keskinlik ve duyarlılık değerlerinin harmonik ortalaması alınarak hesaplanan F-Score değeri (Denklem 4.21’ de gösterildiği gibi) kullanılmalıdır.

Buna ek olarak, Doğruluk Oranı da bu algoritmaların başarı seviyesi için bize yorum olanağı tanımaktadır. Doğruluk Oranı ve F-Score genellikle birbirine paralel sonuçlar verdiğinden, her iki metriğe göre de değerlendirme yapabilmekteyiz. Şekil 5.12’ de sürüş stilleri belirleme için, Şekil 5.13’ de ise yol profilleri belirleme için eğitici sınıflandırma algoritmaları üzerinde yapılan test sonuçlarının grafiksel olarak görebiliriz. Şekil 5.14’ de ise sürüş stilleri belirleme için, Şekil 5.15’ de ise yol profilleri belirleme için eğitici sınıflandırma algoritmalarımızın test sonuçlarına dayanan ROC eğrisi gösterilmektedir.



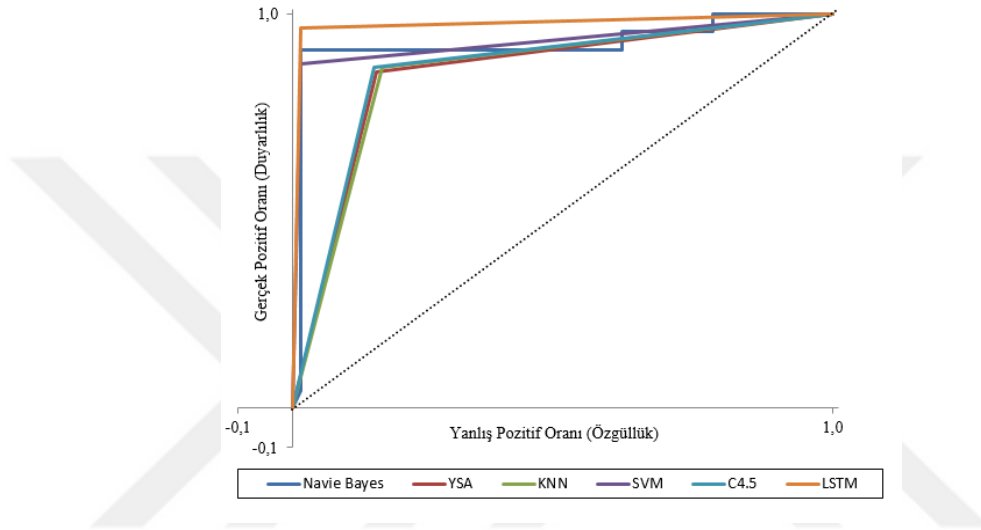
Şekil 5.12. Sürüş stilleri sınıflandırma algoritmaları performans grafiği

Yukarıdaki Şekil 5.12’ de gösterilen sürüş stilleri sınıflandırma algoritmalarının performans değerlendirmesinde çıkan sonuçlara göre, doğruluk, duyarlılık ve f score değerlerinden LSTM modelinin elimizdeki veriler üzerinde en yüksek sonuçları, hata oranında ise en düşük oranı verdiği görülmektedir. Grafikte dikkat çeken bir nokta olarak Navie Bayes keskinlik değerleri içinde sıyrılarak en yüksek başarıyı elde etmiştir. Keskinlik değer sınıflandırmada pozitif tahminlerin ne kadar doğru olduğunu ölçer ama tek başına bir ölçüt olarak kullanılamaz. Bu yüzden sınıflandırmada ölçüt olarak keskinlik ve duyarlılık değerlerinin harmonik ortalaması olan f score başarı ölçütü olarak kullanılmıştır. Bu bilgiler ışığında sürüş stili belirlemede en yüksek başarı değerinden en düşüğe doğru sırasıyla LSTM, Navie Bayes, SVM olarak gittiği sonra gelen J48(C4.5), YSA ve KNN algoritmalarının çok benzer başarılar elde ettiğini söyleyebiliriz.



