



**KABLOSUZ AKILLI YOL DURUM SENSÖRÜ VE GERÇEK ZAMANLI
ARAÇ SÜRÜCÜ UYARI SİSTEMİ**

Mustafa TEKE

**DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

KASIM 2019

Mustafa TEKE tarafından hazırlanan “KABLOSUZ AKILLI YOL DURUM SENSÖRÜ VE GERÇEK ZAMANLI ARAÇ SÜRÜCÜ UYARI SİSTEMİ” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi Bilgisayar Mühendisliği Ana Bilim Dalında DOKTORA TEZİ olarak kabul edilmiştir.

Danışman: Dr. Öğr. Üyesi Fecir DURAN

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.



Başkan: Prof. Dr. O. Ayhan ERDEM

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.



Üye: Dr. Öğr. Üyesi Mustafa YENİAD

Bilgisayar Mühendisliği Ana Bilim Dalı, Ankara Yıldırım Beyazıt Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.



Üye: Dr. Öğr. Üyesi İsmail ATACAK

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.



Üye: Dr. Öğr. Üyesi Hilal KAYA

Bilgisayar Mühendisliği Ana Bilim Dalı, Ankara Yıldırım Beyazıt Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.



Tez Savunma Tarihi: 22/11/2019

Jüri tarafından kabul edilen bu tezin Doktora Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

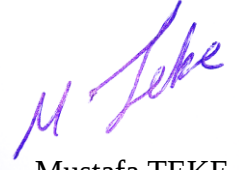
.....
Prof. Dr. Sena YAŞYERLİ
Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.



Mustafa TEKE
22/11/2019

KABLOSUZ AKILLI YOL DURUM SENSÖRÜ VE GERÇEK ZAMANLI ARAÇ SÜRÜCÜ UYARI SİSTEMİ

(Doktora Tezi)

Mustafa TEKE

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Kasım 2019

ÖZET

Bu çalışmada, kablosuz sensör ağı kullanılarak araç sürücülerine yol yüzeyinin durumu hakkında bilgi veren, yolcuların ve araçların güvenliğini sağlamaya yardımcı bir sistem tasarlanmış ve gerçekleştirilmiştir. Yol ve sürüş güvenliği en çok tehdit eden yollardaki buzlanma çalışmanın odak noktası olarak seçilmiştir. Yol yüzeyindeki buzlanmanın tespiti için toprak sıcaklığı, hava sıcaklığı, hissedilen nem, hava basıncı ve yol yüzeyindeki iletkenlik değerleri sınıflandırmanın giriş veri seti olarak kullanılmıştır. Tek kartlı bilgisayar Raspberry Pi kullanarak yol yüzeyinden toplanan veriler ile anlık olarak okunan veriler sınıflandırma algoritmaları kullanılarak yol yüzey durumu tahmin edilmiştir. En yüksek başarıma ve en hızlı cevap süresine sahip olduğu deneysel olarak belirlenen K-NN (K-Nearest Neighbor) algoritmasıyla yol yüzeyi durumu kuru, ıslak ve buzlu yol yüzeyi olarak sınıflandırılmıştır. Gerçek zamanlı ve yüksek başarımla gerçekleştirilen sınıflandırma sonucu, sensör ağı konumlarının koordinatına yaklaşan sürücülere geliştirilen mobil uygulama ile haber verilmektedir. Yapılan mobil uygulama ile yol yüzeyinin buzlu olması durumunda sürücüye sesli ve görsel olarak tehlike uyarısı yapılmakta ve bunun sonucu olarak, sürüş güvenliğini artırdığı görülmüştür.

Bilim Kodu : 92417
Anahtar Kelimeler : Android uygulama, Akıllı yol durum sensörü, Akıllı ulaşım sistemi, Gömülü sistemler, Makine öğrenmesi
Sayfa Adedi : 69
Danışman : Dr. Öğr. Üyesi Fecir DURAN

WIRELESS INTELLIGENT ROAD CONDITION SENSOR AND REAL-TIME
VEHICLE DRIVER WARNING SYSTEM

(Ph. D. Thesis)

Mustafa TEKE

GAZİ UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

November 2019

ABSTRACT

In this study, a system has been designed and implemented by using wireless sensor network to inform vehicle drivers about the condition of the road surface and to help ensure the safety of passengers and vehicles. Icing on roads that most threaten road and driving safety has been chosen as the focus of the study. For the determination of icing on the road surface, soil temperature, air temperature, relative humidity, air pressure and conductivity values on the road surface have been used as the input data set of the classification. Using the single-board computer Raspberry Pi, the road surface condition was estimated by using the classification algorithms collected from the road surface and the data read instantly. The road surface condition has been classified as dry, wet and icy road surface by K-NN (K-Nearest Neighbor) algorithm, which was determined experimentally to have the highest success and fastest response time. As a result of real-time and high-performance classification, drivers approaching the coordinate of sensor network locations are notified via mobile application developed. With the mobile application, if the road surface is icy, the driver is warned visually and audibly, and as a result, it is observed that the driving safety is increased.

Science Code : 92417

Key Words : Android application, Intelligent road condition sensor, Intelligent transportation system, Embedded systems, Machine learning

Page Number : 69

Supervisor : Assist. Prof. Dr. Fecir DURAN

TEŞEKKÜR

Çalışma süresince tez danışmanlığımı üstlenerek bana yol gösteren, çalışmanın planlanması, yürütülmesi, raporlanması ve sonuçlanmasında katkı ve destek sunan değerli sayın hocam; Dr. Öğr. Üyesi Fecir DURAN'a, bu süreçte akademik katkı ve desteklerini benden esirgemeyen tez izleme komitesi üyesi kıymetli hocalarım Prof. Dr. O. Ayhan ERDEM ve Dr. Öğr. Üyesi Mustafa YENİAD'a en içten teşekkür ve saygılarımı sunarım.

Yaşamımın her döneminde ve her türlü kararında yanımda olan, maddi ve manevi desteklerini her zaman yanımda hissettiğim babam Suat TEKE ve annem Sema TEKE'ye, teşekkür ederim. Tez çalışma sürecimde sabırla ve sevgiyle beni destekleyen çocuklarım Sema Zeynep ve Ahmet Suat'a, hayatıma girdiği günden bu yana sevgisini, özverisini ve desteğini bana her zaman sunan eşim Türkan TEKE'ye en içten dileklerimle teşekkür ederim.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
RESİMLERİN LİSTESİ	xi
SİMGELER VE KISALTMALAR.....	xii
1. GİRİŞ.....	1
2. LİTERATÜR TARAMASI	3
3. MATERYAL VE METOT	9
3.1. Materyal	9
3.1.1. Tek kartlı bilgisayar	10
3.1.2. Hava basınç sensörü.....	12
3.1.3. Nem ve sıcaklık sensörü.....	13
3.1.4. Yol sıcaklık sensörü	13
3.1.5. Yol yüzey iletkenlik sensörü.....	14
3.1.6. Android Studio	15
3.1.7. Python	18
3.2. Metot	18
3.2.1. Makine öğrenmesi.....	19
3.2.2. Problemin tanımlanması	20
3.2.3. Veriyi anlama	21

	Sayfa
3.2.4. Veriyi hazırlama.....	21
3.2.5. Modelleme.....	24
3.2.6. Model değerlendirme ve seçimi	24
3.2.7. Modelin uygulamaya geçirilmesi	28
3.2.8. Sınıflandırma algoritmaları	28
3.2.9. Özellik seçimi.....	38
3.2.10. Özellik seçim yöntemleri	40
4. KABLOSUZ AKILLI YOL DURUM SENSÖRÜ VE GERÇEK ZAMANLI ARAÇ SÜRÜCÜ UYARI SİSTEMİ TASARIMI VE GERÇEKLEŞTİRİLMESİ.....	43
4.1. Veri Toplama Sistemi	45
4.2. Özellik Seçimi.....	47
4.3. Sınıflandırma Algoritması Seçimi.....	48
4.4. Akıllı Yol Durum Tespit Gömülü Sisteminin Gerçekleştirilmesi.....	53
4.5. Mobil Uygulama Gerçekleştirilmesi	57
5. SONUÇ VE ÖNERİLER	59
KAYNAKLAR	61
ÖZGEÇMİŞ	69

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.1. Karışıklık matrisi	24
Çizelge 4.1. K-NN algoritması başarımları	49
Çizelge 4.2. Bayes algoritmasının başarımları.....	49
Çizelge 4.3. Karar ağacı algoritmasının başarımları.....	50
Çizelge 4.4. Rastgele orman algoritmasının başarımları	50
Çizelge 4.5. Lojistik regresyon algoritmasının başarımları	51
Çizelge 4.6. Destek vektör makinası algoritmasının başarımları.....	51
Çizelge 4.7. Sınıflandırma algoritmalarının başarımları.....	52
Çizelge 4.8. Sınıflandırma algoritmalarının cevap süresi	53

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 3.1. Tasarlanan sistemin genel şeması.....	10
Şekil 3.2. Yol yüzey iletkenlik sensör şeması.....	15
Şekil 3.3. Android işletim sistemi mimarisi.....	16
Şekil 3.4. Veri madenciliği süreci.....	19
Şekil 3.5. CRISP modeli akış şeması.....	20
Şekil 3.6. K-NN sınıflandırma algoritması.....	30
Şekil 3.7. Karar ağacı algoritması örneği.....	31
Şekil 3.8. Rastgele orman akış diyagramı.....	33
Şekil 3.9. Destek vektör makinesi algoritması.....	37
Şekil 3.10. Özellik seçimi algoritması	39
Şekil 4.1. Akıllı yol durum sensörü ve gerçek zamanlı araç sürücü uyarı sistemi	43
Şekil 4.2. Akıllı yol durum sensörü akış şeması.....	44
Şekil 4.3. Akıllı yol durum uyarı sisteminin akış şeması	45
Şekil 4.4. Özellik sayısı ile performans değeri arasındaki ilişki.....	48

RESİMLERİN LİSTESİ

Resim	Sayfa
Resim 3.1. Tek kartlı bilgisayar	11
Resim 3.2. Hava basınç sensörü	12
Resim 3.3. Nem ve sıcaklık sensörü.	13
Resim 3.4. Yol sıcaklık sensörü.....	14
Resim 4.1. Aktif ARS31, Pasif IRS31 ve WS600 sensörleri.....	46
Resim 4.2. ARS31, IRS31 ve WS600 sensörlerinden toplanan veriler	46
Resim 4.3. Gerçekleştirilen sistemin bütünleşmiş hali	54
Resim 4.4. Gerçekleştirilen programın Python çıktısı	55
Resim 4.5. Web sunucusuna gönderilen yol durum bilgisi	56
Resim 4.6. Web sayfasında listelenen yol durum bilgisi	56
Resim 4.7. Veri tabanı sayfası ve yol durum tahmini.....	57
Resim 4.8. Android cep telefonu uygulaması ekran görüntüsü	58

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Açıklamalar

x'	Normalize edilen değer
x_i	Veri yığınınındaki girdi değeri
x_{max}	Veri yığınınındaki en büyük sayı değeri
x_{min}	Veri yığınınındaki en küçük sayı değeri
μ_i	Veri yığınınının ortalaması
σ_i	Veri yığınınının standart sapması
e	Doğal logaritma değeri

Kısaltmalar

Açıklamalar

DN	Doğru Negatif
DP	Doğru Pozitif
DVM	Destek Vektör Makinesi
KA	Karar Ağacı
K-NN	K- En Yakın Komşu
LR	Lojistik Regresyon
NB	Naive Bayes
RO	Rastgele Orman
SSH	Secure Shell
YN	Yanlış Negatif
YP	Yanlış Pozitif

1. GİRİŞ

Günümüzde ulaşım için, hava yolları, deniz taşımacılığı, demir yolları ve kara yolları kullanılmaktadır. Hava yolları her ne kadar ulaşım hızla yaygınlaşsa da karayolları halen en çok tercih edilen ulaşım türüdür. Karayolu ulaşımında iki tekerlekli araçlar, arabalar ve toplu taşıma araçları kullanılmaktadır. Güvenlik açısından istatistiksel olarak en tehlikeli olan karayollarının tercih nedeni, kısa seyahat mesafelerindeki alternatifsizlik ve seyahat maliyetlerinden kaynaklanmaktadır.

Karayollarında meydana gelen kazaların en önemli etkeni sürücü hatalarıdır. Bu nedenle ulaşım esnasında güvenlik tedbirinin alınmış olması ya da tehlikeli yol ve hava koşuluna karşı sürücülerin uyarılması seyahatin konfor ve güvenliği için önemlidir. Karayolu etrafına yerleştirilen uyarı ve işaret levhaları bazen sürücüler tarafından fark edilememektedir. Sürücülerin dikkatini çeken ışıklı ve sesli uyarı işaretçileri günümüzde kullanılmaktadır. Akıllı cep telefonları ile sürücünün bulunduğu konumdaki yolun risklerinden haberdar edilmesini de günümüzde yaygınlaşmaktadır.

Kış mevsiminde kar ve buzlanmadan dolayı kazaların arttığı bilinmektedir [1]. Oluşan trafik kazaları sonucunda sürücüler ve yolcular psikolojik ve ekonomik olarak olumsuz etkilenmektedir [2]. Olumsuzluklara sebep vermemek için sürücü ve yolcuların her türlü trafik işaret ve işaretçileri aracılığıyla uyarılması gerekmektedir. Çünkü sürücünün varmak istediği hedefe rota belirlerken yol yüzey koşullarını bilmesi hangi yoldan gideceği tercihinin ve yolculuk esnasındaki sürüş biçimini belirlemede yardımcı olacaktır.

Yol bakım birimleri kış aylarında yollardaki buzlanmayı önlemek için tuzlama çalışmaları yapmaktadırlar. Maliyet açısından bakıldığında sadece yol bakım maliyeti değil trafik kazalarından meydana gelen ekonomik zararlarında azaltılmasını hedeflemektedir. Yolun tuzlanması ile trafik kazası ve tıkanıklıkları büyük ölçüde azaltılabilir. Doğru zamanlama tahmini yapılan tuzlama çalışmasının hem kazaları önleyecek hem de tuzun yola verdiği zararı azaltacaktır. Aşırı ve gereksiz tuzlama yol yüzeyinden yol kenarlarına kayması sonucu çevreye de zarar vermektedir. Aşırı tuzlama çevredeki ağaçların kurumasına, suyun kirlenmesine, yol yüzeyinin bozulmasına sebep olur. Aynı zamanda ekonomik anlamda da yol bakım maliyetini artırmaktadır. Bu nedenle en uygun tuzlama miktarını ve süresini

hesaplamak gerekir. Bunu saęlamak içinde yol yüzey koşulunun yüksek doğrulukta tahmin edilmesi gerekmektedir.

Bu tez çalışmasında kış koşullarında sürücülerin güvenli seyahat edebilmesi literatürdeki eksiklikleri gidermeye yönelik; hızlı cevap süresine sahip, düşük maliyetli, güvenilir, yüksek tahmin başarımlarına sahip, etkin bir uyarı sistemi içeren yol durum tespit ve uyarı sistemi gerçekleştirilmektedir. Uyarı sisteminde yol durumunu algılamak için raspberry pi tabanlı temashlı akıllı yol durum sensörü geliştirilmiştir. Bu sensör ile elde edilen veriler geliştirilen mobil uygulama üzerinden gerçek zamanlı olarak sürücülere uyarı olarak iletilmektedir.

Bu tez 5 bölümden oluşturulmuştur. Tezin ikinci bölümünde, bu alanda yapılan önceki çalışmalara yer verilmiştir. Üçüncü bölümde, bu çalışmada geliştirilen uyarı sisteminin bileşenleri olan akıllı yol durum sensörü ve mobil uygulamanın içerięi, materyal ve metot başlıęı altında verilmektedir. Dördüncü bölümde, geliştirilen sistem, sistemde kullanılan özellikler ve sınıflandırıcının belirlenmesi için yapılan çalışmalara yer verilmektedir. Sonuç ve değerlendirme bölümünde ise sistemin genel değerlendirmesi ve sonraki çalışmalara ait öneriler sunulmuştur.

2. LİTERATÜR TARAMASI

Literatürde sürücüler için yol yüzey durumu ve kötü hava koşullarının bildirilmesi için çeşitli çalışmalar yapılmıştır. 1980 yıllardan beri yol ve hava durumu bilgi sistemi aktif olarak sürücü ve yolculara hizmet vermektedir. Yol hava durumu bilgi sistemi Road Weather Information Systems (RWIS) araç sürücüleri için karayolu çevresine yerleştirilmiş uyarı sistemi olarak çalışmanın yanında yol güvenlik ve bakımından sorumlu otoyol birimlerine de yol yüzeyindeki buzlanma ve kar hakkında bilgilendirme yapmaktadır. RWIS çalışma amacı yol yüzeyindeki olumsuz durumu giderirken harcanan maliyeti düşürmek ve israfı önlemekle birlikte öncelikli olarak araç sürücülerinin ve yolcularının güvenli bir şekilde ulaşım yapmasıdır. RWIS çalışmasıyla yol yüzeyindeki olumsuz durumları giderirken zaman ve maliyetten tasarruf ettirmektedir. Karayoluna yerleştirilen bilgi istasyonlarında yol ve hava değerlerinin ölçümü ve yerleştirilen kameralar yardımıyla buz ve kar tespiti yapılır [3].

Karayolu üzerinde bulunan buzlanma ve karlanma tespiti için temelde temaslı ve temassız olmak üzere iki ana kategoride yaklaşımlar bulunmaktadır. Temaslı ve temassız tespit sisteminde farklı fiziksel prensiplere dayalı ısı, ışık ve ses algılayıcılarının yanında kamera gibi aygıtlar da kullanılmaktadır. Kullanılan yol durum tespit sensörleri farklı fiziksel prensiplere dayalı kızılötesi, lazer görüntüleme, fiber optik, rezonans frekansı, ultrason, iletkenlik, kapasite etkileri ve elektromanyetik dalgalardan yararlanan algılayıcılardır [4, 5].

Yol yüzey tespitinde kullanılan kamera, algılayıcı ve diğer cihazların karayolu çevresine yerleştirilerek temaslı yaklaşım grubu oluşurken, bu aygıtların araç üzerine yerleştirilmesiyle temassız yüzey tespit yaklaşımı oluşmaktadır. Temaslı ve temassız yaklaşımların doğru sonuç verebilmesi için yol yüzeyini aydınlatacak ışık kaynağına ve işlem kapasitesi yüksek bilgisayarlara ihtiyaç vardır. Temaslı tespit sistemlerinde bu kaynaklara ulaşmak kolay iken taşınır yani temassız sistemlerde daha zordur [6].

Temassız tespit sistemlerinde görüntü işleme, ultrasonik radar ve infrared detektörler üzerine çalışılmıştır [7]. Yol yüzey koşulunu optik polarizasyon, görsel görüntü işleme ve hava durumu istasyonlarından alınan veriler ile kombinasyonuna dayanarak temassız yol durum sensörü önerilmiştir [8]. Temassız yol durum sensörleri otomobillere veya karayolu

kenarlarında bulunan direklere monte edilmek üzere tasarlanmıştır. Ancak temassız yol durum sensörlerinde çözülmesi gereken, buz veya su kalınlığının ölçülmesinde problemler bulunmaktadır. Casselgren'in sistemi ve Jonsson'ın metodu ince buz tabakası ya da gizli buzlanmayı fark edememektedir. Bunun yanı sıra yaygın kullanım için yüksek maliyetli bir sistem olmuştur.

Yol yüzeyinin buzlu veya karlı olduğunun tespitinde dalga boyuna hassas kameralar kullanarak yapılan çalışmalarda %80-90 doğruluk oranına ulaşılırken, ıslak ve kuru yol durumunun tespitinde %70 doğruluk oranına ulaşılmıştır [9]. Temassız yol yüzey tespitinde görüntü işlemenin yanında görüntünün işlendiği noktanın GPS verileri ile işaretlenerek yol durum haritası oluşturulmaktadır. Yol yüzeyi karla kaplı olduğu durumda kar altında kalan buzun tespiti bu sistemde mümkün değildir. Dahası kameranın ani ışık şiddetindeki değişimler ile gölgelerden kaynaklı yüzey tespitinde hataya sebep olmaktadır [10].

Danimarka da yapılan bir çalışmada yol yüzey koşullarını en çok etkileyen yol yüzey sıcaklığını tahmin eden sistem üzerine çalışılmıştır. Yol yüzeyi hakkında yapılan bir diğer çalışmada atmosferik değerlerin merkezi istasyondan alınıp sayısal tahmin üzerine bir sistem geliştirilmiştir. Bu çalışmanın zayıf yanı ise merkezi veriler ile yapılan tahminler ile riskli bölgelerdeki yolun atmosferik değerlerinin aynı olmamasından kaynaklı tahminlerin birbiriyle örtüşmemesidir. Sonuç olarak atmosferik veriler yardımıyla genel bir yol durum tahmini yapılmıştır ancak güneş ışığı görmeyen gölgede kalan ya da ani hava değişimlerine maruz olan bölgelerdeki yol yüzey tahmininde başarı durumu düşüktür [11].

Gagnon ve arkadaşları ışığın yansıma ve emilme kurallarını temel alan optik sistem üzerine kurulu yol yüzey tespit yöntemini geliştirmiştir. Yol yüzeyine gönderilen ışık huzmeleri kuru yol yüzeyinde dağınık şekilde yansımaya uğrarken, ıslak ve buzlu yüzeylerde oluşan tabakadan ötürü aynadan yansıyan ışık huzmeleri daha doğrusaldır. Ancak gizli buzlanma gibi ince film tabakalarında bu yöntem doğru çalışmamaktadır [12]. Casselgren ve arkadaşları yol durumunu NIR (Near InfraRed) kamera ve güçlü ışık kaynağı yardımıyla yol yüzey hakkında tahminde bulunmuştur ancak kurulan sistem ekonomik değildir [13].

Alonso ve arkadaşları asfalt yol yüzeyini sınıflandırmak için mikrofon aracılığıyla araç lastiğinden gelen gürültünün analizi ile ıslak ya da kuru şeklinde sınıflandırma yapılmıştır. Gerçekleştirilen sistem yol yüzeyindeki ıslak ve kuru durumu yüksek oranda başarılı olarak

sınıflandırmıştır [14]. J. Shao ve P. J. Lister çalışmalarında ıslak ve buzlu yol yüzeyi tahminini meteorolojik değerlerden faydalanarak yapmış ve %82 oranında buzlu yol yüzeyinde, %72 oranında da ıslak yol yüzeyinde tahmin başarımı elde etmiştir. Yol yüzeyinden alınan herhangi bir ölçüm olmadığı için yolun tuzlanması tahmin başarısını düşürmektedir. Yapılan çalışmalarda yol yüzeyinden bilgi alınmaması durumunda yol yüzeyi sınıflandırmasında başarı oranı düşmektedir ve karlı yüzeyin altında bulunan buzun tespiti zorlaştığı gibi tuzlanma sonucunda donma noktasının düşmesi de yüzey tahmininde başarıyı düşürmektedir.

Temaslı yol yüzey tahmini yapan sistemlerin sınıflandırma başarımı daha yüksektir [15]. Temaslı yol yüzey durum sensörleri iletkenlik ve kapasitans ölçüm temellerine dayanarak çalışmaktadır. Tabatabai ve Aljuboori yol yüzeyinde bulunan buz ve su tabakasını tespit etmek için yeni bir iletkenlik sensörü kullanmıştır. Gerçekleştirilen sensör ıslak, kuru ve buz sınıflandırmasında etkili bir kompakt yapıya sahiptir [16]. Troiano ve arkadaşları entegre bir kapasite sensörü tasarlayıp üç temel yol durumunun tespiti için çalışmışlardır. Sensör üzerinde bulunan su ya da buz katmanının kapasite ve iletkenliği ölçülerek sınıflandırma yapılmıştır. Ancak sensör üzerine serpilene tuzun yapılan ölçümlere bozucu etkisi olmuştur [17]. Havacılık alanında buz tespiti ve kalınlık ölçümü için rezonans ve optik sensörler başarıyla uygulanmış ancak algılayıcının çok fazla uzun olması nedeniyle karayolunda uygulamaya imkan vermemiştir [18].

