T.C.
İSTANBUL ÜNİVERSİTESİ-CERRAHPAŞA
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

YAPAY ZEKA YÖNTEMLERİ KULLANILARAK İNSAN EYLEMİ TANIMA

Engin SEVEN

DANIŞMAN
Doç. Dr. Eylem YÜCEL DEMİREL

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

TEMMUZ-2021

Bu çalışma, 27.07.2021 tarihinde aşağıdaki jüri tarafından Bilgisayar Mühendisliği Anabilim Dalı, Bilgisayar Mühendisliği Programında Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Doç. Dr. Eylem YÜCEL DEMİREL İ(Danışman)
İstanbul Üniversitesi-Cerrahpaşa
Mühendislik Fakültesi

Doç. Dr. Rüya ŞAMLI
İstanbul Üniversitesi-Cerrahpaşa
Mühendislik Fakültesi

Dr. Öğr. Üyesi Zeynep TURGUT AKGÜN
İstanbul Medeniyet Üniversitesi
Mühendislik ve Doğa Bilimleri Fakültesi

Unvan Adı SOYADI
Üniversite
Fakülte

Unvan Adı SOYADI
Üniversite
Fakülte

ÖNSÖZ

Çalışmalarım boyunca kıymetli bilgi, birikim ve tecrübeleri ile bana yol gösteren güler yüzünü ve samimiyetini benden esirgemeyen değerli danışman hocam Doç.Dr. Eylem YÜCEL DEMİREL'e teşekkürü bir borç biliyor ve şükranlarımı sunuyorum.

Bana her zaman maddi ve manevi desteklerini esirgemeyen, beni hiçbir zaman yalnız bırakmayan sevgili aileme teşekkürlerimi sunuyorum.

Temmuz 2021 Engin SEVEN

İÇİNDEKİLER

Sayfa No

ÖNSÖZ ..... **iv**
İÇİNDEKİLER ..... **v**
ŞEKİL LİSTESİ ..... **vii**
TABLO LİSTESİ ..... **viii**
KISALTMA LİSTESİ ..... **ix**
ÖZET ..... **x**
SUMMARY ..... **xii**
1 GİRİŞ **1**
2 GENEL KISIMLAR **3**
2.1 LİTERATÜRDEKİ ÇALIŞMALAR 3
2.2 MAKİNE ÖĞRENMESİ 4
2.2.1 Denetimli Öğrenme 4
2.2.2 Denetimsiz Öğrenme 5
2.2.3 Pekiştirmeli Öğrenme 5
2.3 DERİN ÖĞRENME 6
2.3.1 Evrişimli Sinir Ağı 7
2.3.2 Tekrarlayan Sinir Ağı 9
2.4 TOPLULUK ÖĞRENME (ENSEMBLE LEARNING) 9
2.4.1 Torbalama (Bootstrap Aggregating) 10
2.4.2 Yükseltme (Boosting) 11
2.4.3 Yığma (Stacking) 11
2.5 PERFORMANS DEĞERLENDİRME METRİKLERİ 12
2.5.1 Karmaşıklık Matrisi (Confusion Matrix) 12
2.5.2 Hassasiyet ve Kesinlik 14
2.5.3 F1 Değeri 14
3 MALZEME VE YÖNTEM **15**
3.1 VERİ ÖNİŞLEME 15
3.2 İNSAN EYLEMİ TANIMA İÇİN YIĞMA TOPLULUK ÖĞRENME MODELİ 16

4 BULGULAR.....20
4.1 LOJİSTİK REGRESYON META ÖĞRENİCİYE SAHİP YIĞMA TOPLULUK ÖĞRENME MODELİ.....20
4.2 ÇOK KATMANLI SİNİR AĞI META ÖĞRENİCİYE SAHİP YIĞMA TOPLULUK ÖĞRENME MODELİ.....24
5 TARTIŞMA VE SONUÇ 29
6 KAYNAKLAR 31
7 ÖZGEÇMİŞ 34

ŞEKİL LİSTESİ

Sayfa No

Şekil 2.1: Yapay Sinir Ağı Katmanları. 7

Şekil 2.2: Evrişimli Sinir Ağı Modeli 8

Şekil 2.3: Yığma Topluluk Öğrenme Modeli Mimarisi. 12

Şekil 2.4: Karmaşıklık Matrisi. 13

Şekil 4.1: Model A Doğruluk ve Kayıp Değerleri Grafiği. 20

Şekil 4.2: Model B Doğruluk ve Kayıp Değerleri Grafiği. 21

Şekil 4.3: Model C Doğruluk ve Kayıp Değerleri Grafiği. 21

Şekil 4.4: Model D Doğruluk ve Kayıp Değerleri Grafiği. 22

Şekil 4.5: Lojistik Regresyon Meta Öğreniciye Sahip Topluluk Öğrenme Modeli Sınıflandırıcıları. 23

Şekil 4.6: Model I Doğruluk ve Kayıp Değerleri Grafiği. 24

Şekil 4.7: Model II Doğruluk ve Kayıp Değerleri Grafiği. 25

Şekil 4.8: Model III Doğruluk ve Kayıp Değerleri Grafiği. 25

Şekil 4.9: Model IV Doğruluk ve Kayıp Değerleri Grafiği. 26

Şekil 4.10: Çok Katmanlı Sinir Ağı Meta Öğreniciye Sahip Topluluk Öğrenme Modeli Sınıflandırıcıları 28

TABLO LİSTESİ

Sayfa No

Tablo 3.1: Veri Seti İçerisindeki Sınıflardan Örnekler. 15

Tablo 3.2: Evrişimli Sinir Ağları Modellerinin İç Yapıları. 17

Tablo 3.3: Evrişimli Sinir Ağı ve Çok Katmanlı Sinir Ağı İç Yapısı. 18

Tablo 4.1: A, B, C, D ve Topluluk Öğrenme Modellerine ait VA ve VL Oranları. 22

Tablo 4.2: Lojistik Regresyon Meta Öğreniciye Sahip Topluluk Öğrenme Modeli Performans Metrikleri. 23

Tablo 4.3: I, II, III, IV ve Topluluk Öğrenme Modellerine ait VA ve VL Oranları. 26

Tablo 4.4: Çok Katmanlı Meta Öğrenici Modele Sahip Yığma Topluluk Öğrenme Modeli Performans Metrikleri. 27

KISALTMA LİSTESİ

Kısaltmalar	Açıklama
CNN	: Convolutional Neural Network (Evrişimli Sinir Ağları)
YSA	: Yapay Sinir Ağları
VA	: Validation Accuracy (Doğrulama Değeri)
DL	: Deep Learning (Derin Öğrenme)
EL	: Ensemble Learning (Topluluk Öğrenme)
VL	: Validation Loss (doğrulama veri setine ait kayıp değeri)
SEL	: Stacking Ensemble Learning (Yığma Topluluk Öğrenme)
ML	: Machine Learning (Makine Öğrenmesi)

ÖZET

YAPAY ZEKA YÖNTEMLERİ KULLANILARAK İNSAN EYLEMİ TANIMA

YÜKSEK LİSANS TEZİ

Engin SEVEN

İstanbul Üniversitesi-Cerrahpaşa

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Danışman : Doç. Dr. Eylem YÜCEL DEMİREL

İnsan eylemi tanıma son yıllarda sağlık, güvenlik ve oyun gibi birçok alanda yer bulmuş ve araştırmacılar tarafından bu konuda çeşitli çalışmalar yapılmıştır. Bu alanda yapılan çalışmalar giderek artmaktadır. Bilgisayarlı görü alanındaki son teknolojik gelişmeler ile birlikte resim ve video gibi verilerden insan eylem, jest ve mimikleri analiz edilip karar süreçleri için değerlendirilmektedir.