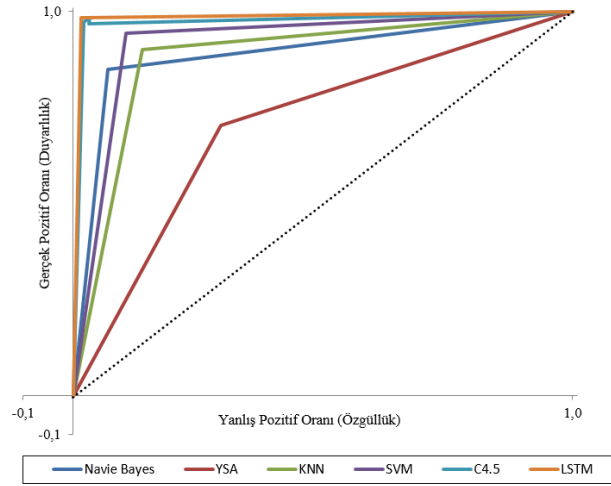
Şekil 5.13. Yol Profilleri sınıflandırma algoritmaları performans grafiği

Aynı şekilde yukarıdaki Şekil 5.13’ de gösterilen yol profiller sınıflandırma algoritmalarının performans değerlendirmesinde çıkan sonuçlara göre, duyarlılık ve f score değerlerinden LSTM modelinin elimizdeki veriler üzerinde en yüksek sonuçları, doğruluk adı verilen doğru tahminlerin toplam tahmine oranından YSA ile aynı başarıya ulaşmıştır. Hata oranında ise en düşük oranı J48(C4.5) algoritmasının verdiği ve aynı şekilde keskinlikte de en iyi başarıyı onun verdiği görülmektedir. Başarı değerlendirme kriterimiz olan f score değerlerine bakıldığında başarı sırası burada sırasıyla LSTM, J48, YSA, SVM, Navie Bayes ve KNN olarak sıralayabiliriz.



Şekil 5.14. Sürüş stilleri sınıflandırma algoritmaları test roc eğrisi

Şekil 5.14’ de gösterilen sürüş stilleri sınıflandırmaları için elde edilen roc eğrisini yorumlayacak olursak, roc eğrileri düzlemi ortadan kesikli çizgilerle kesen eksenin üst kısmında kalması başarının yüksek olduğunu altına doğru inmesi ise başarının düştüğünün göstergesidir. Bir roc eğrisinde $x=0$, $y=1$ doğrusunun devamında ise $y=1$ doğrusu ile devam etmesi en yüksek performansın alındığının göstergesidir. Bu çerçeveden yola çıkılınca $x=0$ $y=0$ noktasından çıkıp y eksenine en yakın olarak hareket eden doğrular sonuç olarak en başarılı olarak sınıflandırmaya sahip olduğunun göstergesidir. Bu bilgiler ışığında grafik ışığında başarı sırası LSTM, Navie Bayes, SVM olarak gittiği sonra gelen J48(C4.5), YSA ve KNN algoritmalarının çok benzer başarılar elde ettiğini söyleyebiliriz.



Şekil 5.15. Yol Profilleri sınıflandırma algoritmaları test roc eğrisi

Yol profili sınıflandırmaları sonucu elde ettiğimiz yukarıdaki Şekil 5.15’ deki Yol Profilleri sınıflandırma algoritmaları test roc eğrisinde, gene y eksenine en yakın olarak çizilen eğrilerin incelendiğinde başarı sıralamaları yorumlanabilmektedir. Grafikteki bilgiler ışığında doğru sınıflandırma başarıları sırasıyla LSTM, J48, YSA, SVM, Navie Bayes ve KNN olarak sıralayabiliriz.

6. SONUÇLAR

- 1) Eğitici sınıflandırma algoritmalarının araç üzerinden alınan sayısal sensör verileri yardımıyla gerek sürüş stili gerekse de yol profili belirlemede yüksek başarıları yüzdelere ulaştığı gözlenmiştir.
- 2) Sürüş stili belirlemede kullanılan eğitici sınıflandırma algoritmalarının agresif ve sakin sürüş stilleri üzerinde sırasıyla Navie Bayes %93, Yapay Sinir Ağları (YSA) %86, Destek Vektör Makineleri (SVM) %87, K-en Yakın Komşu (KNN) %86, J48 (C4.5) %86, Derin Öğrenme (LSTM) algoritması %95 başarı oranlarıyla sınıflandırmalarını gerçekleştirmiştir.
- 3) Yol profili belirlemede kullanılan eğitici sınıflandırma algoritmalarının bozuk ve düz yol profilleri üzerinde sırasıyla Navie Bayes %93, Yapay Sinir Ağları (YSA) %95, Destek Vektör Makineleri (SVM) %95, K-en Yakın Komşu (KNN) %83, J48 (C4.5) %96, Derin Öğrenme (LSTM) algoritması %97 başarı oranlarıyla sınıflandırmalarını gerçekleştirmiştir.
- 4) Tez çalışması için seçilen derin öğrenme algoritmalarından olan LSTM modelinin çalışma tiplerine göre aynı veri seti üzerinde farklı doğruluklar oluşturduğu ve bunun var olan veri setindeki verileri ilişkileri ile alakalı olduğu gözlenmiştir.
- 5) Sayısal veriler üzerine normalizasyon işlemlerinin uygulanması giriş parametrelerinin skala aralıkları arasındaki farklılıklardan kaynaklanacak baskın durumları ortadan kaldırdığı yapılan testler sırasında gözlenmiştir.
- 6) Farklı araç marka veya modelleri üzerinde yapılan veri toplama test sürüşleri sırasında istenilen ortak sensör verilerine ulaşıldığı ve araç marka veya modelinin sınıflandırma üzerine önemli bir etkisinin olmadığı gözlemlendi.