Roy ve arkadaşları piezoelektrik dönüştürücü kullanarak yol yüzeyinin kuru, ıslak ve buz olduğunu tespit etmiştir. Yol yüzeyine yerleştirilen plakanın bir tarafında titreşim yapmak için eyleyici olarak piezoelektrik dönüştürücü, diğer tarafına ise ters piezoelektrik dönüştürücü monte edilmiştir. Plakanın titreşim genliği ölçülmesiyle sonuçta buzlu yüzeyin plakanın esnekliğini azaltması buna bağlı olarak rezonans frekansını arttırması, sulu ve kuru karlı yüzeyinde rezonans frekansını azalttığı gözlenmiştir. Buz kaplı yüzey sensörün bükülmesini ve kütlesini arttırırken, su kaplı yüzeyde sensörün sadece kütlesi artmaktadır [19]. Li ve arkadaşları piezoelektrik dönüştürücü ile yapılan ölçüm aralıklarını genişletmek için çalışmıştır ancak ölçüm aralığını arttırırken sensör araçlar tarafından zarar görebilecek hale gelmiştir [20]. Gui ve arkadaşları Roy ve Li'nin çalışmalarındaki problemleri çözmek için makine öğrenmesi ve piezoelektrik dönüştürücüyü sürmek için farklı frekans taramalarından faydalanmıştır. Uygulanan frekans değerlerine karşı su ve buz tabakalarının verdiği titreşim tepkilerinden elde edilen eğriler yardımıyla bir ilgileşim elde edilmiştir.

Geleneksel yöntemlere nazaran regresyon ve yapay sinir ağına dayalı oluşturulan modelden buz kalınlığı tespitinde yüksek başarımlar elde edilmiştir [21].

Jonsson ve arkadaşları kızıl ötesi kamera kullanarak asfalt yüzeyi tahmini üzerine çalışmıştır. KNN (K Near Neighbours) sınıflandırıcısının SVM (Support Vector Machine) sınıflandırıcısından daha yüksek sınıflandırma başarımları olduğunu belirtmiştir. Sistem için yararlanılan araçların ekonomik ve fiziksel boyutu çok yüksek olması nedeniyle kullanılabilirliği zordur [22]. Li yapmış olduğu sınıflandırmada, kullandığı öznelilikleri SVM algoritmasından yararlanarak özellik seçim işlemi sonucu elde ettiği alt özellik kümesi ile daha iyi sonuç aldığını ifade etmiştir [23].

Casselgren ve arkadaşları asfalt yüzeyinde oluşan durumları optik tabanlı sensörler yardımıyla %80-90 tahmin başarımları elde edilirken, kamera tabanlı optik yansımali sistemlerde %70-80 değerleri arasında başarımlar elde edilmiştir [6]. Cheng ve arkadaşları yapmış olduğu çalışmada geleneksel görüntü işleme tekniklerinin gerçek zamanlı yol yüzey durumu tespitinde kullanmak için yeterli düzeyde hızlı ve yüksek performansa sahip olmadığını belirtmişlerdir. Bu eksikliklerin ve zayıflıkların derin öğrenme aracılığıyla giderilebileceğini, modellere performans iyileştirmesi ve etkin işlevsellik katılabileceğini ifade etmişlerdir. Evrişimli sinir ağı modelinin yüksek sınıflandırma performansına sahip olduğunu ve daha iyi genelleme yaparak %94,89 oranında sınıflandırma başarımları elde edilmiştir. Ancak yol yüzeyinin tuzlanması ve kar altında bulunan buzlu yüzey dikkate alınmadan sınıflandırma başarımları elde edilmiştir [24].

Temaslı yol yüzey durumu tespitinde çeşitli sensörler asfalt yüzeye ya da kaldırıma yerleştirilerek tahminde bulunulur [25]. Metro (Model of the Environment and temperature of Roads) 1999 yılında uygulamaya konularak yol durumu için meteorolojik değerler ile ve yol yüzey sıcaklık değerleri kullanılarak bilgilendirme yapmıştır. Uygulamaya konulan Metro aracılığıyla yıllık bazda kış ayları buzlanma değerleri veri tabanı oluştururken sonraki yıllara önceden yapılacak tuzlanma ve yol bakım maliyetleri gibi hazırlıkların yapılmasını sağlamaktadır. Küresel ısınma sebebiyle yıllık don ve buzlanma süresinin değişimi ve tek tip veri kullanması sebebiyle istatistiksel yaklaşımda hata oluşsa da kış öncesi yaklaşık bir hazırlık için fikir oluşturmaktadır.

Shao çalışmasında yakın zaman asfalt yüzey sıcaklık tahminini yapay sinir ağının üç katmanlı modelini kullanarak gerçekleştirmiştir. Asfalt yakınına yerleştirilen istasyonlardan alınan veriler ile hava durumu verileri sinir ağına (NN, Neurol Network) eğitim verisi olarak girilmiş ve buzlanma tahmini yapılmıştır [26].

Luchetta ve arkadaşları, yol yüzey durumunun tahmininde kurulan sensör ağına NN başarımının, analitik ve uzman sistemlerden daha yüksek oranda doğru sonuç verdiğini ifade etmiştir [27]. Bogren asfalt yüzeyinin sıcaklığını değiştiren güneşlenme, rüzgar hızı ve rakım değerleri arasındaki ilişkiye bakmıştır. Asfalt zemin üzerindeki sıcaklıkta güneşlenme süresinin verdiği etkinin, rüzgar hızının verdiği etkiden daha fazla olduğunu belirtmiştir [28].

Yapılan diğer çalışmalarda yol durumu hakkında verilen ikaz işaretlerinin sürücüler tarafından tam olarak fark edilmediğini söylemektedir [29]. Yol yüzey tespitinin ardından sürücüye bu bilginin dikkat çekecek şekilde iletilmesi gerekmektedir. Oluşan trafik kazaları ile trafik levhalarının doğru konumlandırılması arasındaki ilişki ile ilgili çalışmalar yapılmış ve araç sürücülerinin yol durumu hakkında uyarılması yetersiz görülmüştür [30]. Araç sürücülerinin tehlikeli bölgeye yaklaşmadan önce dikkat çeken işaret ve işaretçiler tarafından ikazı yapılmalıdır.

Literatürde yapılan temassız yol yüzey durumu tanıma çalışmalarında ses, ışık ve görüntü temelli algılayıcılar ile sınıflandırma yapılmıştır. Ancak kullanılan bu algılayıcıların her birinde farklı olumsuzlukların bulunması nedeniyle genel olarak eksiklikler şu şekilde ortaya çıkmaktadır:

- Genel olarak tahmin başarımları temaslı algılayıcılara göre daha düşüktür.
- Ultrasonik algılayıcıları ile yapılan çalışmalarda, kuru ve ıslak zemin tespitinde yüksek başarı elde edilmişken buzlu, karlı ve tuz serpilmiş yol hakkında bir çalışma yapılmamıştır.
- Optik ışık temelli çalışan algılayıcılar ile yapılan çalışmalarda ışığın farklı dalga boylarında yansımaları ya da absorbe edilmesine yönelik çalışmalar gerçekleştirilmiştir. Ancak temiz yüzeye sahip veya hava kabarcığı bulunmayan buzun tespit edilmesi ile sınırlı kalmıştır.

- Görüntü işleme ile yapılan yol yüzey sınıflandırılmasında güçlü aydınlatma kaynağına ihtiyaç duyulması, ani değişen ışık geçişleri, yol yüzey görüntü kalitesi, sis gibi olumsuz hava koşulları, araç hareketliliğinden kaynaklı bulanıklıktan ve elde edilen görüntülerin işlenmesi için güçlü makinalara ihtiyaç olması gibi zayıflıkları bulunmaktadır.
- Tuzlu yol yüzeyi ve karlı yüzey altındaki buzlu zemin durumu göz ardı edilerek yapılan görüntü tabanlı çalışmalarda elde edilen en yüksek başarı yüzdesi 94,89% olarak ifade edilmiştir.
- Yol altı sıcaklığının ve tuzlama olayının temassız ölçümlerde sınıflandırma sürecine dâhil edilmemesi yapılan çalışmalarda yol durumunun tahmin başarımını düşürmektedir.
- Kullanılan optik, ses veya görüntü işleme sistemlerinin gerçek zamanlı sonuç verebilmesi için güçlü makinalara, maliyeti yüksek donanımlara ve hacimsel olarak büyük alanlara ihtiyaç duymaktadır.

Temaslı yol yüzey durumu tanıma ve bilgilendirme çalışmalarındaki zayıflıklar ise şu şekildedir:

- Yol yüzeyine ve altına yerleştirilmesi gerektiği için montaj işlemi zordur.
- Yolun tuzlu olması durumunda yol durumunun tespiti çok daha uzun sürede mümkün olmaktadır.
- Bilgilendirme sistemleri ile sadece lokal ve kablolu olarak haberleşebilmektedir.
- Yüksek maliyete sahiptirler.
- Yol durum bilgisi meteorolojik veriler aracılığıyla sürücülere genel olarak bilgilendirme yapılırken, riskli konumlarda gömülü sistem üzerinde çalıştırılan gerçek zamanlı bir yol durum uyarı ve bilgilendirme sistemi bulunmamaktadır.

3. MATERYAL VE METOT

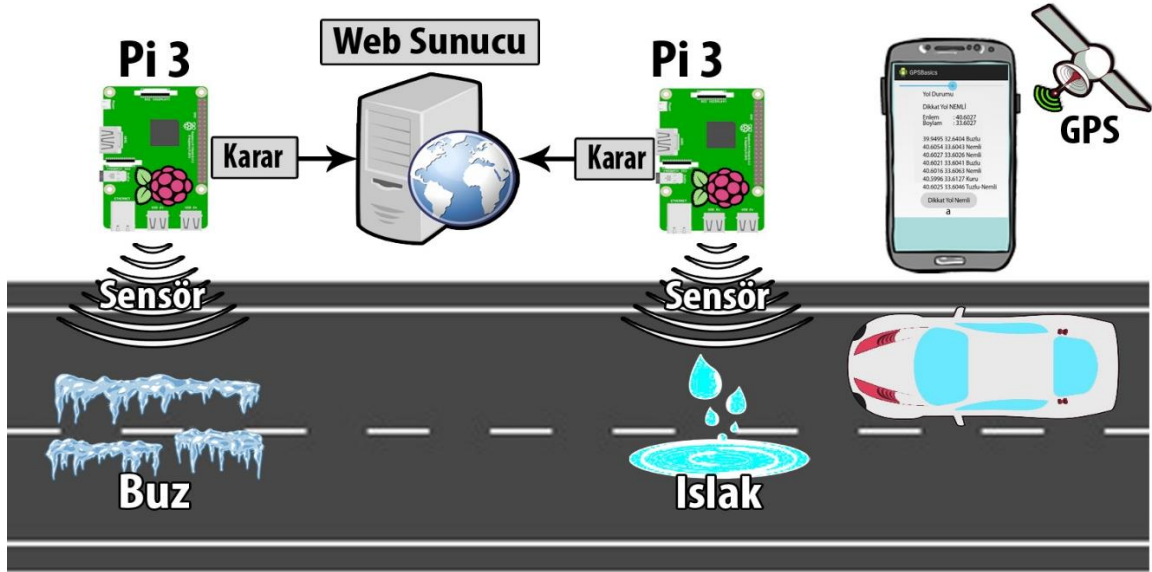
Yol durum sensörü ve uyarı sistemi tasarımında birçok farklı metot ve yöntem kullanılabilmektedir. Temaslı yol durum sensörlerinin temassızlara göre başarımının yüksek olması nedeniyle bu çalışmada temaslı akıllı yol durum sensörüne uygun materyaller ve metotlar tercih edilmiştir. Yol durum sensörlerinden elde edilen verilerle sürücülerin gerçek zamanlı ve etkin olarak uyarılması için mobil uygulama kullanılmaktadır.

3.1. Materyal

Akıllı ulaşım sistemleri sürücülerin ve yolcuların güvenlik ve konforunu artırmak için çalışmaktadır. Emniyetli bir sürüş için yol yüzey durumunu araç sürücülerinin bilmesi gerekir. Yol yüzey tahmininde bulunmak için yapılan çalışmalarda çevresel verilerin toplanması tahmin başarımında önemli bir etkidir. Genel olarak yol durum tahmini için meteorolojik veriler kullanılmaktadır ancak riskli bölgelerde bu veriler tam manasıyla yol durumunu temsil etmemektedir. Bu sebeple riskli bölgelerde yol yüzey tahmini yapılırken bölgenin içinde bulunduğu hava ve yol verilerinin kullanılması gerekmektedir.

Akıllı yol durum tespit ve uyarı sistemi temelde iki kısımdan oluşmaktadır. İlk olarak yol yüzey durumunun tespiti için, karayolu üzerine yerleştirilen iletkenlik ve yol sıcaklık sensöründen alınan değerler ile bulunduğu riskli konuma ait hava, nem ve basınç değerlerinin anlık olarak gerçek zamanlı toplanıp Raspberry Pi gömülü sistemin üzerinde sınıflandırma algoritmasına sokularak yol yüzey tahmini yapılmıştır.

Yol yüzey tahmininin yapılmasının ardından Raspberry Pi gömülü sistemi internet bağlantısı aracılığıyla bulunduğu konuma dair oluşan yol yüzey bilgisini veri tabanına ve oradan da web sayfasına göndermektedir. Android tabanlı akıllı cep telefonları için yapılan uygulama ile araç sürücüsü yol durum sensörünün bulunduğu konuma yaklaştığında telefona yazılan uygulama ile uyarılmaktadır. Tasarlanan sistemin genel şeması Şekil 3.1’de gösterilmektedir.



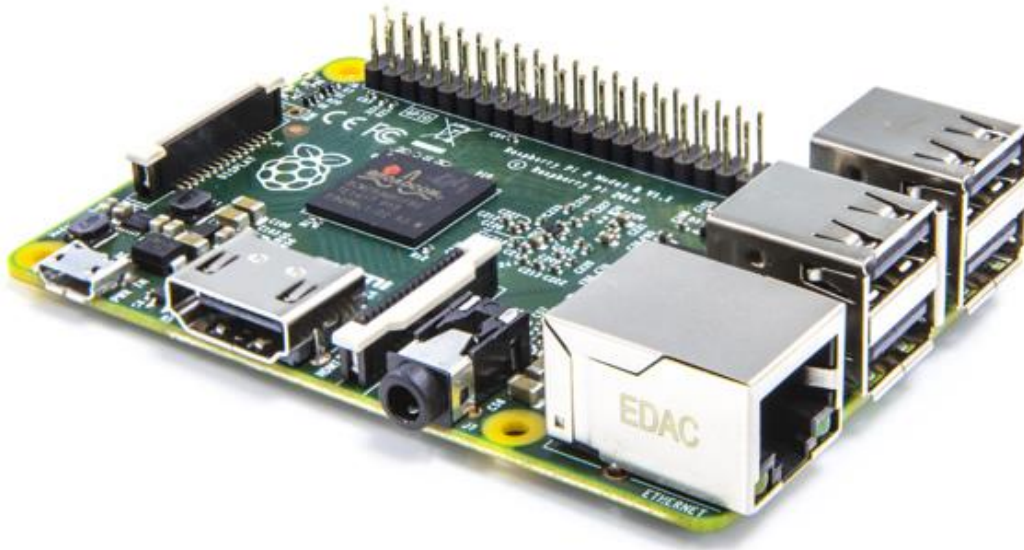
Şekil 3.1. Tasarlanan sistemin genel şeması

Yol altına yerleştirilen sensör yol sıcaklığını ölçmekte ve hava durumu verilerine bakıldığında buz oluşumu mümkün iken yol sıcaklığı nedeniyle yol yüzeyinde buz oluşmamaktadır. Temassız ölçüm sistemlerinde yol sıcaklığı ve tuzluluk verileri göz ardı edildiği için yol yüzey tahmin başarımı düşük çıkmaktadır. Akıllı yol durum tespit ve uyarı sistemi için geliştirilen yol yüzey durum sensörü yüzey iletkenlik sensörü, yol ve hava sıcaklık sensörü, hava basıncı ve bağıl nem sensörü ile Raspberry Pi3'den meydana gelmektedir [31].

3.1.1. Tek kartlı bilgisayar

Tek kartlı bir bilgisayar olan Raspberry Pi geleneksel gömülü sistemlerden daha çok hesaplama kabiliyetine sahip olmanın yanında düşük enerji tüketimine sahip bilgisayar olarak kullanılmaktadır. Fiziksel boyutunu ve maliyetini en aza indirmek için tek bir kart üzerine bilgisayarı oluşturan bütün modüller entegre edilmiştir. Broadcom BCM43438 yonga ile 802.11n kablosuz ağ ve düşük enerji harcayan Bluetooth 4.1 desteğini sağlamaktadır. Kart üzerinde bulunan dahili anten, kablosuz ve bluetooth ağlara erişmesini sağlamaktadır.

Raspberry Pi için özel üretilen Broadcom BCM2837 yonga üzerinde 32 kB L1 ön bellek ve 512 kB L2 ön belleği olan 1.2Ghz hızında ARM Cortex-A53 işlemci çekirdeği ile 1GB boyutunda DDR2 RAM bellek bulunmaktadır. Genel amaçlı giriş, çıkış ve sensör bağlantı işlemleri yapabilmek için kullanılan 40 adet pin bulunmaktadır. Benzer şekilde SMSC LAN9514 yongası üzerinde 10/100 ethernet bağlantısı yanında dört adet USB 2.0 bulundurmaktadır. Bunun yanında seri ekran ara yüz bağlantısı, seri kamera ara yüz bağlantısı ve işletim sisteminin yüklendiği Mikro SD (Secure Digital) kart desteği bulunmaktadır. Raspberry Pi üzerine kurulabilen Linux tabanlı Raspbian Jessie gibi açık kaynak işletim sistemleri bulunmaktadır. Bu işletim sistemleri Python ve Scratch gibi birçok programlama dilini desteklemektedir. Bu dillerde yazılan programlar sistemin açılışında belirtilerek istenildiği zaman otomatik olarak çalıştırılabilmektedir [32]. Resim 3.1’de tek kart bilgisayar olan Raspberry Pi 3 gösterilmiştir.



Resim 3.1. Tek kartlı bilgisayar

Birçok programlama dilini destekleyen yapısı fonksiyonel özellikleri ve elektronik devre tasarımlarına uygun tek kartlı bilgisayar olarak tercih edilmiştir. Yol durum sensör düğümü tasarımında kablosuz bağlantıya sahip olması sınıflandırma algoritmalarının üzerinde koşturulması en önemli tercih nedenlerindendir. Sınıflandırma algoritmalarındaki büyük veri yığını çalıştırıyor olması ve başka bir sunucuya ihtiyaç duymuyor olması makine öğrenmesi projelerinde kullanılmasında etkilidir. Yol durum sensör tasarımında hava sıcaklığı, nemi, basıncı ve tuzluluk oranının algılanması gereken sensör bağlantılarının kolaylığı, maliyeti ve diğer sebeplerden dolayı bu çalışmada Raspberry Pi tercih edilmiştir.

3.1.2. Hava basınç sensörü

Hava basıncı değeri artmasıyla donma noktası düşerken kaynama noktası da yükselmektedir bu nedenle yol yüzeyinin buzlanmasında hava basıncının da etkisi görülmektedir. Toplanan veri kümesinde özellik seçimi işlemi yapıldığında seçilen özellikler içerisinde hava basıncı değeri yüksek oran almış ve sınıflandırma işlemi için öznitelik olarak seçilmiştir. Yol durum tahmininde kullanılan hava basıncı değeri, tüketici uygulamalarında kullanılmak üzere yüksek hassasiyetli, düşük güç tüketimli, elektronik cihaz tasarımına optimize edilmiş BMP180 basınç sensörü ile elde edilmiştir. Resim 3.2’de hava basınç sensörü BMP-180 gösterilmiştir.



Resim 3.2. Hava basınç sensörü

Yüksek doğrusal durum çalışmasında 0,25m hassasiyetle iyi performans göstermektedir. I2C (Inter-Integrated Circuit) arayüzü ile gömülü sistemler ile uyum içerisinde çalışmaktadır. I2C haberleşme protokolü, SDA (Serial Data Line) ve SCL (Serial Clock) kanallarına sahip eşzamanlı haberleşmeyi sağlamaktadır. SCL veri alış verişiindeki bozulmaları engellemek için kullanılan saat darbelerini ifade ederken, SDA gömülü sistem ile sensör arasındaki veri yolunu ifade etmektedir [33]. Piezo dirençli teknolojiye dayanan yapısı ile yüksek doğrulukta sonuç vermektedir. BMP-180 dijital hava basıncı sensörü 3,3V çalışma voltajına sahip olup 300-1100hPa aralığındaki basınç ve 0-85 °C sıcaklık değerini ölçmektedir.

3.1.3. Nem ve sıcaklık sensörü

Özellik seçim algoritmasına uygulanan veri kümesinde bağıl nem değerinin yol yüzey durumuna etkisi yüksektir. Bağıl nem, hava kütlesi içerisinde bulunan mutlak nem miktarının maksimum nem miktarına oranı olarak ifade edilir. Buz oluşumu için yol yüzeyinde su bulunması ya da havada yüksek oranda nem olması gerekir. Tasarlanan yol sensöründe havanın bağıl nem miktarını ölçmek için DHT22 sensörü kullanılmıştır. Resim 3.3’de nem ve sıcaklık sensörü gösterilmiştir.



Resim 3.3. Nem ve sıcaklık sensörü.

3,3-6V aralığında çalışabilen polimer kapasiteli sensör ile %0-100 aralığında %2 oranda doğruluk ve %0,1 çözünürlükle 20 metre mesafeden bağıl nem ölçümü yapabilmektedir. Bunun yanında aynı sensör paketi içinde sıcaklık ölçümünde -40 °C ile 80 °C aralığında $\pm 0,5^{\circ}\text{C}$ doğruluk ve $0,1^{\circ}\text{C}$ çözünürlük değerinde çalışmaktadır. Bu sensörün en büyük zayıflığı ardışık ölçüm yapmak istendiğinde yaklaşık iki saniye gibi bir cevap süresinin bulunmasıdır [34].

3.1.4. Yol sıcaklık sensörü

Yol durum tahmin çalışmalarında buzlanmayı etkileyen bir diğer unsurda yol sıcaklığıdır. Hava sıcaklığı suyun donma noktasına gelse bile yol sıcaklığının etkisinden ötürü buzlanma gerçekleşmeyebilmektedir. Bu nedenle buzlanma tahmini başarımının yüksek çıkması için yol sıcaklığının tahmin algoritmasında olması gerekir. Özellik seçimi işlemi de bu özellik değerinin sınıflandırıcıda olması gerekliliğine yüksek oran vermektedir. Resim 3.4’de yol sıcaklık sensörü gösterilmiştir.



Resim 3.4. Yol sıcaklık sensörü

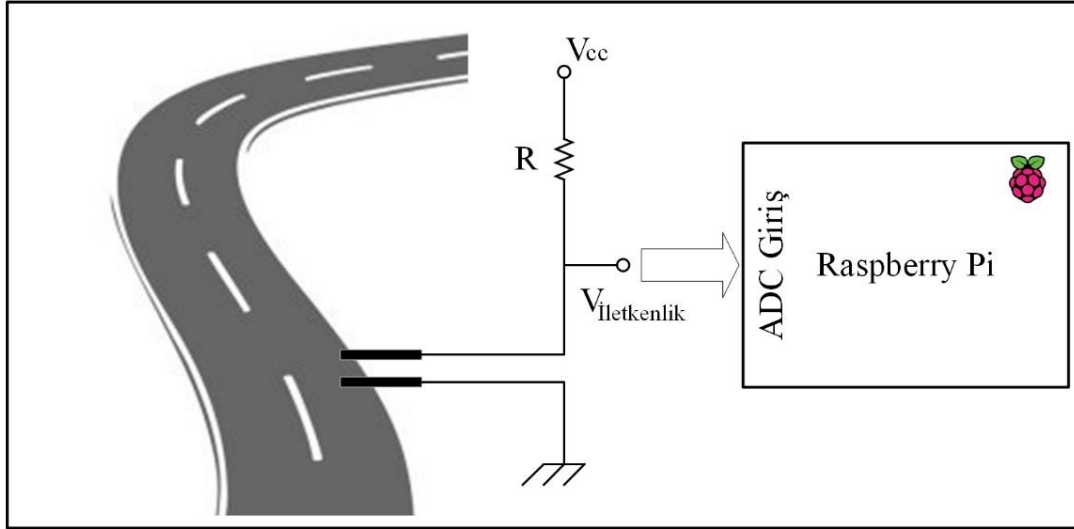
Bu tez çalışmasında yol altı sıcaklığı DS18B20 dijital sensör ile ölçülmüştür. DS18B20 sensörü, 9 - 12 bit arası çözünürlükle tek telli veri yolu üzerinden haberleşme ile sıcaklık ölçümü yapabilmektedir. 3-5,5V aralığında çalışan istenirse harici besleme girişini kullanmadan veri hattı üzerinden besleme yapılarak sıcaklık ölçüm işlemi yapabilmektedir. -55 °C ile 125 °C arasındaki sıcaklık değerini -0,0625 °C hassasiyetle ölçmektedir. DS18B20 sensörünün her birinin 64 bitlik seri numarası sayesinde tek telli hat üzerinde çok sayıda sıcaklık sensörü ile ölçüm yapılabilir.