Gerçek hayat görüntüleri içerisinden insan eylemlerini analiz ederek sınıflandırmak aşılması gereken bir problem alanı olarak karşımıza çıkmaktadır. Bu görüntülerde, farklı arka plan, hareketli kamera ve kötü hava koşulları gibi olumsuz etkenlerden dolayı insan eylemlerinin tanınmasında zorluklar yaşanmaktadır. Kullanılan veri seti içerisinde 3 kanallı (Kırmızı, Yeşil, Mavi) ve 16 insan eylemi sınıfına ait resimler yer almaktadır. Yapılan çalışmada temel olarak iki farklı topluluk öğrenme modeli geliştirilmiştir. Ayrıca bu modeller farklı mimarilere sahip evrişimli sinir ağı modelleri ile karşılaştırılmıştır. Yığma topluluk öğrenme modellerimizden ilkinde temel öğrenici katmanında evrişimli sinir ağı modeli kullanılarak veri seti ile öğrenme gerçekleştirilmiştir. Temel öğrenicilere ait tahminler, meta öğrenici katmanında olan Lojistik Regresyon yöntemi ile öğrenme yapılmıştır. İkinci yığma topluluk öğrenme modeli olarak

temel öğreniciler katmanında farklı evrişimli sinir ağı modeli kullanılarak veri seti ile öğrenme gerçekleştirilmiştir. Daha sonra meta öğrenici katmanında çok katmanlı sinir ağı modeli kullanılarak alt katmandan gelen tahminler üzerinde öğrenme yapılmıştır. Evrişimli sinir ağı modelleri ile yığma topluluk öğrenme modelleri performans metrikleri olarak Hassasiyet, Kesinlik ve F1 değerleri kullanılarak karşılaştırılmıştır. Deneysel çalışmalar sonucunda çok katmanlı sinir ağı meta öğreniciye sahip yığma topluluk öğrenme modeli diğer derin öğrenme modellerinden daha başarılı sonuçlar elde etmiştir.

Temmuz 2021, 47 sayfa.

Anahtar kelimeler: İnsan Eylemi Tanıma, Derin Öğrenme, Evrişimli Sinir Ağı, Topluluk Öğrenme

SUMMARY

HUMAN ACTION RECOGNITION USING ARTIFICIAL INTELLIGENCE METHODS

M.Sc. THESIS

Engin SEVEN

Istanbul University-Cerrahpasa

Institute of Graduate Studies

Department of Computer Engineering

Supervisor : Assoc. Prof. Dr. Eylem YÜCEL DEMİREL

Recognition of human action has taken place in many areas such as health, safety and play in recent years and various studies have been carried out by researchers. Studies in this area are increasing day by day. With the latest technological developments in the field of computer vision, human actions, gestures and facial expressions from data such as pictures and videos are analyzed and evaluated for decision processes.

Analyzing and classifying human actions from real life images is a problem area that needs to be overcome. These images have difficulties in recognizing human actions due to negative factors such as different background, moving camera and bad weather conditions. Our dataset includes 3-channel (Red, Green, Blue) and 16 human action classes. In our study, basically two different ensemble learning models were developed. In addition, these models are compared with convolutional neural network models with different architectures. The first of our stacking ensemble learning models was created using the convolutional neural network model in the basic learner layer. Learning was carried out on our data set using this model. The predictions of the basic learners are fed to the meta learner in the upper layer. Learning was carried out with the Logistic Regression method in the meta learner layer. As the second stacking ensemble

learning model, a different convolutional neural network model was used in the basic learners layer. Learning was carried out with the data set using the stacking ensemble learning model. Then, the training was carried out using the predictions from the meta-learner layer. A multilayer neural network is used in the meta-learning layer of our model. Convolutional neural network models and heap ensemble learning models were compared using precision, F1 scores as performance metrics. As a result of experimental studies, the stacking ensemble learning model, which uses a multilayer neural network in the meta-learner layer, has achieved more successful results than other deep learning models.

July 2021, 47 pages.

Keywords: Human Action Recognition, Convolutional Nerual Networs, Deep Learning, Ensemble Learning

1 GİRİŞ

İnsan vücudu belirli uzuvlardan meydana gelmektedir. İnsanlar birbirleri ile iletişim kurarken sesli iletişimin yanı sıra uzuvların hareketleri ile karşı tarafa belirli mesajlar iletebilmektedir. İnsan faaliyetleri, eylemde bulunan vücut bölümlerinin hareketlerine göre 4 kategoriye ayrılmaktadır. Bunlar sırasıyla jestler, eylemler, etkileşimler ve grup etkinlikleridir [1]. Jestler, yüz ve vücudun belirli bölümleri ile yapılan karşı tarafa kısa mesajlar iletmek için kullanılan bir yöntemdir. Jestler, sözlü ya da sesli bir mesaj değildir. Eylemler, yürümek ve koşmak gibi daha senkronize hareketler bütünüdür. Etkileşimler, en fazla iki aktör tarafından gerçekleştirilen hareketler bütünü olarak tanımlanmaktadır. Hareketlerin, eylemlerin ve etkileşimlerin bir karışımı olan aktiviteler ise grup etkinlikleri olarak tanımlanmaktadır.

İnsan Eylemi Tanıma (Human Action Recognition) resim ya da video görüntülerinden insanın doğasını, yaptığı eylemleri tanımlamayı amaç edinmektedir. Bu alan artan talep nedeniyle araştırmacıların yoğun ilgisini çekmektedir. Güvenlik, sağlık, oyun ve video indeksleme gibi birçok alanda insan eylemini analiz etmek ve anlamlandırmak önemli bir ihtiyaç haline gelmiştir. İnsan eylemi tanımadaki asıl amaç gerçek dünyadaki bir resim ya da sahneyi analiz etmek ve anlamlandırmaktır. Bu amaç doğrultusunda gerçek hayat problemlerinde çok çeşitli zorluklarla karşılaşılmaktadır. Bunlardan bazıları dağınık ve dinamik arka plan, kamera hareketi, kötü hava koşulları ve yetersiz verilerdir.