ÖNERİLER

- 1) Benzer olarak yapılabilecek sürüş stilleri ve yol profilleri üzerine çalışmalarda, araçlar üzerinde uygulanabilecek bir elektronik sistem ile senkron olarak sınıflandırma belirlemesi sürücülere hızlı dönüş sağlayabilir.
- 2) Derin öğrenme algoritmalarını literatürde belirlenen agresif ve sakin sürüş tiplerinin anlayacak şekilde eğitilmesi ile araç üzerine konan bir kamera ve hoparlör sistemi yardımıyla OBD2 bağlantısı aranmaksızın var olan tüm araçlarda sınıflandırma yapabileceği düşünülmektedir.



KAYNAKLAR

- [1] Karayolu Trafik Kaza İstatistikleri, <https://trafik.gov.tr/istatistikler37> Emniyet Genel Müdürlüğü, Erişim: 25 Mayıs 2023
- [2] Karayolu Güvenliğine İlişkin Küresel Durum Raporu <https://www.who.int/publications/i/item/9789241565684> Dünya Sağlık Örgütü, Erişim: 30 Aralık 2018
- [3] Chakravarty, T., Ghose, A., Bhaumik, C. and Chowdhury, C. (2013). MobiDriveScore – A System for Mobile Sensor Based Driving Analysis A Risk Assessment Model for Improving One’s Driving , *International Conference on Sensing Technology (ICST)*, Wellington, New Zealand, ss. 338 – 344
- [4] Sawant, P.R., Pande , Sachin S. (2015). A Mobile Application for Monitoring Inefficient and Unsafe Driving Behavior, *Fourth Post Graduate Conference*, cilt 7, sayı 3
- [5] Meseguer , J. E., Calafate C. T., Cano J. C. and Manzoni P. (2013). DrivingStyles: a smartphone application to assess driver behavior, *IEEE Symposium on Computers and Communications (ISCC)*, Split, Croatia, ss. 535 – 540
- [6] Bergasa, L. M., Almería, D., Almazán, J., Yebes, J. J. and Arroyo R. (2014). “DriveSafe: an App for Alerting Inattentive Drivers and Scoring Driving Behaviors”, *IEEE Intelligent Vehicles Symposium Proceedings*, Dearborn, MI, USA, ss. 240 – 245
- [7] Kumtepe , Ö., Bozdağı Akar , G. and Yüncü, E. (2015). On Vehicle Aggressive Driving Behavior Detection Using Visual Information, *Signal Processing and Communications Applications Conference (SIU)*, Malatya, Turkey, ss. 795 – 798
- [8] Kumtepe, Ö., Yüncü, E. and Bozdağı Akar, G. (2016). A Multimodal Approach for Aggressive Driving Detection, *Signal Processing and Communication Application Conference (SIU)*, Zonguldak, Turkey, ss. 729 – 732
- [9] Hong, J.-H., Margines, B. and Dey, A. K. (2014). A Smartphone-based Sensing Platform to Model Aggressive Driving Behaviors, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Ontario, Canada, ss. 4047-4056
- [10] Fazio, P., Santamaria , A. F., De Rango, F., Tropea, M. and Serianni A. (2016). New Application for Analyzing Driving Behaviour and Environment Characterization in Transportation Systems based on a Fuzzy Logic Approach, *Unmanned Systems Technology XVIII*, Baltimore, Maryland, United States, 13 sayfa
- [11] Meseguer , J. E., Toh, C. K., Calafate, C. T., Cano, J. C. And Manzoni, P. (2017). DrivingStyles: A mobile platform for driving styles and fuel consumption characterization, *Journal of Communications and Networks*, cilt 19, sayı 2, ss. 162 – 168
- [12] Johnson, D. A. and Trivedi, M. M. (2011). Driving Style Recognition Using a Smartphone as a Sensor Platform, *International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington, DC, USA, ss. 1609 – 1615
- [13] Koh, D-W. and Kang, H-B. (2015). Smartphone-Based Modeling and Detection of Aggressiveness Reactions in Senior Drivers, *IEEE Intelligent Vehicles Symposium* , Seoul, South Korea , ss. 12 – 17
- [14] Li, F., Zhang, H., Che, H. and Qiu X. (2016). Dangerous Driving Behavior Detection Using Smartphone Sensors, *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brazil, ss. 1902 – 1907
- [15] Songkroh, A., Fooprateepsiri, R. and Lilakiataskun, W. (2014). An Intelligent Risk Detection from Driving Behavior Based on BPNN and Fuzzy Logic Combination, *IEEE/ACIS 13th International Conference on Computer and Information Science (ICIS)*, Taiyuan, China, ss. 105 – 110
- [16] Wu, B.F., Chen, Y.H. and Yeh, C.H. (2012). Fuzzy Logic Based Driving Behavior Monitoring using Hidden Markov Models, *International Conference on ITS Telecommunications*, Taipei, Taiwan, ss. 447 – 451