3.1.5. Yol yüzey iletkenlik sensörü

Sürücülerin buz gibi tehlikeli yol koşullarına maruz kalmaması için yol bakım birimlerince yol yüzeyi tuzlanmaktadır. Yol yüzeyine serpilen tuz donma noktasını düşüreceği gibi yüzeyde bulunan mevcut buzı da eritmektedir. Uzaktan yol yüzey algılama sistemlerinde yol yüzeyinden örnek alınmamasından dolayı yapılan çalışmaların başarı oranı düşmektedir.

Bu çalışmada yol yüzeyine serpilen tuzun tespiti için iletkenlik sensörü kullanılmıştır. Elektriksel iletkenlik malzemenin üzerinden elektrik akımının geçebilmesidir. Metal malzemelerde atomun son yörüngesinde bulunan boşta elektronlar iletkenliği sağlarken sıvı malzemelerde çözünmüş iyonlar elektrik iletkenliğini sağlar. Sıvıların iletkenliğine elektrolitik iletkenlik denmekte ve birimi S/m (Siemens/metre) olarak kullanılmaktadır. Elektrolitik iletkenlik ölçümü için konduktif ve indüktif iletkenlik problemleri kullanılmaktadır. İndüktif iletkenlik ölçüm probu içine yerleştirilen iki bobinden meydana gelmektedir. İletkenlik ölçümü için sıvı içerisine daldırılan sensör içerisindeki birinci bobin manyetik alan

oluşturur ve sıvı içerisinde çözülmüş iyonları hareketlendirerek ikinci bobin üzerinde akım indüklenmesine neden olur bu indüklenme değerine göre iletkenlik ölçümü yapılmaktadır. Konduktif iletkenlik ölçümünde ise metal iki plaka üzerine verilen sabit gerilim sonucu plakalar arasında kalan sıvı içerisindeki iyon hareketleri sonucunda oluşan akımın değeriyle iletkenlik ölçülmektedir. Şekil 3.2’de yol yüzey iletkenlik ölçüm diyagramı gösterilmiştir.



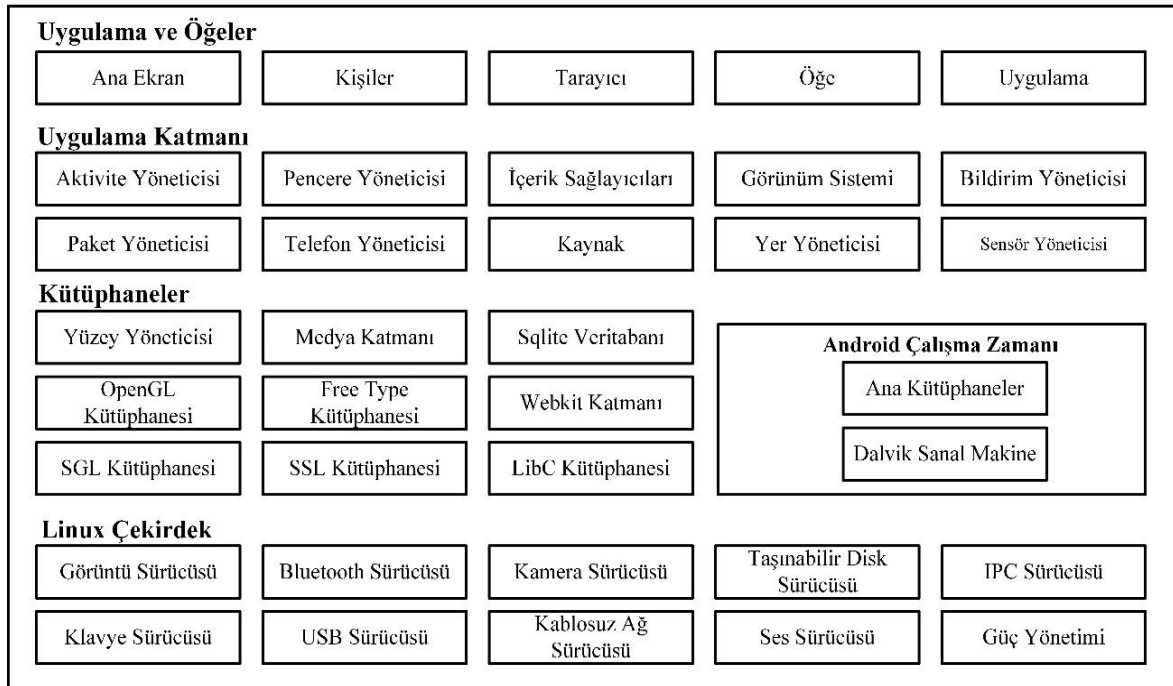
Şekil 3.2. Yol yüzey iletkenlik sensör şeması

Yapılan bu ölçümde elektrot sabiti de hesaba katılmalıdır. Ölçüm elektrot sabiti karşılıklı yerleştirilen metal elemanların birbirlerine olan mesafesinin alan ölçülerine oranıdır. Elektrot sabitinin değeri büyüdükçe elektrotun iletkenlik ölçüm aralığı da büyümektedir. Konduktif ölçüm işleminde alternatif gerilimin kullanılması daha uygundur. Doğru akım uygulanmasında negatif gerilim uygulanan plakaya pozitif iyonlar, pozitif gerilim uygulanan plakaya negatif iyonlar toplanarak metallerin uçlarında kristalleşme olur ve ölçüm hassasiyeti bozulur. Konduktif ölçümün basit yapısı, düşük maliyeti, istenilen prob uzunluğu, ölçüm hassasiyeti ve uygun dönüştürücüleri kullanabilmesi üstünlüğü iken probun direk sıvı teması nedeniyle kirlenmeden kaynaklı ölçüm hassasiyetinin düşmesi zayıflığıdır.

3.1.6. Android Studio

Yol durumu uyarı sisteminin sürücülerini uyarması için Android tabanlı işletim sistemine sahip akıllı cep telefonu uygulaması geliştirilmiştir. Android Studio; tümleşik geliştirme ortamı olarak da bilinen akıllı cep telefonları için uygulama geliştirme platformudur.

Uygulama yapabilmek için Java destekli kütüphanelerin Android Studio'ya eklenmesi gerekmektedir. Android Studio yüklenme esnasında temel özellikler standart olarak kurulmaktadır. Bunun yanında SDK (Software Development Kit) aracılığıyla Emülatör, Sanal cihaz, araçlar ve dokümanların bulunduğu yardımcı eklentilerin yüklenmesi gerekir. Android, taşınabilir cihaz teknolojileri için kullanılan açık kaynak kodlu işletim sistemidir. Linux çekirdek sistemine sahip kapsamlı bir mimarisi bulunmaktadır [35]. Şekil 3.3'de Android işletim sistemi mimarisi gösterilmiştir.



Şekil 3.3. Android işletim sistemi mimarisi

Android işletim sistemi, akıllı cep telefonları gibi taşınabilir cihazlar için ücretsiz açık kaynaklı bir işletim sistemi ve geliştirme platformu olarak çalışmaktadır. Uygulama geliştiricilerin Java programlama bilgisine sahip olması büyük kolaylık sağlayacaktır. Uygulama geliştirme sırasında Activity, Layout, Service ve Intent gibi kullanılan sınıflar bulunmaktadır [36].

Activity; tasarlanan uygulamanın ana sayfası gibi düşünülebilir. Yani uygulamayı kullanan kişinin ekran üzerinde kullandığı ya da faydalandığı sınıflar Activity sınıfıdır. Buton, TextView, EditText ve ListView gibi ekranda kullanılan öğelerdir. Bu Activity sınıfını kullanmak için AndroidManifest.xml içerisinde ne şekilde kullanacağına dair tanımlama

yapmak gerekmektedir. Bu dosya içerisinde tanımlanması yapılmayan sınıf uygulamada çalışmayacak ve hata verecektir.

AndroidManifest.xml dosyasına uygulama açıldığında başlangıç ekranının <intent-filter> etiketi altında belirterek kullanıcıya sunmak mümkündür. Layout klasörü içindeki tasarımı yapılan bileşen ile Activity sınıfı ilişkilendirilerek tasarımın gereken sınıf faaliyeti gerçekleştirilmektedir. Uygulama çalışırken herhangi bir hata vermemesi için aktivite sınıflarının yaşam döngüsüne dikkat etmek gerekmektedir. Kurulan sınıflara ait program çalışma sürecinin yaşam döngüsü bulunmaktadır.

Layout; uygulamada kullanılan ekran bileşenlerinin tasarımını, nasıl gruplanacağı ve ekrandaki görsel düzeninin esas oluşturulduğu alandır. RelativeLayout, LinearLayout gibi çeşitleri bulunan geliştiricinin tercihinin bırakılan ve kodlamayı hızlandıran yapılarıdır.

Service; kullanıcı yüzü olmayan internet, veri tabanı gibi arka planda ya da uygulama kapalıyken dahi çalışabilen kullanıcıdan bağımsız işlem yapan Android sınıflarıdır. Bu özelliklerinden dolayı Activity sınıflarına göre daha öncelikli seviyeye sahiptirler [35]. Service sınıfları ile aktivite sınıfları arasında intent ya da broadcastreceiver kullanılarak iletişim kurmak gerekebilir ve service sınıfları AndroidManifest.xml dosyasında belirtilmelidir.

Intent; bir uygulamadan başka uygulamaya geçişte, Activity arasındaki geçişlerde veya servis sınıflarını çalıştırmada kullanılan nesnelere intent denir. Açık (explicit) ve kapalı (implicit) şeklinde iki tür intent bulunur. Açık intent programın içindeki bir bileşeni ya da sınıfı çağırmak için kullanılırken, kapalı intent ise başka bileşeni örneğin tarayıcıyı açmak gibi işlemlerde kullanılmaktadır.

Android application package (APK) uzantılı dosyalar Android işletim sisteminin çalıştırdığı dosya tipidir. APK uzantılı oluşturulan uygulama dosyaları mobil cihazları yüklenerek kurulumu yapılabilir. Bu tez çalışmasında araç sürücüler için yol durum uyarı sistemi Android uygulama aracılığıyla yapılmış ve APK dosyası kurularak sensör düğüm koordinatlarında test edilmektedir.

3.1.7. Python

Raspberry Pi tek kartlı bilgisayarın işletim sistemi içerisinde hazır gelen python programlama dili oldukça sade, kolay yazılabilen ve akıcı komut yapısına sahip bir programlama dilidir. Guido Van Rossum 1991 senesinde geliştirdiği python dili yüksek seviyeli bir programlama dili olup, herhangi bir platform kısıtlaması bulunmamaktadır. Bu nedenle Windows, Linux, Mac OS, Android, iPhone ve Macintosh gibi farklı donanım ve işletim sistemleriyle çalışmaktadır [37]. Python programlama dilinin komut yapısının kolay olması hem program yazmayı, hem de yazılan programı okumayı diğer programlama dillerine göre daha kolay kılmaktadır. Python dili büyük kullanıcı ağı ve aktif geliştiricileri ile en sık kullanılan ilk on programlama dillerinden biridir. Bu nedenle istikrarlı çalışan ve kuvvetli bir programlama dili olması sebebiyle yalnızca bireysel kullanılan değil büyük şirketler tarafından da yoğun kullanılan programlama dilidir.

Python programlama dili açık kaynak kodlu ve ücretsiz yayınlanan programlama dilidir. Bu sebeple çok geniş kapsamlı makine öğrenme, veri tabanları, grafik ara yüzü gibi birçok kütüphane aracılığıyla geniş uygulama alanına sahiptir. Python dilinin çok yaygın kullanılmasının nedenlerinden biri de veri hesaplama işlemindeki fonksiyonel modüllerin, kütüphanelerin ve paketlerin zenginliğidir. Python programlama dilinde en yaygın Numpy, SciPy ve Matplotlib gibi kütüphaneler hesaplamalı bilimlerde kullanılmaktadır. Numpy kütüphanesi vektör, matris ve dizi işlemleri gibi sıralama, seçim, temel istatistik, mantık ve matematiksel işlemlerin daha fazlasının yapıldığı özel bir kütüphanedir. SciPy kütüphanesi fiziksel ve matematiksel problemlerin dijital ortamda ifade edilmesine, Matplotlib kütüphanesi ise verilerin iki boyutta görselleştirilerek anlaşılmasını kolaylaştırmakta kullanılmaktadır.

Bu tez kapsamında kullanılan gömülü sistem üzerinde Python dili kullanılarak sınıflandırma yapılmış ve sınıflandırma sonucu web sunucusuna gerekli kütüphaneler kullanılarak gönderilmiştir.

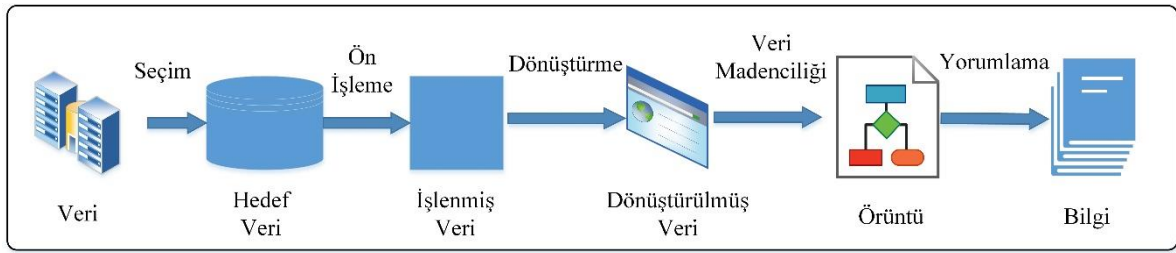
3.2. Metot

Akıllı yol durum tespit ve uyarı sistemi için gerçekleştirilen yol durum sensöründen alınan gerçek zamanlı veriler yardımıyla yol yüzeyi hakkında sınıflandırma yapılmaktadır.

Sınıflandırma işleminden önce özellik seçim ve yüksek performans gösteren algoritma seçim işlemi yapılmaktadır. Sınıflandırma sonucu gömülü sistem üzerinden gerçek zamanlı çalıştırılarak internet aracılığıyla akıllı cep telefonu uygulamasına gönderilmektedir.

3.2.1. Makine öğrenmesi

Veri madenciliği çevreden toplanan veriler içerisinde algoritma ve yardımcı araçlar ile işe yarar bilginin ortaya çıkarılmasıdır. Toplanan verilerin farklılıklarını, birbirleri arasındaki ilişkilerini, desenlerini, analiz ederek bilgiye dönüştürme işlemidir. Veri madenciliği tıp, genetik, eğitim ve bankacılık gibi farklı disiplinlerde kullanılmaktadır. Veri madenciliği büyük veri yığınları içerisinde doğrudan görülemeyen desenleri yani ilişkiyi matematiksel ve istatistiksel yöntemler aracılığıyla açığa çıkaran yöntemdir [38]. Fayyad ve diğerleri tarafından veri madenciliği süreci Şekil 3.4' deki gibi ele alınmıştır [39].



Şekil 3.4. Veri madenciliği süreci

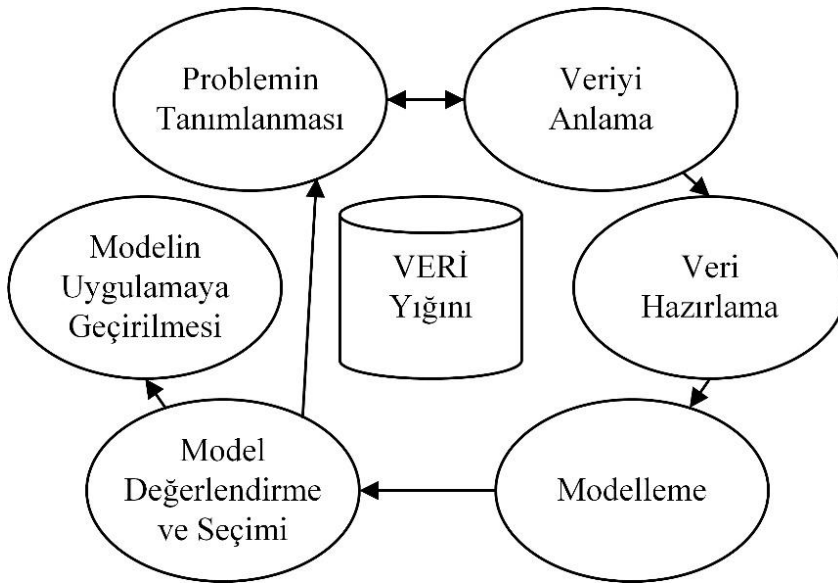
Fayyad'ın sürecine göre ilk aşama verilerin seçilmesi ikinci aşama ham verinin analiz edilebilir sürece hazırlanması (ön izleme), analiz için veri formatının uygunlaştırılması (dönüştürme), veri madenciliği algoritmalarının uygulanması (veri madenciliği) ve son olarak elde edilen sonuçların (değerlendirme) değerlendirilip yorumlama aşamasıdır.

Makine öğrenmesi ve veri madenciliği büyük veri yığınlarından bilgiyi elde ederken ortak bir süreç izlemesi olasıdır. Literatürde büyük veri yığınlarına sahip problemlerin çözülmesinde bazı makine öğrenmesi yaklaşımları bulunmaktadır. Hu çalışmasında makine öğrenmesi modelini veri toplama, veri ön işleme, eğitim verisi üzerinde modelin kurulması, test verisi üzerinde modelin değerlendirilmesi ve alınan iyi sonuç neticesinde uygulamaya geçilmesi olarak beş adımda açıklamıştır. Brownlee gerçek hayattaki problemlerin çözümünü; problemi tanımlama, veri analizi, veri hazırlanması, algoritmanın

değerlendirilmesi, algoritma sonuçlarının iyileştirilmesi ve sonuçların sunulması olarak altı basamakta çözmüştür.

Rossi dört basamakta sınıflandırma problemlerinin çözümünü; girdiye karar verme, eğitim setinin ayrılması, sınıflandırıcı algoritmasının eğitilip performansının tahmini ve sınıflandırıcı algoritmasının sonuçlarının değerlendirilmesi olarak sıralamıştır [40].

İki yüzden fazla uzman ve Endüstri önderleri veri madenciliği için çapraz endüstri standart süreç modeli CRoss-Industry Standard Process for Data Mining (CRISP) geliştirmiştir [41]. CRISP makine öğrenmesi uygulaması problemlerin çözümünü; problemin tanımlanması, veriyi anlama, veriyi hazırlama, modelleme, modelin değerlendirme ve seçimi, modelin uygulamaya geçirilmesi şeklinde altı başlıkta incelemektedir. CRISP Modelinin akış diyagramı Şekil 3.5’de gösterilmektedir.



Şekil 3.5. CRISP modeli akış şeması

3.2.2. Problemin tanımlanması

Seçilecek makine öğrenmesi algoritmasının yüksek performans gösterebilmesi için öncelikli amaç problemin doğru tanımlanmasıdır.

3.2.3. Veriyi anlama

Makine öğrenmesi algoritmasının düzgün çalışabilmesi için çözülecek probleme uygun veriler temin edilmelidir. Bu veriler internet ortamı gibi farklı kaynaklardan yararlanılarak elde edilebilir. Kullanılan makine öğrenmesi algoritmalarının performanslarını geliştirmek açısından internet kaynaklı University of California Irvine Machine Learning Repository yayınlanan veri setleri ile mümkün olmaktadır [42]. Makine öğrenmesi sürecinde en çok vakit alan süreçlerden biri de veri yığınının hazır hale getirilmesi sürecidir.

Öğrenme modelinin yüksek performansta çalışması için veri yığınının tam manası ile anlaşılması gerekir. Temin edilen veri seti üzerinde bazı matematiksel ve istatistiksel hesaplamalar yapılabilir. Eldeki niteliklerin nümerik veya kategorik olma şekline göre min, max, mod, medyan ve ortalama gibi işlemler yapılabilirdiği gibi histogram kutu grafiği, sütun ve pasta grafikleri ile görselleştirilerek ön fikir edinmesi sağlanır.

Kutu grafiği; sürekli değişkenin dağılımının yanı sıra büyüklük değerini gösteren bir grafikdir. Kutu grafiği çarpıklık, basıklık ve yayılma açısından eldeki verileri açıklarken aykırı (outlier) veri değerini rahat görmemizi sağlar. Histogram grafiği; verinin hangi sıklıkla dağıldığını gösteren x ekseninde sürekli değişkene ait aralıkları y ekseninde ise verinin frekans değerini gösteren grafik türüdür [43, 44].

3.2.4. Veriyi hazırlama

Veri analizine geçmeden önceki en önemli basamaklardan biri de verinin işleme hazır hale getirilmesidir. Han ve Kamber(2006) tarafından veri indirgeme, veri bütünleştirme, veri özetleme, veri temizleme, veri ayırıklaştırma ve kavram hiyerarşisi yaratma başlıkları altında incelenmektedir.

Eksik/Kayıp değerler ile çalışma

Toplanan veri yığnında herhangi bir niteliğe/özellğe ait olan değer hastalık, ölüm, analizin hatalı yapılması ve diğr dış etmenlerden dolayı istenmeyen kayıp değerler oluşabilir. Literatürde kayıp değer (missing value) ile algoritmaların çalışması bazı problemlere sebep olmakta bu da sınıflandırma performansını düşürmektedir [45, 46]. Kayıp değerlerin

oluşturduğu boşluğu tamamlama işlemi bu adımda yapılmalıdır ve bunun için çeşitli yöntemler kullanılmaktadır [47].

- Ölçülen niteliklerinde neticesinde oluşan sınıf değerinin bilinmediği ya da kaydedilmedi durumlarda ilgili kaydın tüm satırı silinebilir.
- Kayıp değere “bilinmeyen”, “N/A” şeklinde sabit ataması yapılabilir o niteliklerin önemli olmadığı anlarda bu yöntem tercih edilmektedir.
- Kayıt altına alınamayan nitelikler bulunduğu sütunun ortalama değeri hesaplanarak kayıp değer yerine yazılabilir, sınıf değerlerine ait kayıp değer kendi sınıf niteliklerin ortalama değeri şeklinde yazılır.
- Kategorik özelliklerde kayıp değer yerine (mode) en sık tekrarlayan nitelik yazılır.
- Kayıp değerlerin doldurulmasında istatistiksel yöntemler kullanılabilir. Kayıp değerlerin bulunması da sınıflandırma problemine dönüşebilmektedir.
- Kayıp değeri olmayan sınıfın değerlerinden yararlanarak yapılan tahminlerle diğer sınıfların kayıp değerleri doldurulabilir.
- Regresyon analizi yardımı ile kayıp nitelikler tahmin edilebilir.
- Tekrar eden niteliğin kayıp olması durumunda en son aldığı değer kayıp değer yerine yazılabilir.

Normalizasyon (Veri normalleştirme)

Makine öğrenmesi algoritmalarında nümerik değerlerin aralıkları uygun olmayabilir. Veri yığınınındaki niteliklerin sayısal değerleri birbirinden ayrı olması durumunda normalizasyon yani veri normalleştirme veri setine uygulanır. Çünkü veri yığınınındaki niteliklerin ortalama ve varyanslarının birbirlerinden anlamlı ölçüde aykırı olmaları doğruluk ve algoritma performansını etkilemektedir. Bu sebeple algoritma üzerindeki veri ağırlıklarını eşitlemek için normalizasyon yöntemleri veri seti üzerine uygulanır [48]. Normalizasyon için aşağıdaki yöntemler kullanılmaktadır.

Minimum maksimum normalizasyonu; bu yöntemde öz nitelikleri değerleri doğrusal biçimde normalize edilmektedir. Min-Max normalizasyonu uygulanmış veri setindeki değerlerin birbirleri arasındaki büyüklük sabit kalır ve genelde 0-1 aralığında bulunur [49]. Eş. 3.1’de Min-Max modelinin denklemi verilmiştir.