Dağınık ve dinamik arka plan sorunu başa çıkılması gereken problemlerdendir. Çoğu yöntem arka planın statik ve durağan olduğu iç mekanlarda denenmektedir. Yapılan çalışmalarda dış ortamlarda performansları düşmektedir. Dağınık ve dinamik bir arka plan gürültüye neden olacak ve performansı etkileyecektir [2]. Kamera hareketi diğer başa çıkılması gereken problemlerden birisi olarak karşımıza çıkmaktadır. Sabit kameralarda yapılan çalışmaların hareketli kameralarda performans düşüklüğüne sebep olduğu gözlenmiştir [3]. İnsan eylemi tanımada karşılaşılan bir diğer zorluk ise kötü hava koşullarıdır. Karanlık, yoğun yağmur, kar ve sis gibi hava durumu değişiklikleri görüntüyü önemli ölçüde bozmaktadır [4]. Yetersiz veriler nedeniyle bilgisayarlı görü alanında bazı sorunlar ile karşılaşılmaktadır. Özellikle verinin ölçek aralıkları ve sınıf içi varyasyonları iyi temsil etmesi gerekmektedir. Bunun ile

birlikte insan eylemi için sınırlı sayıda veri setleri olması çalışmalarda aşılması gereken bir problem olarak ifade edilmektedir [5].

İnsanların yemek yemek, uyumak, koşmak ve televizyon izlemek gibi günlük yaşam aktivitelerinin takip edilmesi ve anlamlandırılarak kaydedilmesi gerekliliği ortaya çıkmıştır. Bu artan ihtiyaçları karşılamak için genelde iki yaklaşım benimsenmektedir. Bunlardan ilki algılayıcı (sensor) tabanlı yaklaşımlar ikincisi ise görüntü tabanlı yaklaşımlardır. Bu alandaki öncü yaklaşım görüntü temelli yaklaşım olarak görülmektedir. Görüntü temelli yaklaşımlar algılayıcı temelli yaklaşımlara göre eylemi analiz edilmeye çalışılan kişiye konfor alanı sağlamaktadır. Resim ya da kamera görüntüleri ile yakalanan veriler üzerinden bilgisayarla görme teknikleri uygulanarak insan eylemleri analiz edilip anlamlandırılmaktadır. Bilgisayarlı görü teknikleri kullanımı kolay olmaktadır. İnsan eylemi tanımada algılayıcı temelli yaklaşımlarda vücuda giyilen bir dizi algılayıcı olması ve bunların vücut üzerinde sürekli taşınması belirli zorluklara yol açmaktadır. Sağlık alanında uygulanmış giyilebilir algılayıcılardan oluşan çalışmalar yer almaktadır [6]. Bununla birlikte bir diğer sorun bazı insanların bu algılayıcı ya da etiketleri takmayı unutmalarıdır.

Bu çalışmanın amacı, insan eylemi tanıma alanında karşılaşılan problemlerin üstesinden gelebilecek derin öğrenme modellerinden biri olan Topluluk Öğrenme Modelini test etmektir. Amacımız gerçek hayat görüntülerinden oluşan bir dizi insan eylemini analiz etmek ve tanımlamaya çalışmaktır. Yaptığımız çalışmada topluluk öğrenmesi modellerinden biri olan Yığma Topluluk Öğrenme Modeli kullanılmıştır. Bu modelde iki ayrı bölüm olup birinci bölümde veri setini işleyen öğrenme modelleri yer almakta ikinci bölümde ise bu verilerden çıkarılan tahminler üzerinden eğitilen meta öğrenici model yer almaktadır. Çalışmada iki ayrı Yığma topluluk öğrenme mimarisi oluşturularak test edilmiştir. İnsan eylemi tanıma için kullanılan veri seti belirli ön işlemlerden geçirilerek oluşturulan öğrenme modelinin hesaplama maliyetleri azaltılmaya çalışılmıştır.

2 GENEL KISIMLAR

2.1 LİTERATÜRDEKİ ÇALIŞMALAR

İnsan eylemi tanıma problemi araştırmacıların artan oranda ilgisini çekmiş ve birçok çalışma gerçekleştirilmiştir. Çalışmalar genel olarak iki alt kategoriye ayrılmaktadır. Bunlardan ilki insan eylemi tanıma için algılayıcı (Sensor) tabanlı yaklaşımlar ikincisi ise görüntü işlemeye dayalı yaklaşımlardır. Her iki yöntemin de birbirlerine göre üstünlükleri ve zorlukları vardır. Algılayıcı tabanlı yaklaşımlarda çeşitli algılayıcılardan gelen verilerin belirli parametrelere göre değerlendirilmesi ve insan eylemi tanıma için analizler gerçekleştirilmesi esas alınmıştır [7]. Algılayıcı temelli yaklaşımlarda giyilebilir algılayıcılar yardımı ile spor aktivitelerinin izlenmesi gibi eylemler analiz edilmeye çalışılmıştır [8].

Görüntü temelli yaklaşımlar, insan eylemi tanımada insan bedeninden bağımsız olması gibi belirli kolaylıklar sağlamaktadır. İnsan görüntüsü belirli esnekliğe sahip uzuvlardan meydana gelmektedir. Farklı ortamlarda boyutları farklı olarak görünmektedir. Bu problemin basit bir çözümü görünüş ve şekil özelliklerini çıkarmaktır [9]. Yaşlıların izlenmesi ve sağlık sorunları ile karşılaştıklarında sağlık birimlerine haber verilmesi için dinamik Bayes ağı (Dynamic Bayesian Networks) temelli bir yaklaşım önerilmiştir [10]. Sınırlı miktarlarda eğitim verileri ile başa çıkmak için transfer öğrenme modeli yaklaşımları denenmiştir [11]. Video görüntüleri üzerinde insan eylemi tanıma önemli bir problem olarak karşımıza çıkmaktadır. Video görüntüleri üzerinde optik akış vektörlerini kullanarak uzun vadeli zamansal evrişim ile sinir ağlarını kullanarak insan eylemi tanıma gerçekleştirmişlerdir. İnsan eylemi tanıma alanında yapılan bir diğer çalışmada uzun kısa vadeli bellek ile evrişimli sinir ağlarını birleştiren yöntemler denenmiştir [12].

İnsan eylemi tanıma alanında yapılan çalışmalarda kızıl ötesi, ivme ölçer, derinlik ölçer gibi algılayıcılara dayanan yaklaşımlar kullanılmış bu algılayıcılar insan vücuduna giyilerek eylem tanıma ve analiz etme gerçekleştirilmiştir. Bu algılayıcılardan gelen veriler üzerinde makine öğrenmesi yöntemleri kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Bir diğer çalışmada görüntü tabanlı yöntemler benimsenmiştir. Görüntüler üzerinde evrişimli sinir ağları, destek vektör makineleri, uzun kısa vadeli bellek ya da bir den fazla yöntemin bir arada kullanıldığı modeller önerilmiştir.