- [17] Imkamon, T., Saensom ,P., Tangamchit, P. And Pongpaibool, P. (2008). Detection of Hazardous Driving Behavior Using Fuzzy Logic, *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, Krabi, Thailand, ss. 657 – 660
- [18] Christian G. Q. M., López, J.O. and Pinilla, A.C.C. (2012). Driver Behavior Classification Model based on an Intelligent Driving Diagnosis System, *International IEEE Conference on Intelligent Transportation Systems*, Anchorage, AK, USA, ss. 894 – 899
- [19] Li ,Y., Miyajima, C., Kitaoka, N. and Takeda, K. (2014). Measuring Aggressive Driving Behavior Using Signals from Drive Recorders, *International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Qingdao, China, ss. 1886 – 1887
- [20] Vaitkus, V., Lengvenis, P. and Žylius, G. (2014). Driving Style Classification using Long-Term Accelerometer Information, *International Conference on Methods and Models in Automation and Robotics (MMAR)*, Miedzyzdroje, Poland, ss. 641 – 644
- [21] Sun, J., Zhen, A., Li, C., Zhang, M. and Hu, X. (2016). A Vehicles' CO2 Emission Monitoring Platform Combined with Driver's Driving Behavior, *IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, Seoul, South Korea, ss. 1 – 2
- [22] Dia, H. and Panwai, S. (2015). Impact of Driving Behaviour on Emissions and Road Network Performance, *IEEE International Conference on Data Science and Data Intensive Systems*, Sydney, NSW, Australia, ss. 355 – 361
- [23] Chen, Z., Yu, J., Zhu, Y., Chen Y. and Li, M. (2015). D3: Abnormal Driving Behaviors Detection and Identification Using Smartphone Sensors, *Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Seattle, WA, USA, ss. 524 – 532
- [24] Pholprasit, T., Choochaiwattana, W. And Saiprasert, C.I. (2015). A Comparison of Driving Behaviour Prediction Algorithm Using Multi-Sensory Data on a Smartphone, *IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, Takamatsu, Japan, ss.1-6
- [25] Castignani, G., Frank, R. and Engel, T. (2013). An Evaluation Study of Driver Profiling Fuzzy Algorithms using Smartphones, *IEEE International Conference on Network Protocols (ICNP)*, Goettingen, Germany, ss. 1 – 6
- [26] Doshi, A. and Trivedi M.M. (2010). Examining the Impact of Driving Style on the Predictability and Responsiveness of the Driver: Real-world and Simulator Analysis, *IEEE Intelligent Vehicles Symposium*, San Diego, CA, USA, ss. 232 – 237
- [27] Shirazi, M.M. and Rad, A.B. (2014). Detection of Intoxicated Drivers Using Online System Identification of Steering Behavior, *IEEE Transactions on Intelligent Transportation Systems*, cilt 15, sayı 4, ss. 1738 – 1747
- [28] Gregoriades, A. and Pampaka, M. (2013). Driver Behaviour Analysis Through Simulation, *IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK, ss. 3681 – 3686.
- [29] Dong, W., Li, J., Yao, R., Li, C., Yuan, T. Ve Wang, L. (2016). Characterizing driving styles with deep learning, *arXiv:1607.03611v2*
- [30] Singh, S.K., Anand, U., Patel, A., Boro, D. (2024). Driving Behavior Analysis Using Deep Learning on GPS, *Data*. In: *Deka, J.K., Robi, P.S., Sharma, B. (eds) Emerging Technology for Sustainable Development. EGTET 2022. Lecture Notes in Electrical Engineering*, cilt 1061. Springer, Singapore.
- [31] Peppes, N.; Alexakis, T.; Adamopoulou, E.; Demestichas, K. (2021). Driving Behaviour Analysis Using Machine and Deep Learning Methods for Continuous Streams of Vehicular Data. *Sensors* 2021, no 21, 4704
- [32] Shtayat, A., Moridpour, S., Best, B., Abuhassan, M. (2023). Using supervised machine learning algorithms in pavement degradation monitoring, *International Journal of Transportation Science and Technology*, cilt 12, sayı 2, ss. 628-639
- [33] Ma, N., Fan, J., Wang W., Wu J., Jiang Y., Xie, L., Fan, R. (2022). Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms, *Transportation Safety and Environment*, cilt 4, sayı 4, ss. 1–16