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (3.1)$$

Ondalık ölçekleme normalizasyonu; bu yöntemde veri yığını içerisindeki niteliklerin değerleri basamak sayısı kadar ondalıklı kısım kaydırılarak +1 ile -1 arasında ondalık değere dönüştürülür. Eş. 3.2’de ondalık ölçekleme denklemi verilmiştir.

$$x' = \frac{x_i}{10^j} \quad (3.2)$$

$\text{Max}(|x'|) < 1$ eşitliğini sağlayan minimum tam sayı j olmak üzere

Z-değeri (Z-skor) normalizasyonu; bu normalleştirmede nitelik değerinin standart sapması ve ortalamasına bağlı olarak dönüşüm yapılır Z-değeri eksi sonsuz ile artı sonsuz aralığında iken genellikle -1,5 ile +1,5 aralığına çekilir ve denklemi Eş. 3.3’de gösterilmektedir.

$$x' = \frac{x_i - \mu_i}{\sigma_i} \quad (3.3)$$

Sigmoid normalizasyonu; bu normalizasyonda hiperbolik tanjant ile lojistik sigmoid fonksiyonları kullanılarak veri setindeki değerler 0-1 ya da -1,+1 aralığına çekilir [50, 51]. Eş. 3.4 ve 3.5’de sigmoid fonksiyonlar gösterilmektedir.

$$x' = \frac{1}{1 + e^{-x_i}} \quad (3.4)$$

$$x' = \frac{e^{x_i} - e^{-x_i}}{e^{x_i} + e^{-x_i}} \quad (3.5)$$

Uç noktalar (Outliers)

Veri yığını içerisindeki niteliklerin arasında ölçülen değerlere uzakta olan ya da sıra dışı gözlem olarak kaydedilen verilere uç noktalar (Outliers) denir [52, 53]. Veri yığını kaydı esnasında çevresel gürültü ya da ölçüm hatasından kaynaklanan durumlarda toplanan değerler uç nokta olarak kaydedilir.

3.2.5. Modelleme

Farklı makine öğrenmesi algoritmaları yardımıyla veri yığını içerisinde bulunan gizli deseni ortaya çıkarmak algoritmaların görevidir [54]. Veri yığını içerisindeki nitelik değerlerin sınıflara çıkaran haritalamayı model olarak adlandırılmıştır. Öğrenme algoritmaları tarafından ortaya çıkan problemlerin çözüldüğü ve çözümleme için yapılacak görevlerin modeller aracılığıyla yaptırıldığı bilinmektedir [55].

Makine öğrenmesi alanında çok sayıda model geliştirilmiştir. Bu modeller sınıflandırma, kümeleme, regresyon ve birliktelik kuralı olarak gruplanır. Bu tez çalışmasında sınıflandırma algoritmaları ile çalışıldığı için sınıflandırma modelleri açıklanmıştır.

Sınıflandırma bağımlı değişkeni bilinmeyen veri seti değerlerinin hangi sınıfa ait olduğunu bulmaya yarar. Sınıflandırma algoritmaları temel prensip olarak veri yığınının üçte ikisini eğitim için, üçte birini algoritmanın testinde kullanmaktadır. Eğitilen sınıflandırma modeli sonucunda sınıflandırıcı kuralı oluşturur ve sınıflandırma bu kurala uygun şekilde tahmin edilir [56].

3.2.6. Model değerlendirme ve seçimi

Bilgisayar destekli seçilen modelin başarısını sınıflandırıcıların hangisinin daha iyi olduğunu bulabilmek için çeşitli yöntem ve ölçülere ihtiyaç duyulmaktadır. Kullanılan bu ölçütler modelin ne ölçüde başarılı olduğunu değerlendirmesi bakımından önemlidir. Sınıflandırma sonucu elde edilen veriler karışıklık matrisi olarak tanımlanmaktadır. Karışıklık matrisinin diyagonali algoritmanın doğruluk oranı hakkında bilgi vermektedir. Karışıklık matrisi yatayda gerçek sınıf değerlerini düşeyde tahmin edilen sınıf değerlerini barındıran bir kare maddesidir. Matrisin diyagonali değerleri doğru sonuçların toplamını vermektedir [57, 58]. Çizelge 3.1 ‘de iki sınıfa ait karışıklık matrisi gösterilmiştir.

Çizelge 3.1. Karışıklık matrisi

	Tahmin Pozitif	Tahmin Negatif
Gerçek Pozitif	DP	YN
Gerçek Negatif	YP	DN

Karışıklık matrisinde bulunan Doğru Pozitif (DP) gerçekte buzlu yol yüzeyi sınıfında bulunan örneklerin buzlu yol yüzeyi olarak tahmin sayısını, Yanlış Negatif (YN) gerçekten buzlu yol yüzeyini kuru yüzey olarak sınıflandırılmış örnek sayısını göstermektedir. Yanlış Pozitif (YP) gerçekte kuru yol yüzeyi sınıfı iken tahminde buzlu yol yüzeyi olarak etiketlenen örnek sayısını, Doğru Negatif (DN) gerçekte kuru yol yüzeyi iken tahminde de kuru yüzey olarak örneklenen sayısını ifade etmektedir.

Bu tez çalışması içeriğinde sınıflandırma algoritmalarının başarılı olup olmadığına karar verdiğimiz ölçütler bulunmaktadır. Bunlardan en sık kullanılanları doğruluk, dengeli doğruluk, kesinlik, f ölçütü, duyarlılık, seçicilik ve Mathews Corelation Coefficition (MCC) ölçütleridir.

Doğruluk

Model başarımının kullanımında en yaygın ve popüler olan yöntem doğru tahmin oranıdır sınıflandırma algoritmasının vermiş olduğu doğru yanıtların başarısı hakkında değer verir [59]. Eş. 3.6’da doğruluk oranı formülü verilmektedir.

$$\text{Doğruluk Oranı} = \frac{DP + DN}{DP + DN + YP + YN} \quad (3.6)$$

Doğruluk oranı, karışıklık matrisinin diyagonal değerlerinin toplamının Matrisin bütün değerlerinin toplamının oranı olarak bilinmektedir. Hata oranı ise doğruluk oranının 1’e tümleyeni ile bulunur ve formülü Eş. 3.7’de gösterilmektedir.

$$\text{Hata Oranı} = 1 - \text{Doğruluk Oranı} \quad (3.7)$$

Dengeli doğruluk

Dengeli doğruluk ölçütü algoritmaların değerlendirilmesinde doğruluk ölçütü yoğun olarak kullanılmasına rağmen daha objektif bir yaklaşıma sahiptir. Doğruluk ölçütünde sınıflandırıcı tek bir sınıfın doğruluk oranına bakarken (buzlu veya kuru) dengeli doğrulukta her bir sınıfın doğruluk değerinin ortalaması alınmaktadır. Kısacası yol yüzeyinin kuru tahmin edilmesinin yanı sıra buzlu, tuzlu nemli ve nemli olarak da doğru tahmin edilmesi

gerekmektedir. Her bir sınıfın sınıflandırma tahmini kendi içinde önem arz etmektedir [48]. Dengeli doğruluk denklemi Eş. 3.8 'deki gibidir.

$$\text{Dengeli Doğruluk Oranı} = \frac{1}{2} \left(\frac{DP}{DP + YN} + \frac{DN}{DN + YP} \right) \quad (3.8)$$

Kesinlik

Buzlu olarak doğru tahmin edilen örnek sayısının sınıflandırıcının buzlu olarak tahmin ettiği değerlere oranıdır [48]. Bu doğrultuda kesinlik ölçütü sınıflandırma algoritması tarafından doğru tahmin edilen gerçek değerlerin doğru olarak sınıflandırılmış toplam verilere oranıdır. Eş. 3.9 'daki gibi hesaplanır.

$$\text{Kesinlik Oranı} = \frac{DP}{DP + YP} \quad (3.9)$$

F-ölçütü

Sınıflandırma algoritmaları başarımında kullanılan kesinlik ve duyarlılık ölçütlerinin harmonik ortalaması bize F ölçütünü verir. Bu değerler hesaplanırken sistemin performansını ölçmekte kesinlik veya duyarlılık tek ölçüt olarak uygulandığında yetersiz kalabilmektedir. F ölçütü Eş. 3.10'daki gibi hesaplanır [48].

$$F_Ölçütü = \left(2 * \frac{\text{Duyarlılık} * \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}} \right) \quad (3.10)$$

Duyarlılık

Sistem tarafından buzlu tahmin edilen yolun gerçekten ne kadarının buzlu olduğunu verir. Gerçekten buzlu olan yolların buzlu olanlarını ayırma ölçüsüdür. Duyarlılık doğru sınıflandırılan pozitif örneklerin toplam pozitif örnek sayısına oranıdır [60]. Eş. 3.11'deki gibi hesaplanır.

$$\text{Duyarlılık Oranı} = \frac{DP}{DP + YN} \quad (3.11)$$

Seçicilik

Sınıflandırıcı tarafından kuru olduğu tahmin edilen yolun gerçekten ne kadarının kuru olduğunu verir. Sınıflandırıcının seçicilik ölçütü kuru yollar içerisinde gerçek kuru yolu ayırma kabiliyetidir [60]. Eş. 3.12'deki gibi hesaplanır.

$$\text{Seçicilik Oranı} = \frac{DN}{DN + YP} \quad (3.12)$$

Matthews Correlation Coefficient

MCC ölçütü sınıflandırma algoritmalarının değerlendirilmesinde doğruluk oranına kıyasla daha doğru ve güvenilir değerler verdiği gözlenmiştir. Eşit olmayan veri dağılımına sahip sınıflandırma algoritması için kullanılan doğruluk oranı genellikle daha yüksek değerlerde çıkar. Böyle bir durumda kesinlik duyarlılık ve f ölçütü değerlerine de bakılmalıdır. Ancak MCC ölçüsü diğer sınıflandırma ölçütleri ne bakılmaksızın sınıflandırma işlemlerinde doğru ve güvenilir sonucu veren en iyi ölçüdür.

Karışıklık matrisi hata matrisi olarak da adlandırılan yöntem ölçüsünün hesaplanmasında kullanılmaktadır. Karışıklık matrisi birden fazla sınıflı veri kümesinin sınıflandırıcı tahmin sonuçlarını içerir. Dengesiz sınıf dağılımlarında bile MCC ölçütü en doğru sonucu vermektedir. MCC sonuç aralığı +1 ile -1 arasındadır. MCC değerinin sifıra yakın olma durumu sınıflandırıcının rastgele sınıflandırma yaptığını gösterirken, sonucun -1 çıkması gerçekte olan değer ile tahmin değerlerinin birbirine zıt olduğunu, sonucun +1 olması ise yapılan sınıflandırmanın tam olarak başarılı olduğunu gösterir [61].

Çok sınıflı durumlarda MCC katsayısı sınıflar için bir karışıklık matrisi cinsinden tanımlanabilir aşağıdaki değişkenler göz önüne alındığında tanımlama basitleştirilebilir. MCC değeri Eş. 3.13'deki gibi hesaplanır.

$$\text{MCC} = \left(\frac{c * s - \sum_k^K p_k * t_k}{\sqrt{(s^2 - \sum_k^K p_k^2) * (s^2 - \sum_k^K t_k^2)}} \right) \quad (3.13)$$

$t_k = \sum_i^K C_{ik}$ K sınıfının gerçekte meydana gelme sayısı.

$p_k = \sum_i^K C_{ki}$ K sınıfının kaç kez tahmin edildiğini.

$c = \sum_k^K C_{kk}$ Doğru tahmin edilen toplam örnek sayısı.

$s = \sum_i^K \sum_j^K C_{ij}$ Toplam örnek sayısı.

3.2.7. Modelin uygulamaya geçirilmesi

Modelin uygulamaya geçirilmesi için model değerlendirme ve seçiminde kullanılan yöntemler belirlendikten sonran problemin çözümü için uygun yöntem kullanılmaya başlanır. Makine öğrenmesinde çözülecek olan problemin nitelikleri iyi anlaşılmalıdır. Toplanan veri yığınınındaki eksik veriler tamamlanarak öğrenme algoritmalarına hazır hale getirilir. Öğrenme algoritmaları hali hazırdaki veri yığını üzerinde işleme alınarak en yüksek performansı gösteren algoritma sınıflandırma modeli olarak belirlenir. Yapılan bu çalışmaların sonrasında herhangi bir problemle karşılaşılıyorsa model uygulamaya geçer.

3.2.8. Sınıflandırma algoritmaları

Sınıflandırma kavramı, veri yığını üzerinde tanımlanmış çeşitli sınıfları veri üzerinde dağıtma işlemidir. Başka bir deyişle veri kümesi içerisinde bulunan tanımlı kayıtlar yardımıyla yeni kayıtların bir kümeye dahil edilmesidir. Tahmin edilecek değişkenle hedef değişken arasındaki matematiksel ilişkiyi veri yığını üzerinde yapılan deneylerle bulunmaktadır. Veri yığını sınıflandırılmasında çeşitli makine öğrenmesi algoritmaları kullanılmaktadır.

K- En Yakın Komşu algoritması (K-NN)

Bu algoritma veri kümesi sınıflandırmasında oldukça sık kullanılır. Veriler arasındaki analitik düzlemdeki uzaklıklara bağlı olarak sınıflandırma yapan algoritmadır. K-NN algoritması bilinen en eski, basit ve etkili makine öğrenmesi algoritmalarındandır [62].

İlk aşamada gözlenen değerler birbirinden bağımsız olarak sınıflandırılır. Kademeli olarak gözlenen değerler düzlem içerisinde gruplanır. Bu sınıflandırma algoritmasında ilk olarak gözlenen değerlerin birbirine olan uzaklıkları belirlenir. Gözlem sonucunda uzaklıkları belirlenen komşu sayısı k ile ifade edilir. Bu nedenle algoritmaya k en yakın komşu algoritması denir [63]. K-NN sınıflandırıcısı benzerlik yoluyla eğitim yapan sınıflandırıcıdır.

K-NN algoritmasının güçlü taraflarından birisi sınıflandırma esnasında eğitime ihtiyaç duymaması ve gürültülü veri setine karşı hassas olmamasıdır [64].

Sisteme dahil olan yeni verinin sınıfı belirlenirken tüm veri örnekleri ile arasındaki mesafe hesaplanır. Hesaplanan bu mesafeler büyükten küçüğe doğru sıralanır. K adedince en düşük mesafe değerini sahip gözlemler alınır. Yeni verinin sınıfı k adet gözlemlerde en çok tekrar eden sınıf değeri olarak belirlenir [65]. Son aşamada bahsedilen yeni verinin sınıfını belirlerken en çok tekrar eden sınıf değeri yerine ağırlıklı oylama yöntemi kullanılabilir. Uzaklıkların tersi ya da tersinin karesi ağırlık olarak kullanılabilir. Sınıf değerini bulmak için hesaplanan ağırlıklar toplanır ve en çok ağırlığa sahip olan belirlenir.

Örnek veri setleri arasındaki mesafe algoritmanın performansını etkilemektedir. Veri setinde kategorik ve nümerik niteliklerin birlikte bulunmaları durumunda uzaklık ölçüleri değişkenlik göstermektedir. Bu uzaklığı hesaplamak için sıklıkla kullanılan Öklid uzunluk ölçüsü tüm kategorik veriler nümerik olduğunda kullanılabilir. Uzunluk ölçülerinde sadece Öklid değil Minkowski, Manhattan ve Chebyshev gibi uzaklık ölçütleri ile veri setleri arasındaki yakınlık hesaplanabilir. Uzaklık ölçülerinin denklemleri Eş. 3.14 ile 3.17’de sırası ile verilmektedir.

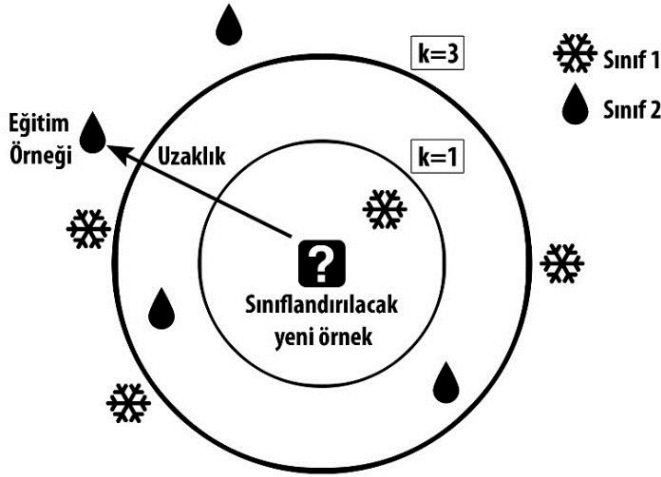
$$d_{Euclidean}(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (3.14)$$

$$d_{Manhattan}(p, q) = \sum_{i=1}^n |p_i - q_i| \quad (3.15)$$

$$d_{Chebyshev}(p, q) = \max_i (|p_i - q_i|) \quad (3.16)$$

$$d_{Minkowski}(p, q) = \left(\sum_{i=1}^n |p_i - q_i|^m \right)^{\frac{1}{m}} \quad (3.17)$$

K-NN algoritmasını açıklamak için 2 boyutlu bir koordinat sistemine Şekil 3.6’daki gibi bir örnek yerleştirilmiştir.



Şekil 3.6. K-NN sınıflandırma algoritması

Sisteme yeni eklenen veri koordinat düzlemi üzerindeki veri kümesinin sınıfı $k=1$ olduğu gözlemlenirken en çok yakın olduğu sınıf buz iken; $k=3$ olduğunda nemli olmuştur. Bu nedenle seçilen k değeri sistemin sınıflandırma başarısını doğrudan etkilediği görülmektedir.

Naive Bayes Algoritması

Sınıflandırma yapmak için istatistiksel verileri kullanan bayes kolay anlaşılabilen ve uygulanabilen bir sınıflandırma tekniğidir. Elde edilen verilerle hedef sınıf arasındaki ilişkiyi değerlendiren ve hedef sınıfın hangi gruba dahil olacağını tahmin eden, tanımlayan bir sınıflandırma algoritmasıdır [66, 67].

Bayes yönteminin makine öğrenmesi problemlerinde yüksek oranda başarılı sonuçlar verdiği gözlenmiştir. Olasılık ilkesinin en üst seviyesi üzerine kurulan bir yöntemdir [68]. Bayes elde edilen olasılıkların maksimum düzeyde doğruluğunu hesaplar. Koşullu olasılıkların belirlenmesinde kullanılan bir yöntemdir. Bayes teoreminde gelişen bir olayın X olma olasılığının bilinmesi durumunda Y olayının olma olasılığı, Y olayının olması durumunda x olayının olma olasılık değeri farklıdır. Bu ters koşulluluk ilişkisini bulan istatistikçi Thomas BAYES'dir [69]. Naive Bayes teoreminin denklemi Eş. 3.18'deki gibidir [70].

$$p(C|F_1, \dots, F_n) = \frac{P(C)p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)} \quad (3.18)$$

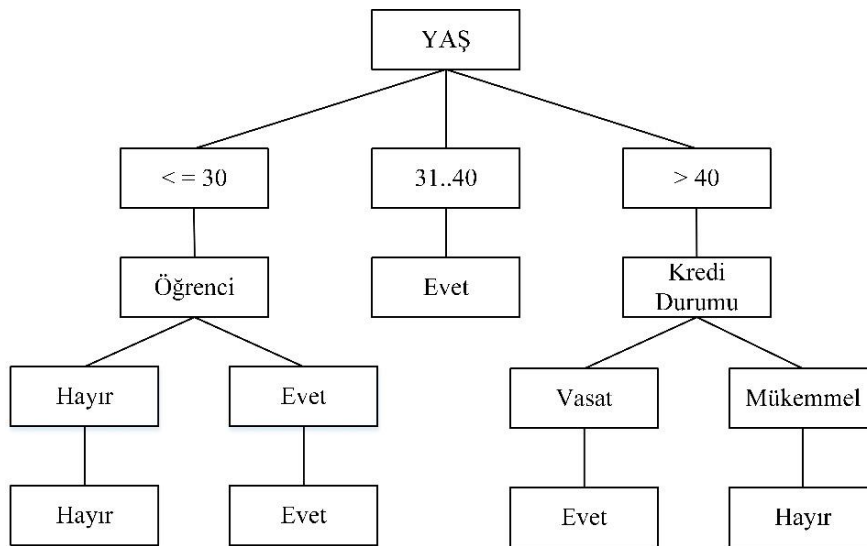
Eşitlik 3.18’de görüldüğü üzere C verilen sınıfı gösterirken F eğitim kümesindeki özellikleri gösterir. Bu sınıflandırıcı tüm koşulluluk olasılıklarının çarpımıdır.

Karar Ağacı Algoritması

Makine öğrenmesi metodunda kullanılan karar ağaçları algoritması kökten başlayarak düğüm noktalarına ilerleyen ve son olarak yapraklara ulaşan ve her ilerlemede incelerek verinin azaldığı ve neticeye daha çabuk kolaylıkla ulaşıldığı bir tahmin algoritmasıdır.

Adından da anlaşıldığı gibi karar ağacı ağaç görünümü verilmiş bir algoritmadır. Verileri uygulanacak olan değerler karar ağacının düğümünü tanımlarken uygulanan değerlerin sonuçlarına dallarda ulaşılır. Eğer bir sonuç belirlenemediği ise yeniden bir karar düğümü oluşturulmalıdır. Uygulamanın sonucunda bir sınıflandırma oluşturulabiliyorsa yaprağa ulaşılmış olur. Bu sayede sınıflandırma da yapmış olunur [71].

Günlük yaşamda karşılaşılan sorunlara uyarılana bilmesi ve hem nümerik hem de kategorik veri ile çalışma imkânı sunar bu sebeplerden dolayı en sık kullanılan metotlardan biridir. Örnek karar ağacı yapısı Şekil 3.7’de verilmektedir.



Şekil 3.7. Karar ağacı algoritması örneği

Quinlan tarafından yapılan bu ağacın yapısı ID3 ayrık değerler ve C4.5 ayrık ve sayısal değerlerle çalışan algoritmaları geliştirilmiştir [72]. Sistemdeki belirsizliğin ölçüsü entropi

kavramı kullanılarak ifade edilir. Entropi bir A kümesindeki olasılıkların toplamını gösteren fonksiyon $S(A)$ olarak Eş. 3.19'daki denklemde ifade edilmiştir.

$$S(A) = - \sum_{i=1}^t p_i \cdot \log_2 p_i \quad (3.19)$$

Eşitlik 3.19'da A indisi sınıfı belirtirken A sınıfı için entropi bulunur. Müşterek hususiyetlerin sınıf bazında ağırlıklı ortalamaları ise Eş. 3.20'deki denklemde görüldüğü gibi hesaplanır [38].

$$S(M, A) = - \sum_{i=1}^t \frac{|A_i|}{|A|} \cdot S(A_i) \quad (3.20)$$

A sınıfına ait müşterek M hususiyet için entropi ölçülür. A toplam değer sayısı A_i ise hususiyeti aldığı değerlerin sayısıdır. $S(A_i)$ ise A hususiyetinin i değeri için ölçülen entropi değeridir. Her hususiyet için bilgi kazanımı denklemi Eş. 3.21'de gösterildiği şekilde belirlenir.

$$Kazanç(M, A) = S(A) - S(M, A) \quad (3.21)$$

En yüksek hususiyet değerlerini sağlayan kazanç ağacın kök durumu olarak belirlenir. Kalan değerler yeniden hesaplanarak en yüksek kazancı veren hususiyet diğer dal ve düğümler olarak belirlenir [60].