2.2 MAKİNE ÖĞRENMESİ

İlerleyen teknolojiler ile birlikte çok büyük miktarlarda veriler ortaya çıkmaktadır. Bu verilerin anlamlı hale getirilmesi ve verilerden yola çıkılarak analizler yapılması zorunlu hale gelmiştir. Bu analizlerin yapılarak büyük veri içerisindeki bilginin çıkarılması ve geleceğe yönelik tahminlerin yapılması makine öğrenmesi yöntemi ile mümkün olduğu görülmektedir. Bu sistemler matematiksel ve istatistiksel yöntemlerden faydalanarak sisteme verilen girdilerdeki örüntülerden yola çıkarak gelecekteki girdilerin çıktılarının ne olacağı ile ilgili tahminler yapılmaktadır [13]. Gelişen dünya ile birlikte birçok alanda insan hatası, verimlilik ve hız gibi problemlerle başa çıkılması zorunlu hale gelmiştir. Bu noktada makine öğrenmesi algoritmaları bu hatalardan arınmış olarak çalışmaktadır. Makine öğrenimi algoritmaları maliyet fonksiyonu ve hiper parametrelerden yola çıkarak en optimum çözümü bulmayı amaçlamaktadır. Maliyet fonksiyonu sistemin ne kadar hata yaptığı ile ilgilenmektedir. Hiper parametreler ise genellikle sistemi tasarlayan tarafından optimize edilmeye çalışılan parametrelerdir. Makine öğrenmesi modelleri aşağıda sırasıyla denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme olmak üzere üç başlık altında açıklanmıştır.

2.2.1 Denetimli Öğrenme

Denetimli öğrenme modeli girdilerin etiketlenerek anlamlandırılması ve sisteme tanıtılması tekniğidir. Sisteme verilecek olan girdinin ne olduğunun söylenerek çıktı arasındaki ilişkiyi sistemin öğrenmesi sağlanmaktadır. Sistem etiketlenen veriler üzerinde eğitildikten sonra girdi ile çıktı arasındaki istatistiksel ve matematiksel ilişkinin ne olduğunu öğrenmektedir. Daha sonra sisteme verilen etiketsiz girdiyi tahmin etmeye çalışmaktadır [14]. Genel olarak kullanılan bazı denetimli öğrenme algoritmaları k-en yakın komşu, karar ağacı, rastgele orman, sinir ağı ve destek vektör makinesi olarak sıralanmaktadır.

Denetimli öğrenme modelinde ortaya çıkan önemli bir sorun ise yeni bir örnekte denetimli öğrenme modelinin nasıl bir performans gösterecek olmasıdır. Denetimli öğrenme belirli bir örnek kümesinden öğrenme gerçekleştirerek tahminde bulunmaya çalışmaktadır. Bu nedenle örnek kümesinin dışında kalacak her bir tahmin performansı olumsuz yönde etkileyecektir. Gerçek bir denetimli öğrenme modeli eğitim verisindeki örüntüleri değil aynı zamanda gürültüleri de öğrenmesi sonucunda performansı olumsuz yönde etkilenecektir. Bu duruma aşırı uyum denir. Aşırı uyum gibi sorunların üstesinden gelmenin en belirgin yolu çapraz doğrulama

yöntemini kullanmaktır. Çapraz doğrulama yöntemi veri setini tek bir parça halinde sistemin öğrenmesi için sunmak yerine onu alt bölümlere ayırarak ayrılmış bölümleri birbirlerini doğrulaması için kullanmaktır [15]. Diğer bir değişle bir alt bölümü model eğitimi için kullanırken kalan diğer bölümleri sonuçların analizi için değerlendirmektedir.

2.2.2 Denetimsiz Öğrenme

Denetimli öğrenmede girdi olarak kullanılan veriler sisteme etiketlenerek verilmekteyken denetimsiz öğrenmede veriler etiketlenmeden sisteme girdi olarak verilmektedir. Bir denetimsiz öğrenme algoritması veriler arasındaki benzerliği ya da ilişkiyi yakalamayı amaçlamaktadır. Verileri kendi aralarında ilişkilerine göre kümelemeye çalışmaktadır [16]. Yaygın olarak kullanılan denetimsiz öğrenme algoritmalarından bazıları K-Ortalamalar (K-Means), Gauss Karışım Modeli (Gaussian Mixture Models – GMM), Temel Bileşen Analizi (Principal Component Analysis – PCA) olarak sıralanmaktadır.

Kümeleme analizi denetimsiz öğrenmede kullanılan başlıca yöntemlerden biridir. Kümeleme analizinde veriler içerisinde bulunan örneklerin özelliklerine göre farklı gruplara ayırma işlemi olarak tanımlanmaktadır. Benzer özelliklere sahip verilerin bir grupta toplanması anlayışına dayanmaktadır. Her veri setinin özelliklerine göre farklı performans sergilemekle birlikte birbirlerine göre artı ve eksileri değişkenlik göstermektedir. Denetimsiz öğrenme modelimizde kullanılan örnek kümemizi belirli özelliklere göre gruplandırmak başarılı olarak görünse de bazı sorunlara açık olarak karşımıza çıkmaktadır. Bu sorun temel olarak öğrenilen örnek kümesi ile ilgili anlamlı bir birliktelik olduğu varsayılmaktadır. Bununla birlikte gerçek bir kümenin ne olduğuna dair kesin bir yargı olmadığı bilinmektedir. Bu örnek kümenin tanımı araştırma alanına ve araştırmanın amacına bağımlıdır.

Bir diğer denetimsiz öğrenme algoritması birliktelik kuralı olarak ifade edilmektedir. Birliktelik kuralında eğitim verisi için bir ölçü belirlenmektedir. Belirlenen ölçü temel alınarak veri seti içerisinde bulunan örneklerin bu ölçüye bakılarak analiz edilmesi ve ayrıştırılması sağlanmaktadır. Böylece veri seti içerisinde bulunan kategorize edilmemiş verilerden faydalanarak özellik vektörleri çıkarmayı amaçlamaktadır.