- [34] Kim, G., Lee, S. Y., Oh J. -S. and Lee, S. (2021). Deep Learning-Based Estimation of the Unknown Road Profile and State Variables for the Vehicle Suspension System, *IEEE Access*, cilt 9, ss. 13878-13890
- [35] Elander, J., West, R., and French, D. (1993). Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings, *Psychological Bulletin*, cilt 113, sayı 2, ss. 279-294
- [36] Lay, M. G. and Benson, F. J. (2022). Road. *Encyclopedia Britannica*, <https://www.britannica.com/technology/road/>, Erişim: 21-05-2024
- [37] Road Profile Users' Group (RPUG), <https://rpug.org/>, Erişim: 21-05-2024
- [38] Global status report on road safety 2023, <https://www.who.int/publications/i/item/9789240086517> , Erişim: 22-05-2024
- [39] CSS Electronics Docs, <https://www.csselectronics.com/> , Erişim: 23-05-2024
- [40] Bozdal, M., Samie, M. ve Jennions, I. (2018). A Survey on CAN Bus Protocol: Attacks, Challenges, and Potential Solutions, *2018 International Conference on Computing, Electronics Communications Engineering (iCCECE)*, ss. 201-205
- [41] Çeven, S., Albayrak A. ve Bayır, R. (2022). Development of CAN and OBD-II Supported Experimental Setup for Vehicle ECU Data Acquisition, *2022 International Conference on Artificial Intelligence of Things (ICAIoT)*, ss. 1-6
- [42] Karabulut, B., Karaduman, Ö., Karabatak, M. & Eren, H. (2020). Performance Evaluation of Major Classification Algorithms for Aggressive Driving Detection using CAN-bus Data, *European Journal of Science and Technology*, sayı 20, ss: 774-782
- [43] Indrakumari, R., Poongodi, T., and Singh, K. (2021). Introduction to Deep Learning, *In EAI/Springer Innovations in Communication and Computing*.
- [44] Understanding LSTM Networks, <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>, Erişim: 06-06-2024

ÖZGEÇMİŞ

Berat KARABULUTER

[REDACTED]

[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

[REDACTED]

AKADEMİK FAALİYETLER

Makaleler:

1. Karabulut, B., Karaduman, Ö., Karabatak, M. & Eren, H. (2020). Performance Evaluation of Major Classification Algorithms for Aggressive Driving Detection using CAN-bus Data, *European Journal of Science and Technology*, sayı 20, ss: 774-782.

Bildiriler:

1. Karabulut, B., Karaduman, Ö., Avcı, E., (2017). Investigation of the Impact of Vehicle Types on the Speed Violations in the Urban / Highway Traffic according to Regions using the Naive Bayes. 8th International Advanced Technologies Symposium, (),209-209.
2. Karabulut, B., Karaduman, Ö., (2017). A Survey on Intelligent Traffic Management with Cooperative Driving in Urban Roads. The ICTM 2017: 19th International Conference on Transport Management, (),2028-2028.
3. Karabulut, B., Karaduman, Ö., (2017). A Survey on Intelligent Connected-Vehicle Applications Based on Intercommunication Techniques in Smart Cities. The ICTM 2017: 19th International Conference on Transport Management, (),2027-2027.