Rastgele Orman Algoritması

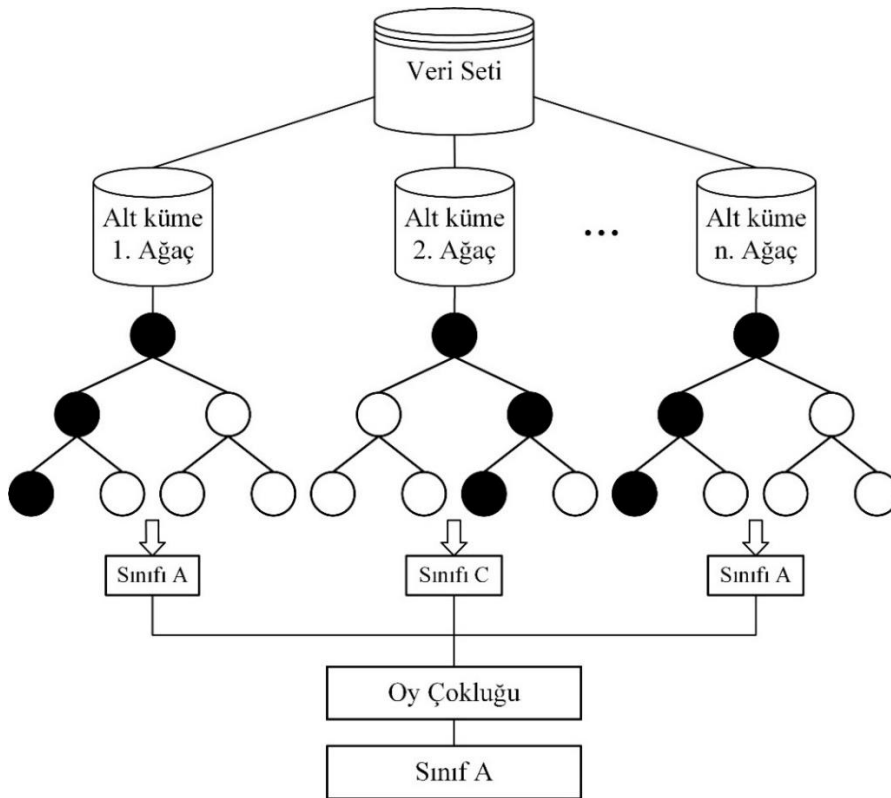
Karar ağaçlarından kullanılan ve bünyesinde oylama yöntemi bulunduran sınıflandırma algoritması olan Rastgele Orman (RO) Leo Breiman ve Adele Cutter tarafından geliştirilmiştir. Müşterek karar ağaçlarının bir araya gelip oylama yapması ile oluşan bir sınıf belirleyicisidir [73].

Veri setinde bulunan bootstrap tekniği ile seçilen örnekler birbirinden bağımsız olan karar ağaçlarını oluşturur. Oluşturulan veri setinin 2/3 ü eğitim 1/3 ü test veri seti olarak ayrılır.

RO de kullanılacak veri kümeleri eşit sınıflar barındıracak şekilde alt k adet veri kümesine dönüştürülür. Bu veri kümeleri de $2/3$ 'ü eğitim ve geriye kalanı da test verisi olarak kullanılır. K adet alt veri seti k adet karar ağacı için kullanılmaktadır.

Oluşturulan ağaçların her birinde test verisi uygulanarak sınıflandırma algoritmalarının hata değerleri hesaplanıp ortalamaları alınır. Bu ortalama hata değeri karar ormanı test hata değeridir. Karar ağaçlarının her birine oluşturulan test hata oranına göre ağırlık verilir. Ağaca verilen ağırlık değeri ile hata oranı ters orantılıdır.

Karar ağacına verilen hata oranı ne kadar yüksek olursa ağaç en düşük ağırlığı alırken verilen hata oranı en düşük olan karar ağacı yüksek ağırlığa sahip olur. Karar ağaçları aldıkları ağırlıklara göre bir sınıflandırma işlemine tabi tutulur ve oylanırlar. Oylanen ağaçlardan en yüksek oyu alan ağaç kazanan sınıf olarak belirlenir [74]. Rastgele orman akış diyagramı Şekil 3.8'deki gibi verilmiştir.



Şekil 3.8. Rastgele orman akış diyagramı

Rastgele orman modeli iki değişken üzerine yapılandırılır ve bir örneklem oluşturulur. Oluşturulan örneklem $2/3$ ü ağacı oluşturmak için kullanılan eğitim veri seti ve geri kalan

1/3'ü oluşturulan modelin iç hata oranını değerlendirmek için test veri seti olarak ikiye ayrılır. Bu örnekleri oluşturacak olan değişkenler için ağaç sayısı (B) ve her düğüm aralığında tesadüfî olarak belirlenecek olan tahminci sayısı (m) dir. Tüm karar ağacı oluşturulurken gerçek veri setindeki gözlem sayısı (n) ile aynı olacak şekilde bootstrap yöntemi ile oluşturulur [75]. Burada dikkat edilmesi gereken en önemli husus oluşturulan karar ağacının budanmaması ve $m < p$ koşulunun sağlanmasıdır. Ancak ağacın aşırı büyümesi de önlenmelidir çünkü aşırı uyum göstermesi istenmez.

Gini indeksi ile değişkenlerin hangi değere göre ayırım yapılacağı belirlenir. Karar ağacında yeni yapılandırılacak dal kalmayana dek bu işlem her düğüm için tekrar edilir. Veri setinin fazla olması ya da veri seti içerisinde kayıp veri olması halinde bile uygulanabilir olması bu algoritmanın üstünlükleri arasında sayılabilir. Oluşturulacak ağaç miktarı algoritmayı kullanacak kişiye bırakılmıştır. Algoritma sürekli ya da ayrık veri değerleri ile de kullanılabilir.

Rastgele orman algoritmasında oluşturulan aşırı ağaç sayısı gerçek zamanlı uygulamalarda algoritmayı yavaşlatır ve istenilen etkiyi vermez. Genellikle hızlı eğitilebilen bir algoritma olmasına rağmen tahmin işlemlerinde geride kalmaktadır. Doğru bir tahmin için daha fazla ağaçla daha yavaş çalışan bir modeldir. Hava koşullarının ani değişime uğradığı kritik yerlerde yol yüzeyindeki ani ve gizli buzlanmaların tahmini için hızlı bir algoritmaya ihtiyaç duyulur. Bu nedenle RO uygulaması gerçek zamanlı çalışma performansı düşük olduğundan ilk sırada tercih edilmemiştir.

Lojistik Regresyon Algoritması

Bağımsız değişkenler örneğin sıcaklık nem basınç ile bağımlı değişken örneğin yol yüzey durumu kuru nemli buzlu arasında bağlantı kuran lojistik regresyon üç başlık altında gruplandırılır. Bağımlı değişken sayısı iki sınıfsa ikili lojistik regresyon ikiden fazla ardışık sınıfsa sıralı lojistik regresyon ikiden fazla sırasız sınıfa sahip ise çok kategorili lojistik regresyon olarak adlandırılır [76].

Lojistik regresyon çalışma mantığı bağımlı değişken değerini bulmak yerine bağımlı değişkenin ikili sınıf örneğin de sınıf değerinin 1 olma olasılığını tahmin etmeyi amaçlar [77]. Çok kategorili lojistik regresyon ile ikili regresyonun çalışma mantığı birbiri ile

örtüşmektedir. Bu sebeple tahmin etmede yardımcı olan X niteliği kullanılarak ikili lojistik regresyona ait π olasılık fonksiyon denklemi Eş. 3.22’de gösterilmektedir.

$$\pi(x) = P(Y = 1|X = x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (3.22)$$

Üstünlük (odds) ifadesi bir durumun olma ihtimalinin olmama ihtimali ne oranı olarak denklemi Eş. 3.23’de ifade edilmiştir.

$$\frac{\pi(x)}{1 - \pi(x)} = e^{\beta_0 + \beta_1 x} \quad (3.23)$$

Bağımlı değişken iki değer aldığında en küçük kareler yöntemi uygulanamadığından logit dönüşüm yapılmaktadır. Bu sayede tahmin yapmak basitleşir ve logit denklemi Eş. 3.24’te gösterilmiştir [78, 79].

$$\text{logit } \pi(x) = g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \ln(e^{\beta_0 + \beta_1 x}) = \beta_0 + \beta_1 x \quad (3.24)$$

Maksimum olabilirlik yöntemi ile lojistik regresyondaki β değeri tahmini yapılmaktadır [80, 81].

Hosmer tarafından çok kategorili lojistik regresyon basit ve anlaşılır şekilde şöyle anlatılmıştır. P adet nitelik ile üç sınıflı bir veri seti ele alalım [82]. İki sınıflı bağımlı değişken olduğunda tek bir logit fonksiyonu ihtiyaç varken üç farklı kategori olduğunda 2 logit fonksiyona gerek duyulur. Bağımlı değişkenlerden biri referans değişken olur geriye kalan iki bağımlı değişken ile karşılaştırılır. Bu duruma ait logit fonksiyonlar Eş. 3.25 ve 3.26’da ve bütün hedef sınıflara ait koşullu ihtimal değerleri Eş. 3.27, 3.28 ve 3.29 de verilmiştir.

$$g_1(x) = \ln\left(\frac{P(Y = 1|x)}{P(Y = 0|x)}\right) = \beta_{10} + \beta_{11}x_1 + \dots + \beta_{1p}x_p \quad (3.25)$$

$$g_2(x) = \ln\left(\frac{P(Y = 2|x)}{P(Y = 0|x)}\right) = \beta_{20} + \beta_{21}x_1 + \dots + \beta_{2p}x_p \quad (3.26)$$

$$\pi_0 = P(Y = 0|x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (3.27)$$

$$\pi_1 = P(Y = 1|x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (3.28)$$

$$\pi_2 = P(Y = 2|x) = \frac{e^{g_2(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (3.29)$$

Bağımlı değişkene ait sınıflar ikili biçimde yani Y yerine Y_0 Y_1 olmak üzere ifade edilir. Bu sebeple n tane tahminden oluşan bir örneklemin koşullu olabilirlik fonksiyonu Eş. 3.30 ve logaritma alınması ile oluşan log-olabilirlik fonksiyonu Eş. 3.31’de verilmiştir.

$$l(\beta) = \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}] \quad (3.30)$$

$$L(\beta) = \sum_{i=1}^n y_{1i} g_1(x_i) + y_{2i} g_2(x_i) - \ln(1 + e^{g_1(x_i)} + e^{g_2(x_i)}) \quad (3.31)$$

Olabilirlik fonksiyonları $L(\beta)$ ’nin birinci kısmı türevleri alınarak bulunmaktadır. $j=1,2$; $k=0,1,\dots,p$ ve $\pi_{ji} = \pi_j(x_i)$ olmak üzere Eş. 3.32’de verilmiştir.

$$\frac{\partial L(\beta)}{\partial \beta_{jk}} = \sum_{i=1}^n x_{ki} (y_{ji} - \pi_{ji}) \quad (3.32)$$

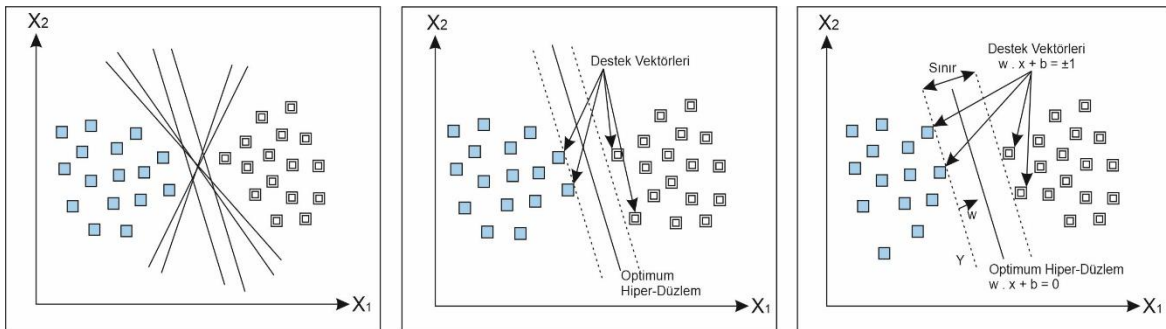
Maksimum olabilirlik tahmincisi β , bu denklemlerin sıfıra eşitlenerek çözülmesi ile elde edilmektedir. Denklemin çözümü varsa maksimum veya minimum bir kritik değeri olduğunu ifade eder. Kritik değer maksimum ise ikinci kısmı türevler negatiftir. Bunun yanında bilgi matrisinin ve kovaryans matrisinin elde edilmesinde de kullanılır [81]. Lojistik regresyonda işlemi alınan niteliklerin birbirleri ile ilişkisinin olmaması gerekir yani nitelikler arasında düşük kovaryans olmalıdır yüksek ilişki olması eş doğrusallık ya da çoklu eş doğrusallık olarak isimlendirilen durumu oluşturur [83, 84].

Destek Vektör Makinesi Algoritması

Destek Vektör Makinesi (DVM) algoritması istatistiksel öğrenme teorisine dayanan regresyon ve sınıflandırma sorunlarına çözüm üreten bir makine öğrenme tekniğidir [85]. DVM'nin en temel uygulaması iki bağımlı değişken veri kümesinin sınıflandırılmasını belirlemek üzere uygulanmıştır. Daha sonrasında da çok sınıflı ve doğrusal olmayan veri kümelerinin sınıflandırılmasında kullanılmıştır. DVM bağımlı değişkenler içerisindeki mesafenin ya da yüzeyin olabildiğince maksimum seviyede ayrışması için çalışır.

Sınıflar arasındaki ayrılmış olan maksimum genişlikteki düzleme hiper düzlem denilir. Burada bahsedilen hiper düzlem kavramı sınıflar arasındaki veri koordinatlarının en yakın mesafesini tanımlamaktadır [86]. DVM algoritması çalıştırılırken karşılaşıcağı iki farklı veri kümesinden bahsedilir. Bunlar doğrusal ayrılabilen ve doğrusal olarak ayrılamayan veri kümeleridir. Gerçek dünyadaki uygulamalarda doğrusal olarak ayrılan veri yapıları kolay sınıflandırılabilirken doğrusal olarak ayrılamayan veri kümelerinde sınıflandırma zorlaşmaktadır. Buna karşın doğrusal olmayan veri setlerinde DVM'nin yüksek oranda doğruluk verdiği gözlenmektedir [38].

DVM algoritması gerçek hayatta birçok alanda uygulandığı gibi akıllı ulaşım sistemlerinde de uygulanmıştır. DVM algoritması iki sınıf arasına sonsuz sayıda sınır koyabilmektedir. Buradaki konulan sınırların en büyük ayrımını oluşturan hiper düzlemi bulmak DVM'nin görevidir [60]. Sınıflar arasında ideal hiper düzlemin oluşumu Şekil 3.9'da gösterilmektedir.



Şekil 3.9. Destek vektör makinesi algoritması

DVM'nin eğitimi için n sayıda örnekten oluşan eğitim verisinin (x_i, y_i) , $i = 1, \dots, n$ olsun. Doğrusal olarak iki sınıfın bir yüzeyle ayrılabilme durumunda düzlemin denklemi Eş. 3.33'deki gibi olur.

$$f(x) = w^t \cdot x + b = \sum_{i=1}^n w_i x_i + b \quad (3.33)$$

Denklemde w ağırlık vektörünü yani düzlemin normalini, b sabiti sapma değerini gösterirken bu değişkenlerin değerleri düzlemin konum pozisyonunu değiştirmektedir. DVM'de sınıflar doğrusal ayrılabiliriyorsa $y_i = +1$ ve $y_i = -1$ etiketli sınıflara aynı uzaklıkta ayıran en iyi düzlemin bulunmasıdır. Eğitim kümesinin Eş. 3.34 ve 3.35'deki denklemleri sağlaması durumunda

$$y_i = +1 \text{ için } w^t \cdot x_i + b \geq +1 \quad (3.34)$$

$$y_i = -1 \text{ için } w^t \cdot x_i + b \leq -1 \quad (3.35)$$

Bu hiper-düzlemleri meydana getiren destek vektörleri $w \cdot x_i + b = \pm 1$ şeklinde ifade edilirler. En uygun düzlemin sınırının maksimum olabilmesi için $\|w\|$ minimum olmalıdır. Buna bağlı olarak sınırlar ise $y_i (w \cdot x_i + b) - 1 \geq 0$ ve $y_i \in [1, -1]$ ifade edilir [87].

Bu optimizasyon eşitliği Lagrange denklemleri aracılığıyla çözümlenebilir ve eşitliğin son hali Eş. 3.36'daki şekli almaktadır [58].

$$L(w, b, a) = \frac{w^2}{2} - \sum_{i=1}^n a_i y_i (w \cdot x_i + b) + \sum_{i=1}^n a_i \quad (3.36)$$

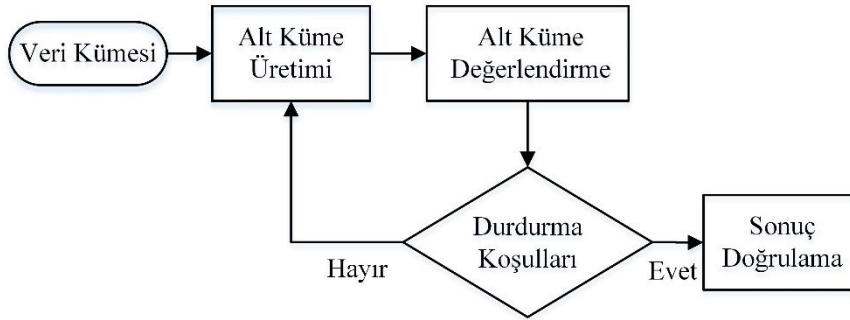
3.2.9. Özellik seçimi

Özellik seçimi bütün niteliklerin tamamının yerine geçecek en ideal alt küme özellikler olarak tanımlanır. Sınıflandırma işleminde eldeki veri yığınının tamamı kullanıldığında sınıflandırıcı en iyi performansı göstermeyebilir. Bunun için veri yığını içerisindeki n adet veriden en iyi sınıflandırma işlemini yapan k adet özellik seçilir [88].

Çözülecek problemin en sade ve en kolay şekilde çözülebilmesi için veri yığını içerisindeki en önemli özellikleri seçmeyi hedefler. Özellik seçme işlemi sayesinde makine öğrenmesi algoritmalarının performanslarını artırıcı yönde etkisi vardır. Toplanan veri içerisindeki gürültü ya da çevresel etmenlerden dolayı hatalı veri kaydı yapılmaktadır ve veri kalitesini düşürmektedir. Özellik seçimi bu durumları telafi ederek gürültülü ve ilgisiz veriyi veri yığınından atarak veri miktarını azaltır ve böylece sistemin hızını artırır.

Veri yığınındaki ilgisiz verilerin atılması ile verinin anlaşılabilirliğine ve içerisindeki gizli deseni ortaya çıkarmamıza yardımcı olur. Özellik seçiminin üstünlükleri kullanılacak algoritmanın işleyeceği veri miktarının azalması sistemin başarısını arttırırken çalışma süresini azaltacak ve depolama alanını düşürecektir.

Özellik seçimi süreci içerisinde öncelikli olarak veri yığınından alt özellik kümesi meydana getirmektir. Sonrasında eldeki nitelikler matematiksel ifadelerle değerlendirilerek niteliğin kümeye alınıp alınmayacağına karar verilir. Kriterleri karşılayan nitelik ile ilgili süreç devam eder kısıtları karşılamayan nitelikler ise alt küme üretimine geri gönderilir. Özellik seçimi süreci Şekil 3.10 'da gösterilmektedir.



Şekil 3.10. Özellik seçimi algoritması

Alt küme üretimi

Eldeki veri yığınından daha düşük boyutta yeni bir alt küme veri yığını oluşturma işlemidir. Alt veri kümesini oluşumu için arama işlemi üç yöntemle; komple arama, ardışık arama ve rastgele arama şeklinde yapılmaktadır. n tane özellik bulunan bir veri yığını için 2^n kadar alt küme meydana gelmektedir. Bu nedenle özellik sayısı arttıkça alt küme arama işlemleri

zorlaşmaktadır. Komple arama yöntemi bütün alt kümeleri değerlendirdiği için en ideal sonucu bulmayı hedeflemektedir [89].

Ardışık arama yöntemi bütün alt kümeleri kullanmayıp onun yerine ileri yönlü arama (sequential forward selection), geri yönlü arama (sequential backward selection) ve iki yönlü arama (bi-directional selection) yöntemlerini kullanarak özellik seçimi yapar. Rastgele arama yöntemi veri yığını içerisinde rastgele seçmiş olduğu niteliklerle klasik ardışık arama yöntemine rasgelelik ilave ederek alt küme oluşturmaktadır [90].

Alt küme değerlendirme

Alt küme üretimi sonrasında meydana gelen her bir alt küme bağımsız ya da bağımlı ölçütlerle değerlendirilir. Değerlendirme aşamasında meydana gelen en yüksek performansa sahip özellik alt kümesi belirlenmeye çalışılmaktadır.

Durdurma ölçütleri

Özellik seçiminde seçilen özelliğin kalıp kalmamasını belirleyen bazı ölçütler vardır. Örneğin seçilmiş alt kümeye yeni eklenen özelliğin algoritma performansını artırıp arttırmaması, seçilecek özellik sayısının minimum değerinin belirlenmesi veya maksimum iterasyon sayısı gibi ölçütler özellik seçiminde ölçüt olarak kullanılmaktadır.

Sonuç doğrulama

Seçilen alt kümenin en doğru küme olduğunu anlamanın yolu algoritmanın çalıştırılarak sınıflandırma performansını ölçmektir. Kullanılan yöntemlerde sınıflandırma performans ölçütü olarak doğruluk oranına bakılmaktadır [91].

3.2.10. Özellik seçim yöntemleri

Özellik seçme işlemini istatistiksel temele bağlı filtreleme (filter) yöntemi, sarmal (wrapper) yöntem ve gömülü (embedded) yöntem olmak üzere üç ana sınıfta toplamak mümkündür [92]. Filtreleme yönteminde makine öğrenmesi algoritmaları işleme sokulmadan önce her bir niteliğe istatistiksel ölçüt ile puanlama yapılarak özellik seçimi yapılır. Sarmal

yöntemlerde özellik seçimi farklı birleşimler ve bunların birbirleri ile karıştırıldığı, değerlendirildiği ve öğrenme algoritmalarının araç olarak kullanılması ile yapılır. Gömülü yöntemlerde ise makine öğrenmesi algoritmaları ile özellik seçimi ölçütleri birlikte kullanılarak çalışmaktadır.

Filtreleme yöntemleri

Filtreleme yöntemleri yakınlık, bilgi, tutarlılık ve bağımlılık gibi istatistiksel yöntemleri kullanarak özellik seçme işlemini yapan en eski metodudur. Filtreleme yöntemi aracılığıyla her bir nitelik üzerinde istatistiksel ölçütler hesaplanarak (örneğin ortalama, standart sapma ve frekans değeri gibi) nitelikler puanlandırılır ve yeni bir alt özellik kümesi oluşturulur. Bu işlemler esnasında herhangi bir makine öğrenme algoritması kullanılmaz ya da sürece dahil edilmez. Fisher skor, t-skör, welch t-istatistiği ki-kare testi, bilgi kazancı ve kazanç oranı gibi teknikler yardımıyla özellik seçimi işlemleri gerçekleştirilebilir.

Sarmal yöntemler

Sarmal yöntemlerde farklı birleşimler değerlendirildiği sınıflandırma ve öğrenme algoritmalarının gösterdiği performans değerlerine göre en iyi tahmini yapan nitelikler alt küme olarak özellik seçim işlemini oluşturur [93]. Sarmal yöntemleri, filtreleme yöntemlerine göre daha iyi alt özellik kümesi oluştururken, filtreleme yöntemindeki hesaplama işlemlerinden daha çok işlem ve hesaplama maliyetine neden olmaktadır. Ardışık ileri yönde seçim boş küme ile başlayıp her döngüde yeni bir niteliği kümeye dahil ederek en iyi performansı gösteren alt özellik kümesini bulmaktadır [94, 95].

Ardışık geri yönde seçim değilse veri yığını içerisindeki bütün nitelikler ile işleme başlanır ve algoritma performansına göre nitelik sayısı azaltılarak en verimli alt küme seçilir bu esnada çıkarılan nitelik tekrar döngüye sokulmadan süreç devam eder [96]. L ekle - R çıkar yönteminde ardışık ileri yönde seçimde tek seferde seçilip sonrasında çıkarılmayan, ardışık geri yönde seçimdeki çıkarılan ve tekrar kümeye sokulamayan özellik değişkenlerinin sorununun çözülmüş hali olarak çalışmaktadır [97]. Sayılan bu yöntemlere benzeyen algoritma ile özellik seçim işlemi gerçekleştirilmektedir [98].