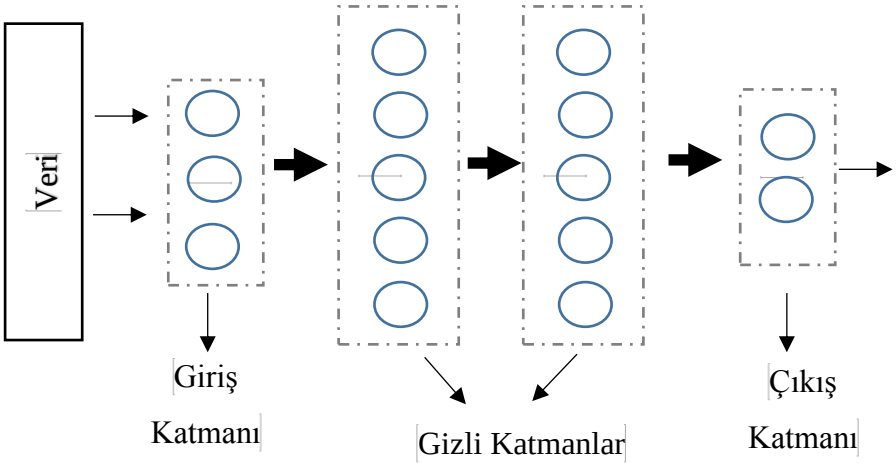
2.2.3 Pekiştirmeli Öğrenme

Pekiştirmeli öğrenme, belirli bir uzay alanında sistemin çevresi ile ilişkiye girmesi ve aldığı tepkilere göre en iyi çözümü bulmaya çalışması olarak ifade edilmektedir [17]. Yöntemde,

çevresi ile kurduğu ilişkide her yapılan doğru davranışta ödül puanı verilmektedir. Yaptığı her yanlış davranışta ceza puanı verilmektedir. Çevresi ile sürekli etkileşimde bulunarak en olası çözümü aramaktadır. Denetimli ve denetimsiz öğrenmede olduğu gibi önceden verilen bir veri yoktur. Kendisi sonsuz sayıda çevresini deneyimlemektedir. Böylece diğer yöntemlerde olduğu gibi ham verinin ön işleme tabi tutulması sisteme uygun hale getirilmesi gibi maliyetler ortadan kaldırılmıştır. Sistemin amacı en uygun davranış politikasını en iyi şekilde öğrenmeye çalışmaktır. Sistem davranış politikasına bakarak en iyi olası davranışı seçmek için hareket etmektedir. Olası durumlar ve yapılabilecek eylemler ile bu eylemler karşısında alınabilecek ödülleri tutan Q-Öğrenme (Q-Learning) kullanarak davranış politikası tablosu elde edilmektedir.

2.3 DERİN ÖĞRENME

Derin öğrenme ilk olarak Nörofizyolog Warren McCulloch tarafından nöronların nasıl işlev göreceğini açıklaması ile başlamıştır. 1950 yıllarına gelindiğinde ise soyut bir sinir ağı geliştirilmiştir [18]. İlerleyen yıllarda çalışmalar devam ederek bügün kullanılan sistemler meydana getirilmiştir. Derin öğrenme modelleri makine öğrenme modeline göre ham verilerin geleneksel yöntemler ile temsili için harcanan mühendislik maliyetinin önüne geçmektedir. Derin öğrenme modelleri ham verilerden anlamsal temsilleri öğrenebilmektedir. Verilerin temsillerini öğrenmek için çok katmanlı yapılar kullanılmaktadır. Sistem içerisinde alt katmanlar verilerin basit özelliklerini yakalarken üst katmanlara doğru veri ilerledikçe daha karmaşık özellikler öğrenilmektedir. Derin öğrenmenin temel yapı taşı yapay sinir ağlarıdır. Sistem içerisinde katman sayıları arttıkça derin öğrenme olarak isimlendirilmektedir. Yapay sinir Ağları genel olarak üç bölümden oluşmakta olup ilk kısımda giriş katmanı ikinci kısımda gizli katmanlar ve son kısımda ise sonuçların üretildiği çıkış katmanı yer almaktadır. Şekil 2.1’de yapay sinir ağlarına ait temel katmanlar gösterilmektedir.



Şekil 2.1: Yapay Sinir Ağı Katmanları.

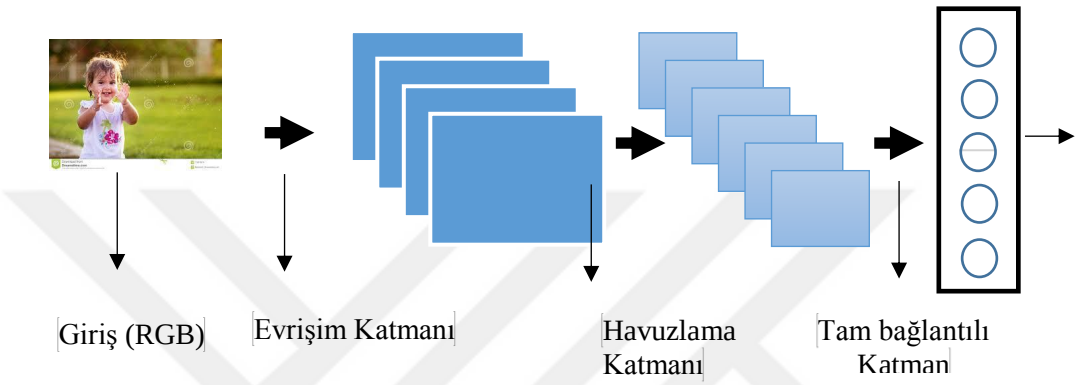
Derin öğrenme modellerinden bazıları mimarilerine göre Evrişimli Sinir Ağı (Convolutional Neural Network-CNN) ve Tekrarlayan Sinir Ağı (Recurrent Neural Network - RNN) olarak sıralanmaktadır. Gelişen teknolojiler ile birlikte her geçen gün yeni mimariler ortaya çıkmakta ve bu konu sürekli gelişim göstermektedir.

2.3.1 Evrişimli Sinir Ağı

Evrişimli sinir ağları ilk olarak görüntüler üzerinde kullanılmış ve başarılı sonuçlar elde edilmiştir [19]. İlerleyen teknolojiler ile birlikte bilgisayarlı görü alanında evrişimli sinir ağları bir çok problem alanına uygulanmış ve başarılı sonuçlar elde edilmiştir. Resim işlemeden video görüntülerine birçok alanda kullanılmaya devam etmektedir. Evrişimli sinir ağının temel amacı canlıların görme yeteneğini taklit etmeye çalışmasıdır. Günümüzde kullanılan evrişimli sinir ağları onlarca katman meydana getirilerek oluşturulmuştur [20]. CNN mimarisi sırasıyla evrişim, havuzlama ve tam bağlantılı katmanlardan meydana gelmektedir. Bu katmanlar problem alanına göre çeşitli olasılıklarla denenerek en olası çözüm aranmaktadır. Kurulan mimaride farklı sayıda ve çeşitte katmanlar yan yana getirilerek oluşturulabilmektedir. Katman sayıları ve sıralaması tamamen tasarlayanın problem alanına çözüm üretecek en iyi mimariyi bulmasına göre şekillenmektedir.

Evrişimli sinir ağlarının ilk katmanı genellikle evrişim katmanı olarak sıralanmaktadır. Bu katmanda girdi olarak verilen görüntünün daha basit olan kenar ve köşe gibi özellikleri öğrenilmektedir. Bu basit özellikler üst katmanlarda birleştirilerek daha karmaşık özelliklerin

öğrenilmesi planlanmaktadır. Evrişim katmanında resim üzerinde filtre denilen bir ızgara resmin pikselleri üzerinde dolaştırılarak özelliklerin çıkarılması sağlanmaktadır. Filtre boyutları farklılık göstermekte ayrıca kullanım amaçlarına göre de farklı filtreler denenmektedir. Genellikle kullanılan filtreler sistemin başlangıcı esnasında tasarlanmakta olup filtrelerin değiştirilmesi başarıyı büyük ölçüde etkilemektedir. Şekil 2.2'de evrişimli sinir ağı modeli gösterilmektedir.



Şekil 2.2: Evrişimli Sinir Ağı Modeli.

Evrişimli sinir ağlarının sonraki katmanında genellikle havuzlama katmanı tasarlanmaktadır. Havuzlama katmanının en büyük kullanım amacı bir önceki evrişim katmanından gelen verilerin görüntü için içerdiği özellikleri kaybetmeden daha az boyutta ifade edilmeye çalışılmasıdır. Böylece evrişim katmanından gelen yüksek hesaplama maliyeti olan verinin daha az maliyetle özelliklerini korumaya çalışmaktır. En sık kullanılan havuzlama yöntemleri minumum, ortalama ve maksimum havuzlamadır. Örneğin ilk katmanda görüntü pikselleri üzerinde filtrelerin dolaştırılmasından sonra ortaya çıkan özelliklerin belirli sayıda piksellerin ortalamasının alınması ortalama havuzlama olarak isimlendirilmektedir. Aynı işlemler maksimum havuzlamada belirli bir bölgedeki piksellerin maksimumunun alınması olarak ifade edilmektedir. Böylece her bir pikselin sistem tarafından öğrenilmeye çalışılmasındansa sadece belirli bölgeleri temsil eden piksellerin öğrenilmesine çalışılmaktadır.

Son olarak evrişimli sinir ağları tüm özellikler düzleştirilerek tam bağlantılı katmana iletilmektedir. Bu katmanda nöron adı verilen yapılar bulunmakta ve görüntünün özellikleri öğrenilmeye çalışılmaktadır.

2.3.2 Tekrarlayan Sinir Ağı

Tekrarlayan sinir ağları, özellikle sıralı ve birbiri arasında bağıntılar bulunan verileri modellemede en uygun algoritmaların başında gelmektedir. Tekrarlayan sinir ağları basit bir ifade ile önceki girişi ve mevcut girişi hatırlama üzerine kurulmuş bir mimaridir. Böylece içerisinde bellek barındırma özelliği bulunmaktadır. Bununla birlikte bir zaman adımındaki çıktının özelliği ondan önceki zaman adımlarındaki çıktının özelliklerine bağlı olarak değişim göstermektedir. Tekrarlayan sinir ağlarının bu özelliği onu doğal dil işleme, duygu analizi ve video analizi gibi konularda başarılı sonuçlar vermesine yol açmaktadır [21].

RNN mimarisi katman sayısı bakımından CNN mimarileri kadar derin olmasa da eğitim esnasında kaybolan gradyan sorununa karşı bir çok araştırmacı çeşitli çalışmalarda bulunmuştur. Tekrarlayan sinir ağları eğitimlerinde ayrıca uzun vadede veriler arasındaki bağımlılıkları öğrenmek için bir çok sorunun üstesinden gelinmesi gerekmektedir. RNN mimarilerinin sıralı verileri modelleme yeteneğinin başarılı olmasına karşın eğitiminin çok zor olduğu ve bu sebeple kullanımının kısıtlandığı görülmektedir [22].

2.4 TOPLULUK ÖĞRENME (ENSEMBLE LEARNING)

Makine öğrenimi alanında yapılan bir çok çalışmada çeşitli yöntemler önerilmiştir. Bu yöntemler modele girdi olarak verilen verilerin özelliklerine göre başarı oranı yüksek ya da başarı oranı düşük olabilmektedir. Bu problemler genelde veri dengesizliği, hata düzeltme, kayıp verilerin değerlendirilmesi gibi bir dizi sorun olarak karşımıza çıkmaktadır. Topluluk öğrenme mimarisinin, bu gibi durumlara çözüm getirdiği ifade edilmektedir [23]. Topluluk öğrenimi birden fazla öğrenen algoritmanın ortak karar vermesi olarak düşünülmektedir. Daha önceki makine öğrenimi modellerinden farklı olarak aynı veri seti üzerinde birden fazla öğrenme modeli çalıştırılmakta ve belirli kurallara göre ortak karar alınmaktadır [24]. Bu tıpkı bir problem alanı için birden fazla öğrenciye eğitim verip problemi ortak karar alarak birlikte çözmelerine benzetilmektedir. Bir topluluk öğrenme modelinde başa çıkılması gereken problem öğrenen arasındaki ilişkinin nasıl olacağı ve nasıl bir mimari oluşturulması gerektiğidir. Model içerisinde yer alan her bir öğrenen modelin farklı varyasyonlarda oluşturulması farklı topluluk öğrenme modelleri meydana getirmektedir. Topluluk öğrenme yapı taşlarından biri olan öğrenen modeller paralel tarzda ya da sıralı bir şekilde öğrenme

gerçekleştirebilmektedir [25]. İyi bir topluluk öğrenme yönteminin, mümkün olduğunca farklı ve doğru seçilmiş olmasının, başarıyı arttırdığı bilinmektedir [26].

Topluluk öğrenme modeli temel olarak iki alt bölüme ayrılmaktadır. Birinci kısımda veri üzerinde işlemler yapan modeller bulunmaktadır. İkinci kısımda ise bu alt modellerden gelen verileri değerlendiren ve karar sürecini işleten bir model yer almaktadır. Öğrenen modeller alt kısımda veri üzerinde işlemler yapmaktadır. Bazen sisteme girdi olarak sunulan veriler bütün olarak her model tarafından öğrenilmekte böylece farklı modellerin öğrenmeleri üst modelde birleştirilmektedir. Bazen de sisteme girdi olarak sunulan veri alt parçalara bölünerek farklı türde öğrenen modellere beslenmekte böylece daha optimal bir çözüm elde edilmektedir.

Topluluk öğrenme modelinin ikinci kısmında meta model yer almakta olup öğrenen modellerin kararlarından yola çıkarak öğrenme gerçekleştirmektedir. Topluluk öğrenme modeli sırasıyla torbalama, yükseltme, oylama ve istifleme olmak üzere farklı mimarilere sahiptir. Gelişen teknolojik gelişmeler bu mimarilerin farklı türlerini ortaya koymaya devam edecektir.

2.4.1 Torbalama (Bootstrap Aggregating)

Torbalama yöntemi topluluk öğrenme yöntemlerinden biri olarak bir çok uygulama alanı bulmuştur. İlk olarak 1996 yılında kullanılmış olup bir temel öğrenicinin farklı versiyonlarının oluşturulması ve veri üzerinde işlem yaparak sonuçların değerlendirilmesine dayanmaktadır [27]. Bir topluluk modelinin bir tek modelden daha iyi sonuçlar elde etmesinin en temel sebebi bias, varyans ve kovaryans olarak ifade edilmektedir [28]. Torbalama yöntemi üç ana aşamadan oluşmaktadır. Bunlar sırasıyla temel öğrenenlerin eğitilmesi, veri setinin alt bölümlere ayrılması ve eğitim için kullanılmayan veri seti ile hiper parametrelerin ayarlanması ve temel öğrenenlerden gelen çıktıların karar mekanizmasında kullanılması olarak ifade edilmektedir. Temel öğrenen yöntemler için birçok model olsa da veri setindeki çeşitliliğin üstesinden gelebildiği için yapay sinir ağları genellikle tercih edilmektedir.