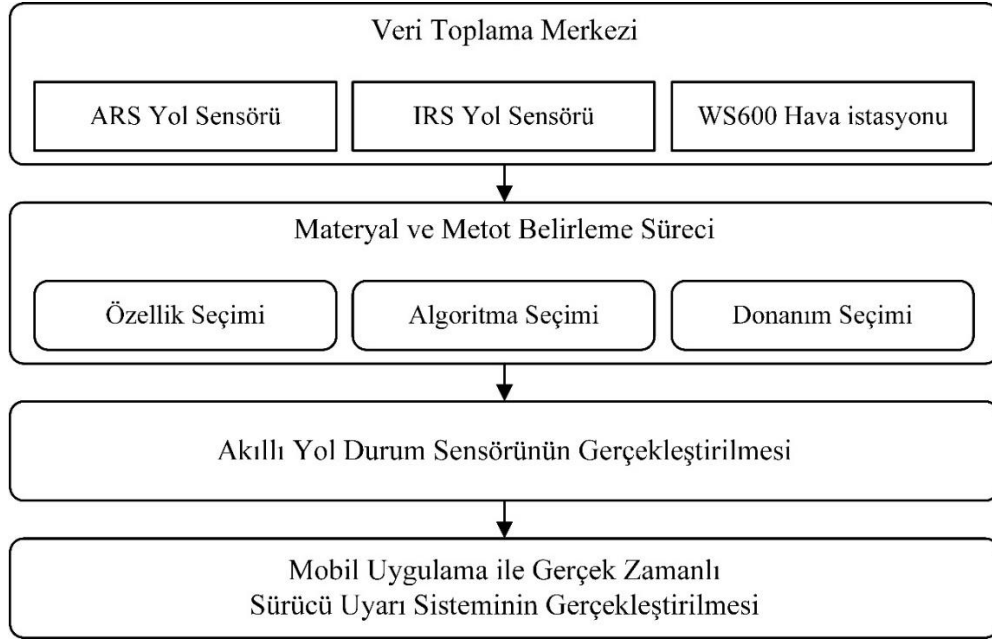
Gömülü yöntemler

Gömülü yöntemler sınıflandırma modeli oluştururken özellikler içerisinde modelin performansını etkileyen en iyi özellikleri hem sınıflandırma hem de özellik seçme algoritmaları yardımıyla seçmektedir. Gömülü yöntemlerin hesaplama ve işlem süreci sarmal yöntemlere göre daha da maliyetlidir. Karar ağaçları ve Destek Vektör Makineleri özyinelemeli özellik elemesi gibi bir çeşit geriye doğru alt küme seçimi yöntemleri kullanılmaktadır [99-101].

Özyinelemeli özellik seçimi önem derecesine göre özellikleri kendi arasında sıraya koyan bir yöntemdir. Sıralama işlemini de bazı ölçütlere göre yapmaktadır. Her döngüde özelliğin önem değeri ölçülür ve ilgisi az olan özellik listeden ya da kümeden çıkarılır. Özyinelemeye ihtiyaç duyulur çünkü elenen ya da listeden çıkarılan özellik diğer bir başka özellikle küme oluşturduğunda daha farklı sonuç verebilmektedir. Bu nedenle her bir özellik bütün alt kümeler ile ayrı ayrı değerlendirilmeli ancak bu sebeple oluşan yineleme hesaplama süresini de aşırı uzatmaktadır. Bu tez kapsamında alt özellik kümesi seçiminde gömülü yöntemler kullanılmıştır.

4. KABLOSUZ AKILLI YOL DURUM SENSÖRÜ VE GERÇEK ZAMANLI ARAÇ SÜRÜCÜ UYARI SİSTEMİ TASARIMI VE GERÇEKLEŞTİRİLMESİ

Bu tez çalışmasında araç sürücüleri için gerçek zamanlı zeki karar destek sistemi tasarlanmış ve gerçekleştirilmiştir. Çalışmanın süreci Şekil 4.1'deki blok diyagramda verilmektedir.

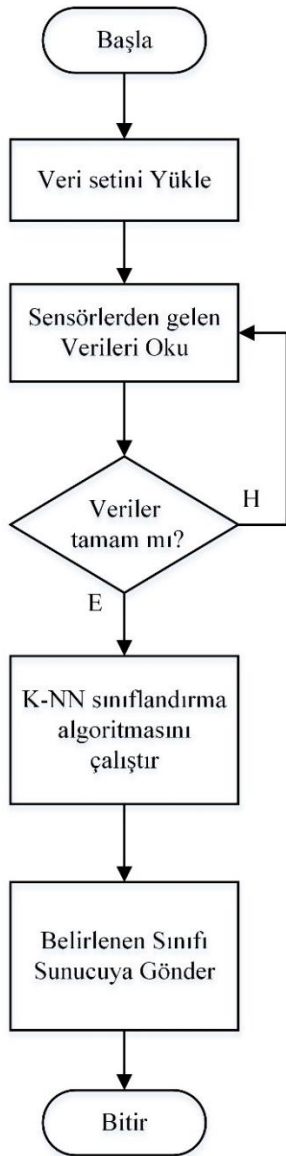


Şekil 4.1. Akıllı yol durum sensörü ve gerçek zamanlı araç sürücü uyarı sistemi

Birinci aşamada tasarım için gerekli veriler mevcut sistem üzerinden elde edilmiştir. Veriler Santez projesi kapsamında Ankara'ya kurulan buzlanma önleyici sistemde var olan ARS, IRS ve WS600 sensörlerinden alınmıştır.

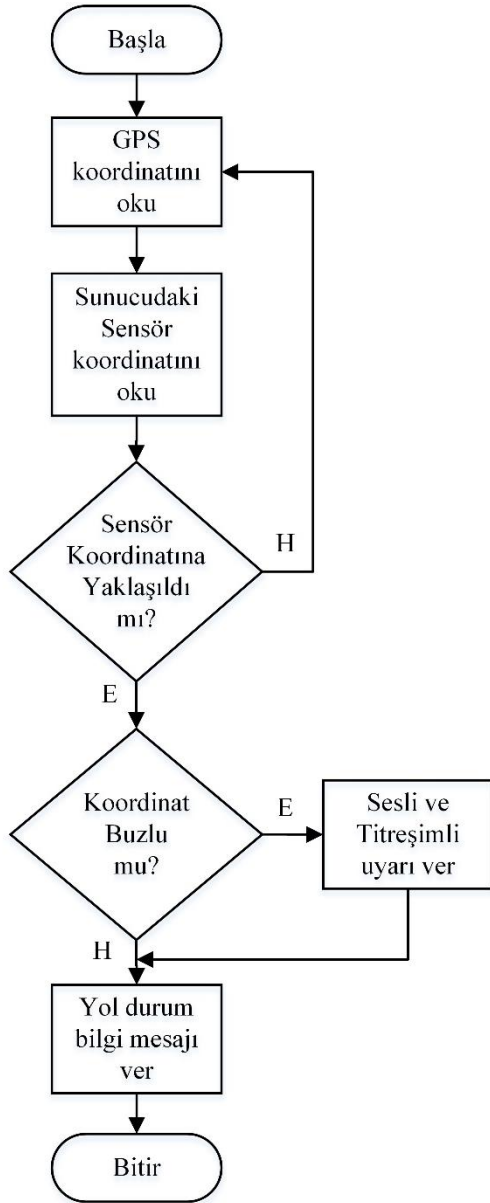
İkinci aşamada ise özellik, algoritma ve donanım seçimi işlemleri yapılmıştır. Özellik seçiminde toplanan özellikler arasında en verimli ve yüksek doğruluk değerlerini veren alt özellik kümesi seçilmiştir. Makine öğrenmesi metotlarında sınıflandırma işlemi yaparken bilgisayara fazla işlem yükü veren gereksiz özellik sayısıdır. Özellik sayısı arttıkça işlem süresi artar ve sonuca ulaşmak vakit almaktadır. Toplanan veri kümesindeki sınıflanma sonucunu en az etkileyen özelliği algoritma işleminden çıkartarak sistemin performansını arttırmaktadır. Bunun için kullanılan veri setinde en iyi performansı almak için özellik seçme işlemi yapılır. Performans ve cevap süresi göz önüne alınarak, seçilen özellik kümesinin hangi makine öğrenme metodu ile kullanılacağına karar verilmiştir. Özellik seçiminden

sonra ihtiyaç duyulan sensörler ve yöntemin koşturulacağı tek kartlı bilgisayar belirlenmiştir. Üçüncü aşamada ise en iyi sonucu veren özelliklerin algılanması için kullanılan sensörler ve tek kartlı bilgisayar donanımsal olarak entegre edilmiştir. Tek kartlı bilgisayar üzerine yüklenen algoritma yardımıyla yol durumunu sınıflandırılmıştır. Son aşama da ise yol kenarına konulan yol durum sensörü yol yüzeyini üzerinde bulunan algılayıcılar ve makine öğrenme algoritmasıyla sınıflandırarak internet üzerinden bir web sayfasına ve daha sonrasında akıllı cep telefonu uygulaması aracılığıyla sürücülere bildirmektedir. Gerçekleştirilen sistemin yol durum sensörü ve akıllı cep telefonu uyarı uygulamasının akış şeması Şekil 4.2’de verilmiştir.



Şekil 4.2. Akıllı yol durum sensörü akış şeması

Akıllı yol durum sensöründen alınan verilerle gerçekleştirilen uyarı sisteminin akış şeması Şekil 4.3’de verilmektedir.

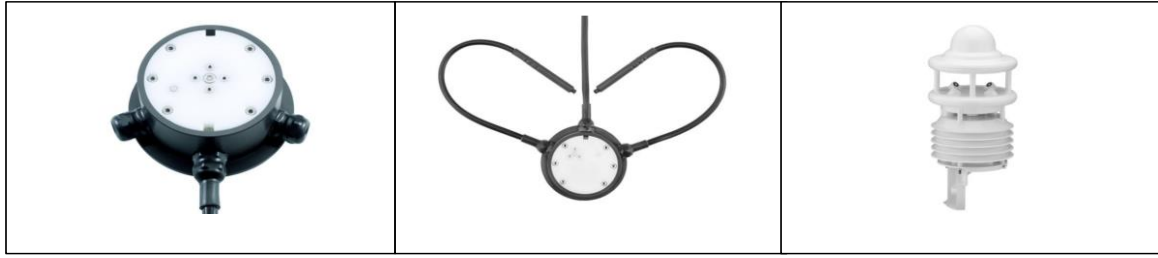


Şekil 4.3. Akıllı yol durum uyarı sisteminin akış şeması

4.1. Veri Toplama Sistemi

Yol durum sensörünün yol yüzeyini tahmin edebilmesi için gerekli bazı fiziksel verilere ihtiyaç duymaktadır. Yol yüzey durumunu rüzgar yönü, rüzgar hızı, hava sıcaklığı, havada bulunan nem ve hava basınç değerinin yanında yol sıcaklık değerleri ile yol yüzeyine serpilen buz çözücü kimyasallar etkilemektedir. Donma noktasını hava basıncı ve yol

yüzeyine serpilen kimyasallar etkilemektedir. Yol yüzeyinin buzlanması için su ya da nem olmazsa olmazlardan olup donma olayının olması için gereken önemli elementlerdendir. Tasarlanan sistemin yol yüzey tahmini yapabilmesi için riskli görülen bölgede yol ve meteorolojik veriler toplanmıştır. Veri toplama sisteminde, WS600 hava durumu istasyonunun yanında iki adet Luft marka ARS31 ile IRS31 model yol sensörü kullanılmıştır. Resim 4.1’de Aktif Ars31, Pasif IRS31 ve WS600 sensörleri gösterilmiştir.



Resim 4.1. Aktif ARS31, Pasif IRS31 ve WS600 sensörleri

Pasif çalışan IRS31 modeli iki parçalı gövde tasarımı ile asfaltın alt ve üst sıcaklığını, yol yüzeyinde bulunan sıvı miktarını, sensör üzerindeki sıvının donma noktasını ve sonuçta yol yüzey durumunu kuru, nemli, ıslak, buz veya kar ve tuzlu nemli olarak sınıflandıran yol sensörüdür. Toplanan veri seti Resim 4.2’de gösterilmektedir.

Hava_Sıcaklık	Çiğ_Noktası	Yüzde_Nem	Basınç_hpa	Rüzgar_Hızı	Rüzgar_Yönü	ARS_Yol_Sıcaklığı	IRS_Yol_Sıcaklığı	IRS_Sıvı_Kalınlığı	I_Sürtünme	I_Yol_Durumu
4.253	-2.909	59.670	923.456	1.918	295.284	10.430	8.286	0.000	0.820	Kuru
4.327	-2.830	59.713	923.407	1.157	61.758	10.430	8.283	0.000	0.820	Kuru
4.327	-2.830	59.713	923.407	1.456	346.692	10.430	8.260	0.000	0.820	Kuru
4.673	-2.507	59.700	923.375	1.684	255.653	10.430	8.252	0.000	0.820	Kuru
4.816	-2.491	59.175	923.384	2.498	282.309	10.350	8.247	0.000	0.820	Kuru
4.816	-2.491	59.175	923.384	2.105	335.233	10.350	8.255	0.000	0.820	Kuru
4.926	-2.743	57.639	923.409	2.141	281.556	10.350	8.278	0.000	0.820	Kuru
5.064	-2.381	58.641	923.351	1.765	267.212	10.350	8.283	0.000	0.820	Kuru
5.064	-2.381	58.641	923.351	3.002	280.104	10.260	8.283	0.000	0.820	Kuru
4.933	-2.561	58.392	923.348	3.119	85.460	10.260	8.255	0.000	0.820	Kuru
4.770	-2.696	58.471	923.287	1.600	151.235	10.260	8.286	0.000	0.820	Kuru
4.770	-2.696	58.471	923.287	1.316	44.736	10.180	8.344	0.000	0.820	Kuru
4.856	-2.522	58.877	923.363	2.025	228.138	10.180	8.406	0.000	0.820	Kuru
4.927	-2.478	58.780	923.337	0.000	0.000	10.180	8.521	0.000	0.820	Kuru
4.927	-2.478	58.780	923.337	1.901	281.429	10.090	8.583	0.000	0.820	Kuru
4.938	-2.537	58.477	923.336	4.269	36.337	10.090	8.557	0.000	0.820	Kuru
4.781	-2.666	58.556	923.310	0.000	0.000	10.090	8.505	0.000	0.820	Kuru
4.781	-2.666	58.556	923.310	1.258	267.084	10.090	8.477	0.000	0.820	Kuru
4.866	-2.602	58.490	923.298	3.211	258.175	10.000	8.427	0.000	0.820	Kuru
5.031	-2.394	58.718	923.299	3.086	293.456	10.000	8.357	0.000	0.820	Kuru
5.031	-2.394	58.718	923.299	2.495	322.878	10.000	8.315	0.000	0.820	Kuru
5.024	-2.426	58.605	923.272	1.198	245.522	9.920	8.247	0.000	0.820	Kuru
5.025	-2.548	58.076	923.267	1.725	267.871	9.920	8.199	0.000	0.820	Kuru
5.025	-2.548	58.076	923.267	1.448	266.491	9.920	8.205	0.000	0.820	Kuru
5.104	-2.444	58.203	923.138	1.349	278.091	9.840	8.202	0.000	0.820	Kuru

Resim 4.2. ARS31, IRS31 ve WS600 sensörlerinden toplanan veriler

Yol sensörleri yol yüzeyine gömülüp seri haberleşme ile verileri ana bilgisayara göndermektedir. Donma noktasını, sensör yüzeyini ısıtıp soğutarak hesaplayan ARS31 modeli aktif yol durum sensörüdür ve istenirse IRS31 modeli ile birlikte çalışmaktadır. İki farklı modelin kullanılma amacı yol yüzeyi verilerinin sağlamlasının yapılmasıdır. WS-600 hava istasyon sensörü bağıl nem, sıcaklık, basınç miktarı, rüzgâr yönü ve hızı, yağış şekli ve yoğunluğu gibi meteorolojik verileri ölçmek için kullanılmıştır.

4.2. Özellik Seçimi

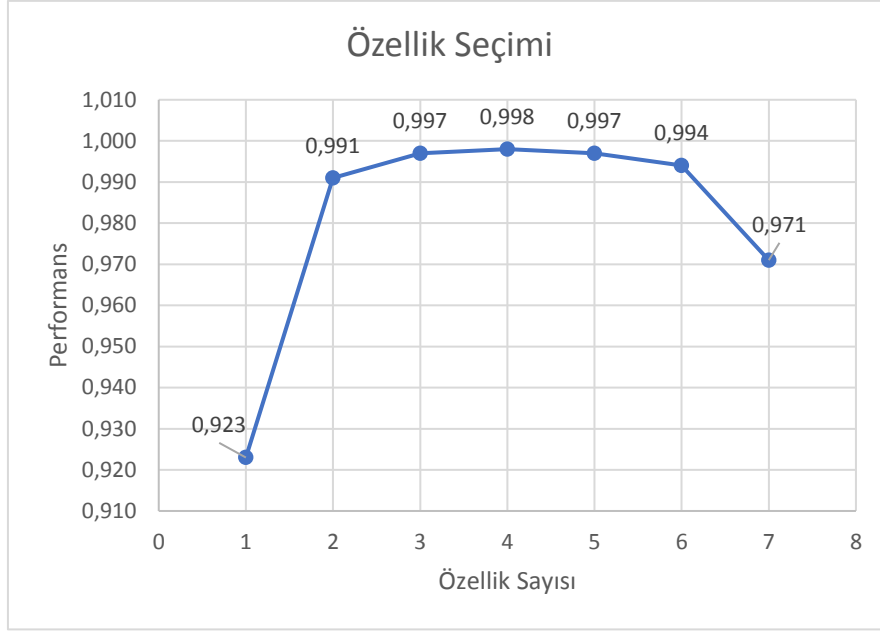
Veri toplama sisteminden çığ düşme noktası, hava basıncı, hava sıcaklığı, bağıl nem değeri, rüzgar yönü, rüzgar hızı, yol altı ve üstü sıcaklığı, yağış miktarı, yol yüzeyine serpilen buz çözücü kimyasal miktarı olmak üzere çeşitli değerler elde edilmiştir.

Makine öğrenmesi algoritmalarında bütün veri setindeki özelliklerin tamamı kullanılmayabilir. Bazı durumlarda veri setindeki bütün özelliklerin kullanılması tahmin başarımını düşürmektedir. Sınıflandırma işleminde performansı etkileyen etkenlerden biriside özellik seçme işlemidir.

Yol yüzey durumunu etkileyen nicelikler arasında rüzgar hızı, yönü ve yol sıcaklığı değerleri bulunmaktadır. Rüzgar hızı ile rüzgar yönünün buzlanmaya olan etkisinin, yol yüzeyinin güneşlenmeden kaynaklı sıcaklık etkisinden daha az etkili olduğu bilinmektedir [28]. Buna bağlı olarak yapılan özellik seçimi işleminde makine öğrenmesi algoritmasında kullanılmak üzere çığ düşme noktası, rüzgâr yönü ve rüzgâr hızı değerleri çıkarılmıştır.

Diğer yandan özellik kümesinin tamamı sınıflandırma algoritmasına girmeden önce, özellik alt küme seçme işlemine tabi tutulmuştur. Performansa dayalı yapılan özyinelemeli özellik seçimi işleminde nitelikler arasında önem derecesine göre özellikler kendi aralarında sıralanmış ve çığ düşme noktası, rüzgâr yönü ve rüzgâr hızı özellik alt kümesinden çıkarılmıştır.

Hava sıcaklığı, bağıl nem, hava basıncı, yol sıcaklığı özelliklerinin aynı anda kullanıldığında en yüksek performans değerini aldığı Şekil 4.4’de gösterilmiştir.



Şekil 4.4. Özellik sayısı ile performans değeri arasındaki ilişki

Özellik seçme işlemine tabi tutulan bu özellikler sistemin doğruluk oranını düşürdüğü için elenmiştir. Oluşan yeni alt özellik kümesi hava sıcaklığı, bağıl nem, hava basıncı, yol sıcaklığı olarak belirlenmiştir. Yol kenarına yerleştirilecek olan gömülü sistemde en yüksek performansı veren sensör ölçümleri arasında seçim yapılarak sınıflandırma algoritması işlemine geçilmiştir.

Yol durum sensörü, Raspberry Pi tek kartlı bilgisayarı üzerinde DHT22 sensörü ile hava sıcaklık ve nem değerini, DS18B20 sensörü ile yol sıcaklığını, BMP180 sensörü aracılığıyla hava basıncını ve iletkenlik sensörü ile de yol yüzeyine serpilmiş tuzu algılayacak biçimde oluşmaktadır. Tasarlanan sistemde yol yüzey sıcaklığını, hava sıcaklığını, algılanan nemi, hava basıncını ve yol yüzeyindeki tuz miktarını özellik olarak kullanmaktadır.

4.3. Sınıflandırma Algoritması Seçimi

Akıllı yol durum sensörü gerçek zamanlı topladığı asfalt sıcaklığını, hava sıcaklığını, nemi, basıncı ve yol yüzey iletkenlik verilerini kullanarak yol yüzeyini buzlu, kuru, nemli ve tuzlu nemli olarak sınıflandırmaktadır. Sınıflandırma algoritmasında veriler işlenmeden önce, özellik seçme aşamasında sınıflandırma algoritma başarısını en az etkileyen nitelikler elenmiştir. Günümüzde en çok kullanılan K-NN, Bayes, Karar Ağacı, Rastgele Orman, Lojistik Regresyon ve Destek Vektör Makine sınıflandırıcıları yardımıyla yol yüzey tahmini

yapılmıştır. Sınıflandırma algoritmalarının başarılı olup olmadığına karar verdiğimiz ölçütler bulunmaktadır. Bunlardan en sık kullanılanları Doğru Pozitif, Doğru Negatif, Doğruluk (accuracy), Dengeli Doğruluk (balanced accuracy), Kesinlik (precision), F-ölçütü (F-measure), Duyarlılık (sensitivity / recall), ve MCC ölçütüdür.

K-NN algoritması bu algoritmalar içinde kolay, eski ve performansı yüksek algoritmadır. Veri setinde sınıflandırma yaparken gürültülü veriden az etkilenmektedir. Çevreden toplanan gerçek zamanlı anlık verileri, daha önceden toplanmış olan verilerle arasındaki mesafeyi kıyaslayarak yol yüzey tahmini yapmaktadır. Mesafe ölçümlerinde Öklid, Minkowski, Manhattan ve Chebyshev gibi uzaklık ölçütleri kullanılmaktadır. Sınıflandırma algoritmalarının başarımlarını ölçütleri her bir sınıf için ayrı ayrı hesaplanarak ortalama başarımları elde edilmiştir. K-NN algoritmasının başarımları Çizelge 4.1’de gösterilmiştir.

Çizelge 4.1. K-NN algoritması başarımları

K-NN	DP	DN	Hassasiyet	Anma	F_Ölçütü	MCC	Doğruluk	Dengeli Doğruluk
Buz	0,9955	0,9987	0,9932	0,9955	0,9943	0,9941	0,9964	0,9925
Kuru	0,9975	0,9964	0,9988	0,9975	0,9982	0,9939		
Nemli	0,9612	0,9994	0,9519	0,9612	0,9565	0,9612		
Tuzlu	0,9924	0,9989	0,9849	0,9924	0,9887	0,9914		
Ortalama	0,9866	0,9983	0,9822	0,9866	0,9844	0,9852		

Bayes sınıflandırıcısı istatistiksel veriler aracılığıyla makine öğrenmesi problemlerinde kullanılan bir algoritmadır. İstatistiksel veriler ışığında özellikler ile sınıflar arasındaki ilişkiyi değerlendirerek tahmin yapan sınıflandırma algoritmasıdır. Çizelge 4.2’de Bayes algoritmasının başarımları gösterilmiştir.

Çizelge 4.2. Bayes algoritmasının başarımları

NB	DP	DN	Hassasiyet	Anma	F_Ölçütü	MCC	Doğruluk	Dengeli Doğruluk
Buz	0,7462	0,9841	0,9028	0,7462	0,8171	0,7519	0,7459	0,8530
Kuru	0,7382	0,8077	0,9225	0,7382	0,8201	0,5472		
Nemli	0,9612	0,9421	0,1781	0,9612	0,3005	0,9035		
Tuzlu	0,7913	0,8535	0,2758	0,7913	0,4090	0,6460		
Ortalama	0,8092	0,8969	0,5698	0,8092	0,6687	0,7121		

Çizelge 4.2’de görüldüğü gibi bayes sınıflandırıcısının doğruluk değeri 0,7459, dengeli doğruluk değeri 0,8530 olarak hesaplanmıştır. MCC ve diğer ölçütlerin de başarımlarının düşük olması sebebiyle gömülü sistem üzerinde bu algoritma sınıflandırıcı olarak tercih edilmemiştir.

Karar ağacı algoritmaları ağaç görünümünde olup, kökten uca doğru karar düğümleri ile ilerleyerek yaprak kısmına yani sınıf değerine ulaşarak çalışan bir sınıflandırma algoritmasıdır. Kategorik ve nümerik veri setleriyle çalışması çok kolay olan bir sınıflandırıcı olduğu için sıklıkla kullanılmaktadır. Karar Ağacı algoritmasının başarımlar değerleri Çizelge 4.3’de gösterilmiştir.