Temel öğrenicilerin birbirinden farklı sonuçlar elde etmesi topluluk öğrenme modelinin başarısını arttıran önemli sebeplerden biridir. Bu nedenle veri seti alt parçalara ayrılarak her bir temel öğrenicilere farklı veri seti verilmektedir. Bir tek model kullanıldığında veri setine bağlı olarak meydana gelen varyans problemini torbalama yöntemi ile en aza indirmek mümkün

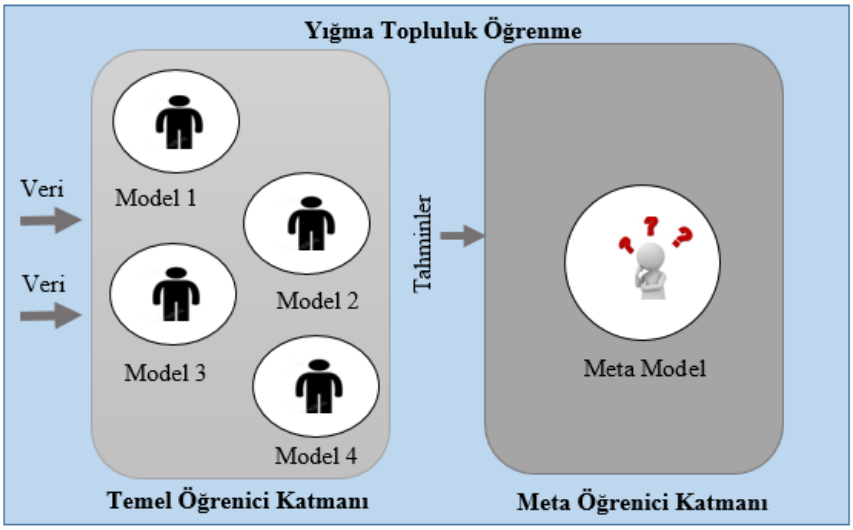
olmaktadır. Torbalama yönteminin kullanılmasında veri setinin yeterince büyük olması önem arz etmektedir. Torbalama yönteminin bir diğer uygulanma biçimi rastgele orman algoritması olarak karşımıza çıkmaktadır.

2.4.2 Yükseltme (Boosting)

Bir veri seti üzerinde temel öğrenicilerin sırasıyla birbirlerinin hatalarından yola çıkarak daha iyi bir öğrenmeye ulaşmaya çalışmasıdır. İlk temel öğrenici model hatayı en çok yapan model olabilirken bu hataların bir sonraki temel öğrenici modelde giderilmesi için aktarılması ve bu şekilde en son temel öğrenici modelde hataların minimize edilmesi amaçlanmaktadır [29]. Bir veri seti üzerinde yükseltme (boosting) algoritması ilk çalıştırıldığında temel öğrenici bazı örnekleri doğru tahmin edecek bazılarını ise yanlış tahmin edecektir. Örneğin 1000 örnekli bir veri setinde 690 tane örneğin doğru, 310 tane örneğin yanlış tahmin edildiği varsayımından yola çıkılarak eğitilen model için yanlış sınıflandırma oranı hesaplanmaktadır. Daha sonra tahminler için ağırlık hesaplaması yapılmaktadır. Burada amaç ilk temel öğrenici model tarafından yanlış tahmin edilen örneklere yüksek ağırlık vermek doğru tahmin edilen örneklere ise düşük ağırlık vermektir. Böylece bir sonraki temel öğrenici model yanlış tahmin edilen örnekleri daha fazla öğrenme eğilimine girecektir.

2.4.3 Yığma (Stacking)

Yığma Topluluk Öğrenme modeli heterojen öğrenme modeli olarak adlandırılabilmektedir. Bu model de daha güçlü bir model elde edebilmek için alt modellerin veri seti üzerinde öğrenmesi ve sonuçların bir üst katmanda yer alan meta öğreniciye entegre edilmesi ile sonuçlanmaktadır [30]. Yığma topluluk öğrenme modeli iki alt bölümden meydana gelmektedir. İlk kısımda temel öğrenici modeller yer almaktadır. İkinci kısımda ise tahminlerden öğrenen model, meta model yer almaktadır. Bir veri seti üzerinde öncelikle temel öğrenici modeller koşulur. Temel öğrenici modellerde bütün bir eğitim setinde öğrenme gerçekleştirilir. Her bir temel öğrenici modelin tahminleri bir meta modele entegre edilerek meta modelin bu tahminlerden yola çıkarak öğrenme gerçekleştirmesi amaçlanmaktadır. Topluluk öğrenme modeli hem temel öğrenicinin yeteneklerinden hem de meta öğrenicinin yeteneklerinden yola çıkarak modelin öğrenme oranını ve doğruluğunu önemli ölçüde arttırmaktadır. Şekil 2.3'de yığma topluluk öğrenme modeli mimarisi gösterilmektedir.



Şekil 2.3: Yığma Topluluk Öğrenme Modeli Mimarisi.

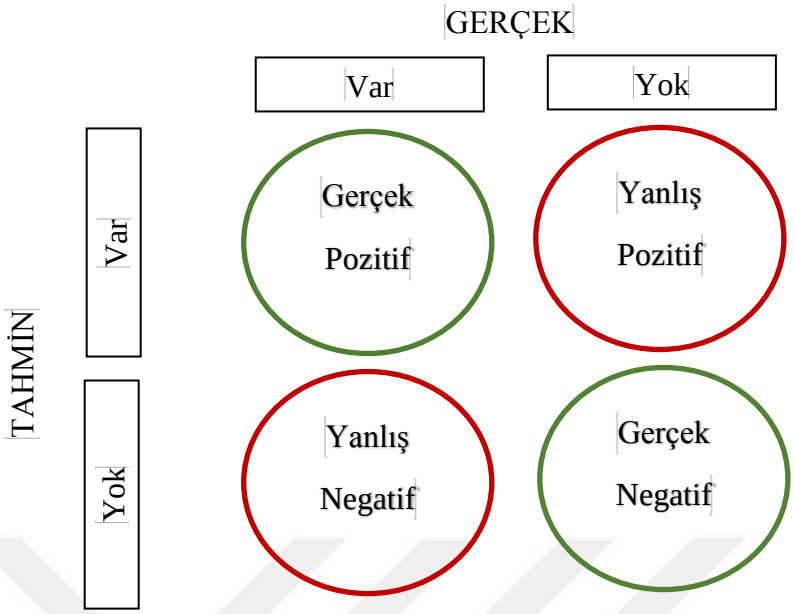
Yığma topluluk öğrenme modelinin başarısı iki temele dayanmaktadır. Bunlardan ilki temel öğrenicilerin veri setinden ne kadar iyi öğrendiği ve ne kadar başarılı tahminlerde bulunduğudur. Temel öğrenici modeller problemi ne kadar çözüme yaklaştırırlarsa nihai modelin başarısı o kadar artma eğilimi göstermektedir. İkinci olarak ise modellerin birbirinden farklı olması ve böylece bu farklılıkların öğrenme başarısından avantaja dönüşmesi söz konusu olmaktadır.