Çizelge 4.3. Karar ağacı algoritmasının başarımlar değerleri

KA	DP	DN	Hassasiyet	Anma	F_Ölçütü	MCC	Doğruluk	Dengeli Doğruluk
Buz	0,9992	0,9993	0,9962	0,9992	0,9977	0,9985	0,9948	0,9874
Kuru	0,9965	0,9918	0,9974	0,9965	0,9969	0,9883		
Nemli	0,9417	0,9985	0,8899	0,9417	0,9151	0,9417		
Tuzlu	0,9734	0,9988	0,9828	0,9734	0,9781	0,9725		
Ortalama	0,9777	0,9971	0,9666	0,9777	0,9721	0,9753		

Rastgele orman sınıflandırıcısı, aynı karar ağaçları gibi kökten dallara doğru oluşturulan birden fazla ağacın oylama yöntemi ile sınıflandırma işleminin yapıldığı algoritmadır. Oluşan bu karar ağaçlarının fazlalığından ötürü başarımları yüksek ancak gerçek zamanlı çalışmalarda ani değişimler için cevap süresi diğer algoritmalara göre yavaş sonuç vermektedir. Rastgele orman algoritmasının başarımlar değerleri Çizelge 4.4’de gösterilmiştir.

Çizelge 4.4. Rastgele orman algoritmasının başarımlar değerleri

RO	DP	DN	Hassasiyet	Anma	F_Ölçütü	MCC	Doğruluk	Dengeli Doğruluk
Buz	1,0000	0,9991	0,9955	1,0000	0,9977	0,9991	0,9967	0,9892
Kuru	0,9990	0,9938	0,9980	0,9990	0,9985	0,9929		
Nemli	0,9515	0,9999	0,9899	0,9515	0,9703	0,9524		
Tuzlu	0,9715	0,9991	0,9865	0,9715	0,9790	0,9710		
Ortalama	0,9805	0,9980	0,9925	0,9805	0,9865	0,9788		

Lojistik regresyon sınıflandırıcısı hava sıcaklığı, hava nemi, hava basıncı gibi bağımsız niteliklerin buzlu veya kuru gibi bağımlı niteliklerle olan ilişkisini kurarak sınıflandırma

yapan algoritmadır. Çalışma prensibi olarak sınıf değerinin ne olduğunu değil, sınıf değerinin ne olacağına dair olasılık tahmini yapmayı amaçlamaktadır. İkili, sıralı ve çok kategorili lojistik regresyon olmak üzere üç gruba ayrılırken çalışmada kuru, nemli, buzlu ve tuzlu nemli olmak üzere dört sınıf olmasından dolayı çok kategorili regresyon kullanılmıştır. Algoritmanın doğruluk değeri 0,9129 iken dengeli doğruluk değeri 0,7047 olarak hesaplanmıştır. Öte yandan nemli ve tuzlu sınıfların doğru pozitif oranlarının çok düşük olduğu görülmektedir. Bunun anlamı sınıflandırıcı tuzlu ve nemli sınıf tahminlerinde başarısız olmuş ve ortalama tahmin başarımını düşürmüştür. Lojistik regresyon algoritmasının başarımlar değerleri Çizelge 4.5’de gösterilmiştir.

Çizelge 4.5. Lojistik regresyon algoritmasının başarımlar değerleri

LR	DP	DN	Hassasiyet	Anma	F_Ölçütü	MCC	Doğruluk	Dengeli Doğruluk
Buz	0,9841	0,9922	0,9615	0,9841	0,9727	0,9763	0,9129	0,7047
Kuru	0,9924	0,6692	0,9030	0,9924	0,9456	0,6991		
Nemli	0,0000	1,0000	0,0000	0,0000	0,0000	0,0000		
Tuzlu	0,0000	1,0000	0,0000	0,0000	0,0000	0,0000		
Ortalama	0,4941	0,9154	0,4661	0,4941	0,4797	0,4189		

Destek Vektör makinası, regresyon ve sınıflandırma işlemleri için verilerin bulunduğu uzayda sınıfları bulmak için sonsuz boyutlu düzlemler oluşturmaktadır. DVM nem, basınç, sıcaklık ve tuzluluk gibi çok boyutlu özellik alanlarının sınırlarını kolay ve hızlı şekilde yüksek doğrulukla ayırabilmektedir.

DVM, doğrusal ve doğrusal olmayan yöntemle sınıflar arasını ayırabilme özelliği bulunan çekirdek yapıya sahiptir. Yapılan çalışmada doğruluk değeri 0,9050 iken dengeli doğruluk değeri 0,6948 olarak hesaplanmıştır. Destek vektör makinası algoritmasının başarımlar değerleri Çizelge 4.6’ da gösterilmiştir.

Çizelge 4.6. Destek vektör makinası algoritmasının başarımlar değerleri

DVM	DP	DN	Hassasiyet	Anma	F_Ölçütü	MCC	Doğruluk	Dengeli Doğruluk
Buz	0,9023	0,9961	0,9786	0,9023	0,9389	0,9024	0,9050	0,6948
Kuru	0,9964	0,6236	0,8915	0,9964	0,9410	0,6681		
Nemli	0,0000	1,0000	0,0000	0,0000	0,0000	0,0000		
Tuzlu	0,0398	1,0000	1,0000	0,0398	0,0766	0,1426		
Ortalama	0,4846	0,9049	0,7175	0,4846	0,5785	0,4283		

Sensörlerden alınan ve özellik seçimi gerçekleştirilen veri setindeki sınıf değerlerine bakıldığında eşit olmayan bir dağılım görülmektedir. Veri setinin sınıflara dağılımı eşit olduğunda doğruluk değeri ile dengeli doğruluk değeri benzer sonuç verirken, veri setinin eşit olmayan dağılımında dengeli doğruluk değeri doğruluk değerinden daha düşük performans göstermektedir. Sınıflandırma algoritmalarının sonuçları incelendiğinde doğruluk ölçütünün sonucuna bakıldığında en yüksek başarıyı rastgele orman sınıflandırıcısı vermektedir.

Eşit dağılmış bir veri setiyle çalışma durumunda, rastgele orman sınıflandırıcısının doğruluk değerinin yüksek olması nedeniyle tercih edilebilir. Nitekim kullandığımız veri setinin sınıflara dağılımı eşit olmamakla beraber aralarında çok büyük sayısal oranda fark bulunmaktadır. Bu nedenle doğruluk değerine bakmak yerine dengeli doğruluk ölçütüne bakıldığında başarımları en yüksek olan sınıflandırıcı K-NN olmaktadır. Tek kartlı bilgisayar olan Raspberry Pi kullanılarak tasarlanan sistem, yol yüzey sıcaklığını, hava sıcaklığını, algılanan nemi, hava basıncını ve yol yüzeyindeki tuz miktarını okumaktadır.

Sensörlerden okunan verileri tek kartlı bilgisayarda K-NN sınıflandırma algoritmasıyla işleme aldığımızda, % 99,64 doğruluk değeri, % 99,25 dengeli doğruluk ve %98,52 MCC değerleri elde edilmiş ve yol yüzey tahmini için K-NN sınıflandırıcısı seçilmiştir. Diğer bir yandan kullanılan algoritmaların bilgisayar üzerinde çalıştırılarak algoritmaların sonuç vermesi için geçen süreye bakıldığında rastgele orman sınıflandırıcısının K-NN sınıflandırıcısından daha yavaş sonuç verdiği görülmektedir. Hem yüksek doğrulukta sonuç verip hem de daha hızlı cevap veren algoritma olan K-NN algoritması tasarlanan sistem için tercih edilmiştir. Çizelge 4.7’de sınıflandırma algoritmalarının vermiş olduğu başarımları gösterilmiştir.

Çizelge 4.7. Sınıflandırma algoritmalarının başarımları

Sınıflandırıcı	K-NN	RO	KA	DVM	LR	NB
Dengeli Doğruluk	0,9925	0,9892	0,9874	0,6948	0,7047	0,8530
Doğruluk	0,9964	0,9967	0,9948	0,9050	0,9129	0,7459
MCC	0,9852	0,9788	0,9753	0,4283	0,4189	0,7121

Çizelge 4.8’de sınıflandırma algoritmalarının bilgisayar üzerinde verdiği cevap süresi verilmiştir.

Çizelge 4.8. Sınıflandırma algoritmalarının cevap süresi

Sınıflandırıcı	K-NN	RO	KA	DVM	LR	NB
Cevap Süresi (ms)	89	219	83	1200	103	50

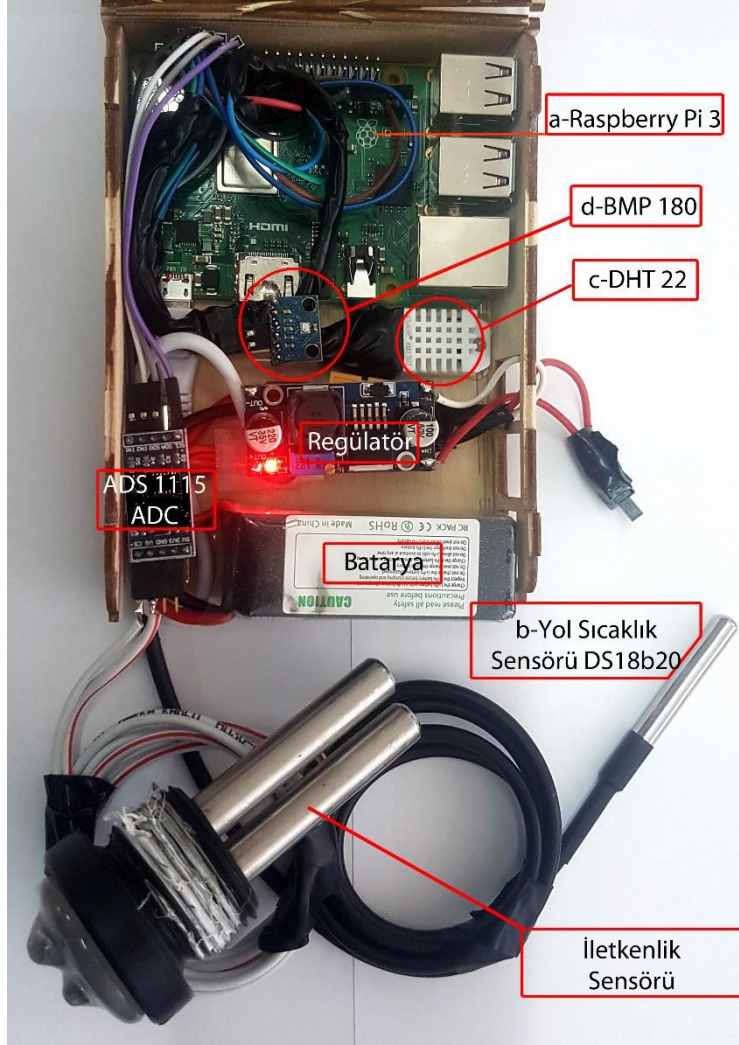
Riskli konum üzerinden alınan anlık verilerin algoritmaya sokulup sonucun hızlı bir şekilde elde edilmesi ani gizli buzlanma gibi durumlarda sürücüler için önem arz etmektedir.

Kullanılan sensörlerin okunması, algoritmanın çalıştırılması ve sonucun sürücülerin akıllı cep telefonuna gönderilmesi için geçen süre zarfında her bir saniyenin önemi bulunmaktadır. Bu nedenle en hızlı ve en yüksek doğrulukla çalışan sınıflandırma algoritması olan K-NN seçilmiştir.

4.4. Akıllı Yol Durum Tespit Gömülü Sisteminin Gerçekleştirilmesi

Akıllı yol durum sensörü tek kartlı bilgisayar olan Raspberry Pi, hava basınç sensörü BMP180, nem ve sıcaklık sensörü DHT22, yol sıcaklık sensörü DS18B20 ve iletkenlik sensöründen meydana gelmektedir. Yol yüzey durum tahmini yapılabilmesi için geçmiş yol ve hava durumu verilerinin toplanması ve tahmin algoritmasından yüksek verim alınabilmesi için elde edilen verilerin özellik seçim işlemine tabi tutulması gerekmektedir.

Raspberry Pi gömülü sistemi üzerine entegre edilen BMP180, DHT22, DS18B20 ve iletkenlik sensörleri aracılığıyla riskli konumdan gerçek zamanlı olarak gerekli verileri sisteme kaydedilmektedir. Daha sonra alınan anlık veriler ile geçmiş veriler daha önceden seçilmiş olan K-NN sınıflandırma algoritmasına ile işleme alınarak yol yüzey sınıflandırması yapılmıştır [102]. Gerçekleştirilen sistem Resim 4.3’de gösterilmektedir.



Resim 4.3. Gerçekleştirilen sistemin bütünleşmiş hali

Raspberry Pi gömülü sistemi üzerine Python dili aracılığıyla program yazılmaktadır. Raspberry Pi üzerinde bulunan HDMI çıkış aracılığıyla ekran bağlantısı yapılarak istenirse kod yazımı yapılabilir istenirse çalıştırılan kod çıktısı görülebilmektedir. Yol durum sensörü uygulamasında gerçekleştirilen sensör üzerinde ekran takılmaması nedeniyle çalışan yazılımın çıktısını Raspberry Pi bilgisayarına SSH (Secure Shell) protokolü ile aynı ağ üzerinden uzak bağlantı ile gösterilebilir. Çalışan programın Python çıktısı Resim 4.4’de gösterilmektedir.

```

pi@raspberrypi ~ $ sudo python tek.py
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[-8.0, 71.0, 925.0, -5.0]
> predicted='Buz'
Train set: 15998
[8.0, 80.0, 900.0, 5.0]
> predicted='Kuru'
Train set: 15998
[8.0, 80.0, 900.0, 5.0]
> predicted='Kuru'

```

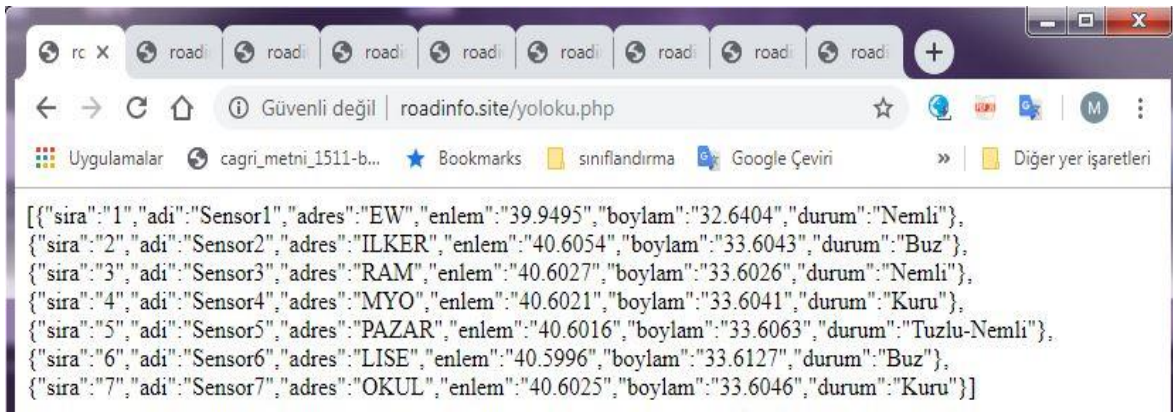
Resim 4.4. Gerçekleştirilen programın Python çıktısı

Resim 4.4’de görüldüğü gibi çevresel sensörlerden anlık olarak alınan hava sıcaklık, nem, basınç ve yol sıcaklık değerleri sınıflandırma algoritmasında işleme alınıp geçmiş veri setleri ile arasındaki mesafeye bakılarak K-NN algoritması aracılığıyla tahmin yapılmıştır. Yapılan tahmin sınıf değerine Raspberry Pi üzerine yazılan program, alt program ve kullanılan kütüphaneler aracılığıyla tahmin sonucunu web sunucusuna gönderilerek son kullanıcı olan sürücülerin erişimi sağlanmıştır. Resim 4.5’de tahmini yapılmış ve web sunucusuna gönderilmiş örnek yol durum sensör çıktılarının web sayfası gösterilmektedir.



Resim 4.5. Web sunucusuna gönderilen yol durum bilgisi

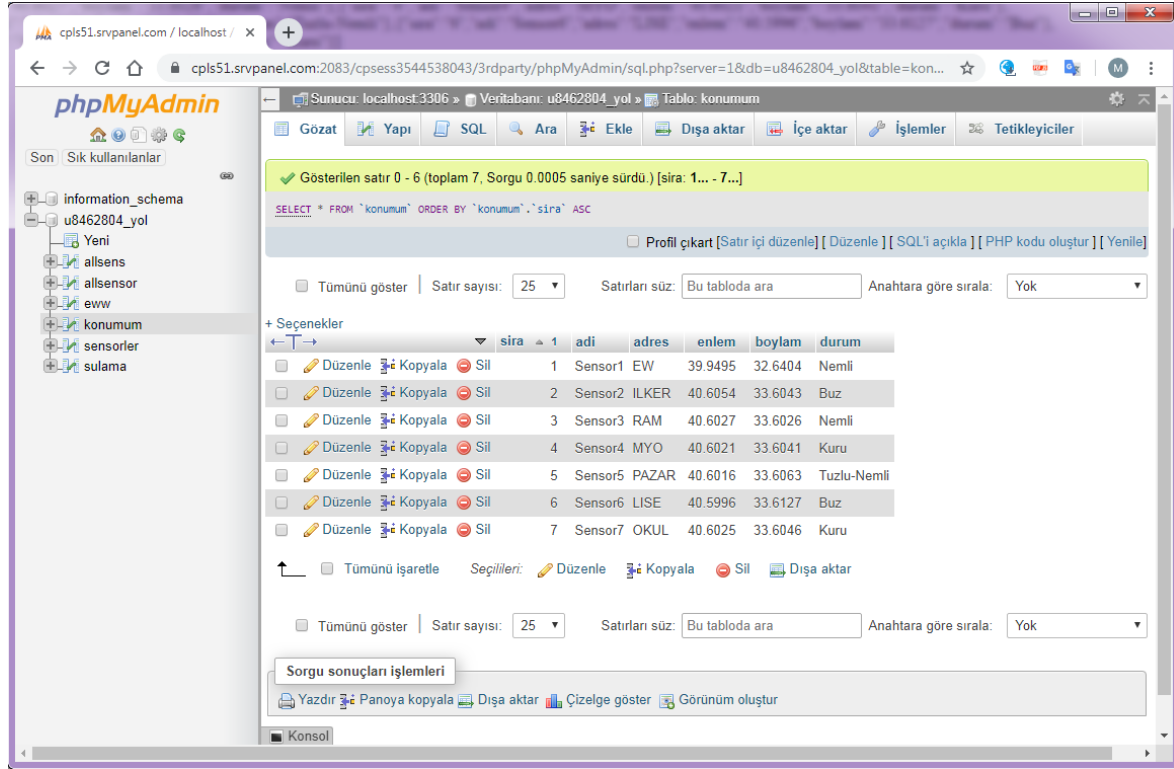
Tasarlanan yol sensörü içerisine yazılan programda hangi düğüm olduğu belirlenerek istenilen sayı kadar sensör çıktısı web sunucusuna gönderilerek kullanılabilir. Web sayfasında sensörün sıra numarası, adı, adresi, coğrafi konumu ve yol durum bilgisi gibi veriler görüntülenmektedir. Gerçekleştirilen ve sonucunda tahmin yapan her bir yol durum sensörü bulunduğu konumun yol durum sonucunu Python'da Requests modülünü kullanarak http isteği ile parametre cinsinden sonuç gönderilmektedir. Yapılan bu istem ile web sayfası üzerinden veri tabanına giriş yapılmakta ve diğer bir web sayfasında ise bu girişler Resim 4.6' daki gibi listelenmektedir.



Resim 4.6. Web sayfasında listelenen yol durum bilgisi

Raspberry Pi üzerinden internet aracılığıyla web sayfası isteminde bulunularak esasında web sayfası adresi üzerinde durum kaydedicisine sensör tahmini get parametresi ile gönderilir. Çağırılan web sayfası http link adresi üzerinde bulunan durum bilgisini mysql veri tabanına

yükleyerek istenildiği zaman başka web adresleri üzerinde listelenmesini sağlamaktadır. Resim 4.7’de Mysql veri tabanı sayfası ve konumlara göre yol durumları gösterilmiştir.



Resim 4.7. Veri tabanı sayfası ve yol durum tahmini

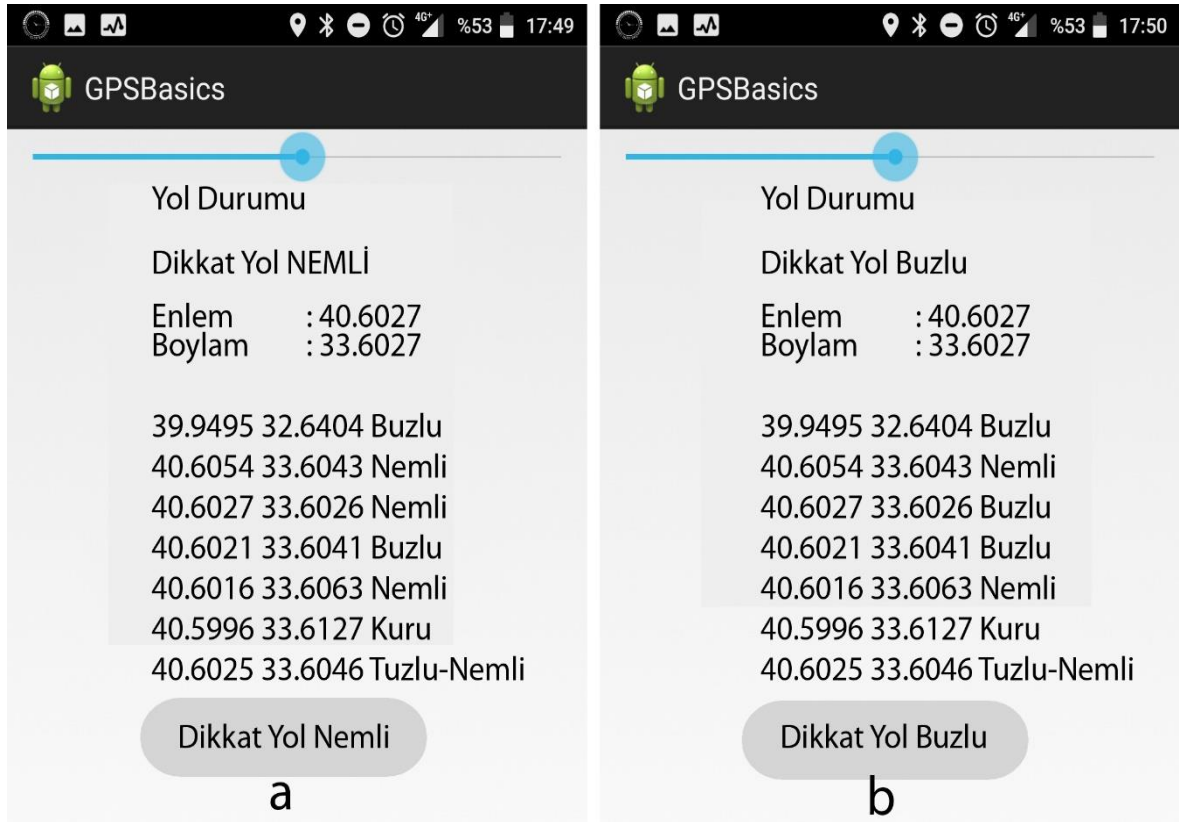
Web sayfası üzerinde yol durum sensörlerinin yapmış oldukları tahminleri görmek mümkün olduğu gibi yapılan akıllı cep telefonu uygulaması aracılığıyla da bu bilgilere kolay yoldan erişilmektedir.

4.5. Mobil Uygulama Gerçekleştirilmesi

Araç sürücülerini yol durum bilgisini yol kenarına yerleştirilen trafik işaret ve işaretçileri aracılığıyla erişmektedir. Bu işaretçiler kimi zaman sürücülerin dikkatinden kaçmakta ya da kötü hava koşulları nedeniyle fark edilememektedir. Sürücülerin yol durumunu fark edebilmesi için günlük yaşamın hemen hemen her anında yanlarında bulunan cep telefonlarından faydalanılmıştır.

Akıllı cep telefonuna yazılan uygulama sayesinde riskli bölgeye yaklaşan sürücüler yol sensöründen alınan bilgi ile uyarılarak kötü yol koşuluna tedbir alabilmeleri sağlanmıştır.

Yol durum sensörü yapmış olduğu sınıflandırmayı internet bağlantısı aracılığıyla web sunucusuna göndermektedir. Yazılan akıllı cep telefonu uygulaması üzerinde bulunan GPS sensöründen bulunduğu koordinatı her konum değiştiğinde güncellemektedir. Cep telefonu konum bilgisi her güncellendiğinde yol durum sensörlerinin tahmin sonuçlarını gönderdikleri web sunucusunda bulunan konum bilgileri ile karşılaştırılarak eşleşme olması durumunda yol bilgisi cep telefonu ekranına sesli uyarı olarak gelmektedir. Yazılan cep telefonu uygulaması Resim 4.8’de gösterilmiştir.



Resim 4.8. Android cep telefonu uygulaması ekran görüntüsü

Uygulama üzerinde bulunan hassasiyet ayar düğmesi ile riskli konuma varmadan önce ne kadar uzaklıkta uyarı vereceği ayarlanabilmektedir. Tasarlanan akıllı cep telefonu uygulaması, yol kenarına yerleştirilen trafik işaretlerine nazaran sürücüler için yol koşulları için uyarda daha etkili olmuştur.

5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında sürücülerin kış koşullarında güvenli seyahat edebilmesi için; hızlı cevap süresine sahip, düşük maliyetli, güvenilir, yüksek tahmin başarımlarına sahip, etkin bir uyarı sistemi içeren yol durum tespit ve uyarı sistemi gerçekleştirilmiştir. Çalışmada, sınıflandırma başarımları daha yüksek olduğu için temaslı yol yüzey tahmin yaklaşımı kullanılmıştır.

Sınıflandırma işleminde performansı etkileyen etkenlerden biriside özellik seçme işlemidir. Tasarlanan sistemde yol yüzey sıcaklığını, hava sıcaklığını, algılanan nemi, hava basıncını ve yol yüzeyindeki tuz miktarını özellik olarak kullanmaktadır. Bu özelliklerin yanı sıra sistemimize kaydedilen çığ düşme noktası, rüzgâr yönü ve rüzgâr hızı bilgileri de bulunmaktadır. Özellik seçme işlemine tabi tutulan bu özellikler sistemin doğruluk oranını düşürdüğü için dikkate alınmamıştır.

Akıllı yol durum sensörünün hızlı cevap süresi ve yüksek başarımla çalışması için özellik kümesi hava sıcaklığı, bağıl nem, hava basıncı, yol sıcaklığı olarak kullanılmıştır. Belirlenen bu özellikler ile yol durumunu algılamak için düşük maliyetli algılayıcılar ve tek kartlı bilgisayar olan Raspberry Pi kullanılarak temaslı akıllı yol durum sensörü geliştirilmiştir.

Yol durumunun tahmini için K-NN, Bayes, Karar Ağacı, Rastgele Orman, Lojistik Regresyon ve Destek Vektör Makine sınıflandırıcıları ile uygulamalar yapılmıştır. K-NN sınıflandırma algoritmasının başarımlar ölçütlerine göre, Doğruluk oranı %99,64 gibi yüksek bir değer bulunmuştur. Ancak toplanan veriler içerisinde daha az sayıda bulunan buzlu veri sayısı, riskli bölgeden toplanan kuru veri sayısının yanında oldukça düşük bir oran olduğundan, elde edilen verilerde sınıflar arası dağılımda dengesizlik bulunmaktadır. Bundan dolayı diğer başarımlar ölçütleri de hesaplanmıştır.

Kesinlik ve anma ölçütlerinin harmonik ortalamasını veren F-Ölçütünün sonucuna bakıldığında, 0,98 gibi yüksek bir oran çıktığı görülmektedir. Yapılan çalışmalarda doğruluk ölçütü değerinin yüksek olduğu sınıflandırma modelinin performansı daha iyi gözükürken, aslında sınıflar arası eşit veri dağılımı olan doğruluk değeri sınıflar arasında dengesiz veri dağılımına göre daha yüksek performans gösterir. Veri setinin sınıflara dağılımı eşit olduğunda doğruluk değeri ile dengeli doğruluk değeri benzer sonuç verirken, veri setinin eşit olmayan dağılımında dengeli doğruluk değeri doğruluk değerinden daha az performans göstermektedir.

Sınıflandırma işlemlerinin doğruluk oranına bakıldığında en yüksek doğruluk oranını rastgele orman algoritması vermektedir. İkinci en yüksek doğruluk oranını ise K-NN algoritması vermektedir. Bunun yanı sıra bilgisayar üzerinde çalıştırılarak algoritmaların sonuç vermesi için geçen süreye bakıldığında rastgele orman sınıflandırıcısının daha yavaş sonuç verdiği görülmektedir. Hem yüksek doğrulukta sonuç verip hem de daha hızlı cevap veren algoritma olan K-NN algoritması tasarlanan sistem için tercih edilmiştir.

Riskli konum üzerinden alınan anlık verilerin algoritmaya sokulup sonucun hızlı bir şekilde elde edilmesi ani gizli buzlanma gibi durumlarda sürücüler için önem arz etmektedir. Kullanılan sensörlerin okunması, algoritmanın çalıştırılması ve sonucun sürücülerin akıllı cep telefonuna gönderilmesi için geçen süre zarfında her bir saniyenin önemi bulunmaktadır. Bu nedenle en hızlı ve en yüksek doğrulukla çalışan algoritma seçilmiştir. Çalışmadaki veri setinde sınıflar arasındaki dağılım eşit olmadığından sistemin başarımında Dengeli Doğruluk sonuçlarına bakıldığında en iyi sonucu K-NN sınıflandırıcısı vermiştir. Bu nedenle de tasarlanan sistemde K-NN algoritmasıyla %99,25 başarımla sınıflandırma yapılmıştır.

Çalışmada ayrıca geliştirilen mobil uygulama ile araç sürücülerini riskli konuma varmadan 200 metre önce sesli, görüntülü ve titreşim mesajları ile uyarılmaktadır. Böylece etkin ve gerçek zamanlı bir uyarı sistemi oluşturulmuştur.

Akıllı yol durum sensörü alüminyum kasa içerisine yerleştirilerek, güneş paneli desteği ile enerjisi temin edilerek yol yüzeyine montajı yapıldığında daha güvenli tahmin yapılması sağlanabilir. Gelecek çalışmalarda tehlikeli yol ve hava koşuluna sahip koordinatlara yaklaşan araç sürücülerine yol yüzey durumuna yönelik araç hızını ayarlaması için karar destek sistemi geliştirilebilir.

KAYNAKLAR

1. Andersson, A. K., and Chapman, L. (2011). The impact of climate change on winter road maintenance and traffic accidents in West Midlands, UK. *Accident Analysis & Prevention*, 43(1), 284-289.
2. Lazarev, Y., Medres, C., Raty, J., and Bondarenko, A. (2017). Method of assessment and prediction of temperature conditions of roadway surfacing as a factor of the road safety. *Transportation Research Procedia*, 20, 393-400.
3. Crevier, L. P., and Delage, Y. (2001). A new model for road-condition forecasting in Canada. *Journal of Applied Meteorology*, 40(11), 2026-2037.
4. Zhao, X., and Rose, J. L. (2016). Ultrasonic guided wave tomography for ice detection. *Ultrasonics*, 67, 212-219.
5. Fuleki, D., Sun, Z., Wu, J., and Miller, G. (2017, 5-9 June). *Development of a Non-Intrusive Ultrasound Ice Accretion Sensor to Detect and Quantify Ice Accretion Severity*. 9th AIAA Atmospheric and Space Environments Conference, Denver, Colorado, 42-47.
6. Casselgren, J., Kutila, M., and Jokela, M. (2012). *Slippery road detection by using different methods of polarised light*. Advanced Microsystems for Automotive Applications, 207-220.
7. Gresham, I., Jain, N., Budka, T., Alexanian, A., Kinayman, N., Ziegner, B., Brown, S., and Staecker, P. (2001). A compact manufacturable 76-77GHz radar module for commercial ACC applications. *IEEE Transactions on Microwave Theory Techniques*, 49(1), 44-58.
8. Jonsson, P. (2011, 28-31 October). *Remote sensor for winter road surface status detection*. IEEE Sensors Conference, Limerick, Ireland, 1285-1288.
9. Kuehnle, A., and Burghout, W. (1998). Winter road condition recognition using video image classification. *Transportation Research Record*, 1627(1), 29-33.
10. Omer, R., and Fu, L. (2010, 19-22 September). *An automatic image recognition system for winter road surface condition classification*. 13th international IEEE conference on intelligent transportation systems, Funchal, Portugal, 1375-1379.
11. Sass, B. H. (1997). A numerical forecasting system for the prediction of slippery roads. *Journal of Applied Meteorology*, 36(6), 801-817.
12. Gagnon, R., Groves, J., and Pearson, W. (2012). Remote ice detection equipment. *Cold Regions Science Technology*, 72, 7-16.
13. Casselgren, J., Rosendahl, S., Sjö Dahl, M., and Jonsson, P. (2016). Road condition analysis using NIR illumination and compensating for surrounding light. *Optics and Lasers in Engineering*, 77, 175-182.

14. Alonso, J., López, J., Pavón, I., Recuero, M., Asensio, C., Arcas, G., and Bravo, A. (2014). On-board wet road surface identification using tyre/road noise and support vector machines. *Applied acoustics*, 76, 407-415.
15. Shao, J., and Lister, P. (1996). An automated nowcasting model of road surface temperature and state for winter road maintenance. *Journal of Applied Meteorology*, 35(8), 1352-1361.
16. Tabatabai, H., and Aljuboori, M. (2017). A novel concrete-based sensor for detection of ice and water on roads and bridges. *Sensors*, 17(12), 2912.
17. Troiano, A., Pasero, E., and Mesin, L. (2010). New system for detecting road ice formation. *IEEE Transactions on Instrumentation and Measurement*, 60(3), 1091-1101.
18. Jackson, D. G., Liao, J. Y., and Severson, J. A. (2003, 16-20 June). *An Assessment of Goodrich Ice Detector Performance in Various Icing Conditions*. SAE Technical Paper, Chicago, Illinois, 15-21.
19. Roy, S., Izad, A., DeAnna, R. G., and Mehregany, M. (1998). Smart ice detection systems based on resonant piezoelectric transducers. *Sensors Actuators A: Physical*, 69(3), 243-250.
20. Li, X., Shih, W. Y., Vartuli, J., Milius, D. L., Prud'homme, R., Aksay, I. A., and Shih, W. H. (2002). Detection of water-ice transition using a lead zirconate titanate/brass transducer. *Journal of Applied Physics*, 92(1), 106-111.
21. Gui, K., Ge, J., Ye, L., and Huang, L. (2019). The piezoelectric road status sensor using the frequency scanning method and machine-learning algorithms. *Sensors Actuators A: Physical*, 287, 8-20.
22. Jonsson, P., Casselgren, J., and Thörnberg, B. (2014). Road surface status classification using spectral analysis of NIR camera images. *IEEE Sensors Journal*, 15(3), 1641-1656.
23. Li, J. (2010, 6-7 March). *A combination of DE and SVM with feature selection for road icing forecast*. 2nd International Asia Conference on Informatics in Control, Automation and Robotics, Wuhan, China, 509-512.
24. Cheng, L., Zhang, X., and Shen, J. (2019). Road Surface Condition Classification using Deep Learning. *Journal of Visual Communication Image Representation*, 2.
25. Baby, C., and George, B. (2012, 16-18 November). *A capacitive ice layer detection system suitable for autonomous inspection of runways using an ROV*. IEEE International Symposium on Robotic and Sensors Environments Proceedings, Magdeburg, Germany, 127-132.
26. Shao, J. (1998). Application of an artificial neural network to improve short-term road ice forecasts. *Expert Systems with Applications*, 14(4), 471-482.
27. Luchetta, A., Manetti, S., and Francini, F. (1998). Forecast: A neural system for diagnosis and control of highway surfaces. *IEEE Intelligent Systems and Their Applications*, 13(3), 20-26.

28. Bogren, J., Gustavsson, T., and Postgård, U. (2000). Local temperature differences in relation to weather parameters. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 20(2), 151-170.
29. Andrey, J., and Yagar, S. (1993). A temporal analysis of rain-related crash risk. *Accident Analysis & Prevention*, 25(4), 465-472.
30. Carson, J., and Mannering, F. (2001). The effect of ice warning signs on ice-accident frequencies and severities. *Accident Analysis & Prevention*, 33(1), 99-109.
31. Duran, F., ve Teke, M. (2019). Akıllı Yol Durum Sensörü Tasarımı. *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 11(1), 396-401.
32. Donat, W., and Krause, C. (2018). *Learn Raspberry Pi Programming with Python* (Second Edition). California: Apress, 1-26.
33. Leens, F. (2009). An introduction to I2C and SPI protocols. *IEEE Instrumentation & Measurement Magazine*, 12(1), 8-13.
34. Harvey, A. H., McLinden, M. O., and Tew, W. L. (2013, 24 September). *Thermodynamic analysis and experimental study of the effect of atmospheric pressure on the ice point*. Proceedings of the 9th International Temperature Symposium, Anaheim, California, 221-226.
35. Tasdemir, S., and Caliskan, A. (2018). A Practical Approach to Android Mobile Application Security. *International Journal of Applied Mathematics, Electronics Computers*, 6(4), 53-58.
36. Şimşek, M. A., Erdemli, T., ve Taşdelen, K. (2013). Android cihazlarda konum tespiti ve aktarılması. *Akademik Bilişim*, 6(4), 53-58.
37. Harwani, B. M. (2011). *Introduction to Python programming and developing GUI applications with PyQt* (First Edition). USA: Course Technology Inc, 2-10.
38. Özkan, Y. (2016). *Veri madenciliği yöntemleri* (3.Baskı). Türkiye: Papatya Bilim Yayınevi, 11-21.
39. Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37-54.
40. Kartal, E. (2015). *Sınıflandırmaya Dayalı Makine Öğrenmesi Teknikleri ve Kardiyolojik Risk Değerlendirmesine İlişkin Bir Uygulama*, Doktora Tezi, İstanbul Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 13-15.
41. Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 13-22.
42. Bache, K., and Lichman, M. (2013). UCI Machine Learning Repository. *School of Information Computer Science*, 28.
43. Han, J., Pei, J., and Kamber, M. (2011). *Data mining: concepts and techniques* (Third Edition). USA: Elsevier, 56-64.

44. Tan, P. N., Steinbach, M., and Kumar, V. (2006). *Introduction to Data Mining*, Pearson New International Edition (First Edition). United Kingdom: Pearson Education Limited, 5-24.
45. Suthar, B., Patel, H., and Goswami, A. (2012). A survey: classification of imputation methods in data mining. *International Journal of Emerging Technology Advanced Engineering*, 2(1), 309-12.
46. Somasundaram, R., and Nedunchezian, R. (2011). Evaluation of three simple imputation methods for enhancing preprocessing of data with missing values. *International Journal of Computer Applications*, 21(10), 14-19.
47. Bedogni, G., and Tsybakov, A. B. (2009). *Clinical prediction models - A practical approach to development, validation and updating* (2009 Edition). Berlin: Springer, 115-118.
48. Bulut, F. (2016). *Performance evaluations of supervised learners on imbalanced datasets*. Electric Electronics, Computer Science, Biomedical Engineerings' Meeting, İstanbul.
49. Cihan, P., Kalıpsız, O., and Gökçe, E. (2017). Hayvan Hastalığı Teşhisinde Normalizasyon Tekniklerinin Yapay Sinir Ağı ve Özellik Seçim Performansına Etkisi. *Electronic Turkish Studies*, 12(11).
50. Bishop, C. M. (1995). *Neural networks for pattern recognition* (First Edition). New York: Oxford University Press, 295-298.
51. Jayalakshmi, T., and Santhakumaran, A. (2011). Statistical normalization and back propagation for classification. *International Journal of Computer Theory Engineering*, 3(1), 1793-8201.
52. Cheng, J. G. (2000). *Outlier management in intelligent data analysis*, Doctor of Philosophy Thesis, Department of Computer Science Birkbeck College University of London, London, 16-21.
53. Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, 11(1), 1-21.
54. Vellido, A., Martín-Guerrero, J. D., and Lisboa, P. J. (2012, 25-27 April). *Making machine learning models interpretable*. European Symposium on Artificial Neural Networks, Bruges, Belgium.
55. Flach, P. (2012). *Machine learning: The art and science of algorithms that make sense of data* (First Edition). England: Cambridge University Press, 32-38.
56. Roiger, R. J. (2017). *Data mining: A tutorial-based primer* (Second Edition). New York: CRC Press, 34-40.
57. Coşkun, C., ve Baykal, A. (2011). Veri Madenciliğinde Sınıflandırma Algoritmalarının Bir Örnek Üzerinde Karşılaştırılması. *Akademik Bilişim*, 1-8.

58. Marsland, S. (2015). *Machine learning: An algorithmic perspective* (Second Edition). New York: CRC Press, 179-183.
59. Nizam, H., ve Akın, S. S. (2014, 27-29 Kasım). *Sosyal Medyada Makine Öğrenmesi ile Duygu Analizinde Dengeli ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması*. XIX. Türkiye'de İnternet Konferansı, İzmir, 1-6.
60. Alpaydin, E. (2014). *Introduction to Machine Learning* (Second Edition). London: MIT Press, 489-493.
61. Liu, Y., Cheng, J., Yan, C., Wu, X., and Chen, F. (2015). Research on the Matthews correlation coefficients metrics of personalized recommendation algorithm evaluation. *International Journal of Hybrid Information Technology*, 8(1), 163-172.
62. Cover, T. M., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27.
63. Çalışkan, S. K., ve Soğukpınar, İ. (2008). KxKNN: K-Means ve K En Yakın Komşu Yöntemleri ile Ağlarda Nüfuz Tespiti. *EMO Yayınları*, 120-124.
64. Bhatia, N. (2010). Survey of nearest neighbor techniques. *International Journal of Computer Science and Information Security*, 8, 302-305.
65. Pan, J. S., Qiao, Y. L., and Sun, S. H. (2004). A fast K nearest neighbors classification algorithm. *Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 87(4), 961-963.
66. Al-Hudairy, H. (2004). *Data Mining and Decision Making Support in The Governmental Sector*, Master of Science Thesis, Department of Computer Engineering and Computer Science University of Louisville, Kentucky, 33-37.
67. Duda, R. O., and Hart, P. E. (2000). *Pattern Classification* (Second Edition). New York: John Wiley Sons, 112-118.
68. Domingos, P., and Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, 29(2-3), 103-130.
69. Stigler, S. M. (1982). Thomas Bayes's bayesian inference. *Journal of the Royal Statistical Society: Series A*, 145(2), 250-258.
70. Çentik, G. (2013). *Makine öğrenmesi yöntemlerinin polisomnografik verilere uygulanması*, Doktora Tezi, Trakya Üniversitesi Fen Bilimleri Enstitüsü, Edirne, 33-34.
71. Safavian, S. R., and Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man and Cybernetics*, 21(3), 660-674.
72. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81-106.
73. Korkem, E. (2013). *Mikroarray Gen Ekspresyon Veri Setlerinde Random Forest ve Naive Bayes Sınıflama Yöntemleri Yaklaşım*, Yüksek Lisans Tezi, Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara, 4-9.

74. Akman, M., Genç, Y., ve Ankaralı, H. (2011). Random forests yöntemi ve sağlık alanında bir uygulama. *Türkiye Klinikleri Journal of Biostatistics*, 3(1), 36-48.
75. Yılmaz, H. (2014). *Random Forests yönteminde kayıp veri probleminin incelenmesi ve sağlık alanında bir uygulama*, Yüksek Lisans Tezi, Eskişehir Osmangazi Üniversitesi Sağlık Bilimleri Enstitüsü, Eskişehir, 23-28.
76. Alpar, R. (2017). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler* (5. Baskı). Ankara: Detay Yayıncılık, 396-400.
77. Hilbe, J. M. (2009). *Logistic Regression Models* (First Edition). New York: CRC Press, 63-68.
78. Ulutürk Akman, S. (2011, 10 Ekim). *Tüketicilerin Fiyat Bilinci Üzerinde Etkili Olan Faktörlere İlişkin Bir İnceleme*, Maliye Araştırma Merkezi Konferansları, 46, 129-147.
79. Karkacier, O., ve Karabaş, S. (2013). İyi tarım uygulamaları ve tüketici davranışları (logit regresyon analizi). *Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Dergisi*, (2), 71-79.
80. Topuz, D., ve Çakır, M. (2003). Lojistik regresyon analiz tekniğinin eğitim bilimleri araştırmalarında uygulanabilirliği ile ilgili bir araştırma. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 56-80.
81. Albert, A., and Anderson, J. A. (1984). On The Existence of Maximum Likelihood Estimates in Logistic Regression Models. *Biometrika*, 71(1), 1-10.
82. Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied Logistic Regression* (Third Edition). New Jersey: John Wiley & Sons, 35-42.
83. Tu, Y. K., Kellett, M., Clerehugh, V., and Gilthorpe, M. S. (2005). Problems of correlations between explanatory variables in multiple regression analyses in the dental literature. *British Dental Journal*, 199(7), 457.
84. Menard, S. (2002). *Applied Logistic Regression Analysis* (Second Edition). United States of America: Sage Publications, 41-63.
85. Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
86. Eray, O. (2008). *Destek vektör makineleri ile ses tanıma uygulaması*, Yüksek Lisans Tezi, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli, 44-56.
87. Akpınar, H. (2014). *Data: Veri madenciliği veri analizi* (2. Baskı). Türkiye: Papatya Yayıncılık, 230-238.
88. Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, 1289-1305.
89. Narendra, P. M., and Fukunaga, K. (1977). A branch and bound algorithm for feature subset selection. *IEEE Transactions on Computers*, (9), 917-922.

90. Doak, J. (1992). An evaluation of feature selection methods and their application to computer security. *Techninal Report, CSE*, 3-18.
91. Liu, H., and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge Data Engineering*, (4), 491-502.
92. Saeys, Y., Inza, I., and Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19), 2507-2517.
93. Guyon, I., and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 1157-1182.
94. Whitney, A. W. (1971). A direct method of nonparametric measurement selection. *IEEE Transactions on Computers*, 100(9), 1100-1103.
95. Pratama, S. F., Muda, A. K., Choo, Y. H., and Muda, N. A. (2011). Computationally inexpensive sequential forward floating selection for acquiring significant features for authorship invarianceness in writer identification. *International Journal on New Computer Architectures and Their Applications*, 1(3), 581-598.
96. Marill, T., and Green, D. (1963). On the effectiveness of receptors in recognition systems. *IEEE Transactions on Information Theory*, 9(1), 11-17.
97. Stearns, S. (1976). On selecting features for pattern recognition. *International Conference, Coronado*, 71-75.
98. Pudil, P., Novovičová, J., and Kittler, J. (1994). Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11), 1119-1125.
99. Eskidere, Ö. (2012). Ses Ölçümlerinden Parkinson Hastalığının Teşhisi İçin Öznitelik Seçme Yöntemlerinin Karşılaştırılması. *Sigma: Mühendislik ve Fen Bilimleri Dergisi*, 30(4), 402-414.
100. Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3), 389-422.
101. Ladha, L., and Deepa, T. (2011). Feature selection methods and algorithms. *International Journal on Computer Science*, 3(5), 1787-1797.
102. Teke, M., and Duran, F. (2019). The design and implementation of road condition warning system for drivers. *Measurement and Control*, 52(7-8), 985-994.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : TEKE, Mustafa
 Uyruğu : T.C.
 Doğum tarihi ve yeri : 12.11.1984, Ankara
 Medeni hali : Evli
 Telefon : 0 (376) 213 11 95
 e-mail : mustafateke@karatekin.edu.tr



Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Doktora	Gazi Üniversitesi	Devam ediyor.
Yüksek lisans	Karabük Üniversitesi Elek. ve Bilg. Eğitimi	2012
Lisans	Gazi Üniversitesi Elektronik Öğretmenliği	2007

İş Deneyimi

Yıl	Yer	Görev
2009-Halen	Çankırı Karatekin Üniversitesi	Öğretim Görevlisi

Yabancı Dil

İngilizce

Yayınlar

1. Teke, M., and Duran, F. (2019). The design and implementation of road condition warning system for drivers. *Measurement and Control*, 52(7-8), 985-994.
2. Duran, F., ve Teke, M. (2019). Akıllı Yol Durum Sensörü Tasarımı. *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 11(1), 396-401.

Hobiler

Yüzmek, Seyahat etmek, Sinemaya gitmek



GAZİ GELECEKTİR..