2.5 PERFORMANS DEĞERLENDİRME METRİKLERİ

İnşa edilen bir sistemi değerlendirmek ve sistemin hatalarını görmek en az o sistemi oluşturmak kadar önem arz etmektedir. Bu sebeple her mimarinin, belirli ölçütlerden yola çıkılarak performansı değerlendirilmektedir. Alınan sonuçlar inşa edilen sistemin başarılı olup olamayacağını göstermektedir. Makine öğrenimi alanında en çok başvurulan başlıca ölçütler Karmaşıklık Matrisi, Doğruluk, Hassasiyet, Geri Çağırma ve F1 değeri gibi değerlerden oluşmaktadır. Bu metrikler aşağıda açıklanmıştır.

2.5.1 Karmaşıklık Matrisi (Confusion Matrix)

Sınıflandırma problemlerini çözerken sıklıkla kullanılan ölçülerden biri Karmaşıklık Matrisidir. İkili sınıflandırma başta olmak üzere çoklu sınıflandırma problemlerinde de başarılı bir şekilde ölçüt olarak kullanılmaktadır. Şekil 2.4'de karmaşıklık matrisi ile ilgili ikili sınıflandırmaya ait ölçütler verilmiştir.



Şekil 2.4: Karmaşıklık Matrisi.

Karmaşıklık matrisi gerçek değerler ve sistemin bu değerlere karşı ürettiği tahmini değerlerden meydana gelmektedir. Birinci bölgede yer alan ölçüt gerçek pozitif (True Pozitive-TP) olarak isimlendirilmektedir. TP sistem tarafından doğru bir şekilde sınıflandıran örneklerin sayısını ifade etmektedir. Gerçek Negatif (True Negative-TN), sistem tarafından doğru olarak sınıflandırılan yanlış örneklerin sayısını ifade etmektedir. Yanlış pozitif (False Pozitive-FP), pozitif olarak sınıflandırılan yanlış örnek sayısını ifade etmektedir. Yanlış Negatif (False Negative-FN), sistem tarafından negatif olarak sınıflandırılan gerçek pozitif örneklerin sayısını vermektedir. Sınıflandırma problemlerinden en sık başvurulan ölçütlerden birisi doğruluk (Accuracy) ölçütüdür. Bir modelin doğruluk değeri karmaşıklık matrisi hesaplandıktan sonra aşağıdaki yer alan denklem (2.1)'deki gibi hesaplanmaktadır.

$$\text{Doğruluk (Accuracy)} = \frac{Doğru\ Tahminlerin\ Sayısı}{Toplam\ Tahminlerin\ Sayısı} \tag{2.1}$$

İkili sınıflandırma için karmaşıklık matrisine göre doğruluk değeri denklem (2.2)'deki gibi hesaplanmaktadır.

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.2}$$

Doğruluk ölçütü başlangıçta sistemin değerlendirilmesinde yeterli bir ölçüt olarak görünmesine rağmen sisteme verilen veri setine bağlı olarak performans değişkenlik göstermektedir.

Doğruluk değeri dengeli dağılmış veri setlerinde işe yaramaktadır. Dengesiz verilerde yanıltıcı sonuçlar elde edilebilmektedir.

2.5.2 Hassasiyet ve Kesinlik

Sınıflandırma problemlerinde yaygın olarak kullanılan ölçütlerden bazıları Hassasiyet (Recall) ve Kesinlik (Precision) ölçütleridir.

Kesinlik, modelin pozitif değerleri tahmin etmede ne kadar başarılı olabildiğinin ölçüsünü vermektedir. Böylece öngörülen pozitif bir sonucun doğru ölçülüp ölçülmediğini test etmektedir [31]. Bir diğer deyişle kesinlik, sezgisel olarak sınıflandırıcının negatif olan bir örneği pozitif olarak etiketleme yeteneğidir. Bir diğer metrik olan hassasiyet ise bir modelin gücünü gösteren başlıca ölçüttür. Bir modeldeki olumlu sonuçları öngörmektedir. Burada bahsedilen iki ölçütteki temel amaç kesinlik değerlerini etkilemeden hassasiyeti arttırmaya çalışmaktır. Kesinlik değerini hesaplamak için kullanılan denklem (2.3)'de gösterilmiştir. Hassasiyet değerini hesaplamak için kullanılan denklem (2.4)'de gösterilmiştir.

$$\text{Kesinlik (Precision)} = \frac{TP}{TP+FP} \tag{2.3}$$

$$\text{Hassasiyet (Recall)} = \frac{TP}{TP+FN} \tag{2.4}$$

2.5.3 F1 Değeri

F1 değeri bir modelin olumlu örneklerinin sınıflandırılması için kullanılmaktadır. Bu değer üretilirken hassasiyet ve kesinlik değerlerinden faydalanılmaktadır. Hassasiyet ve kesinlik değerlerinin ağırlıklı armonik ortalaması alınmaktadır. F1 değerini hesaplamak için kullanılan denklem (4.4)'de gösterilmektedir.

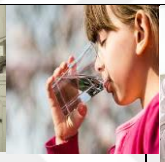
$$\text{F1 değeri} = 2 * \frac{Kesinlik * Hassasiyet}{Kesinlik+Hassasiyet} \tag{4.4}$$

Yukarıda yer alan formülde armonik ortalama alınmadan önce beta değeri kullanılmaktadır. Bu değer genellikle 1 olarak alındığı için F1 değeri olarak da isimlendirilmektedir.

3 MALZEME VE YÖNTEM

İnsan Eylemi Tanıma güvenlikten sağlık sektörüne kadar birçok alanda çalışılmaya devam eden bir problem alanıdır. Yapılan çalışmada, github isimli açık kaynak erişimli web sitesinde yer alan Action-Net isimli insan eylemlerini içeren veri seti kullanılmıştır [32]. Action-Net, bu alanda yapılan çalışmalara katkı sunmak amacıyla insan eylemlerini, jest ve etkinliklerini anlamak üzere toplanmış bir veri kümesidir. Action-Net içerisinden 16 sınıf içermektedir. Toplamda 19.200 resim yer almaktadır. Her sınıf için yaklaşık 1000 resim eğitim verisi için ayrılmış 200 resim ise test verisi için ayılmıştır. Tablo 3.1'de veri setinde yer alan sınıflara ait bazı örnekler yer almaktadır.

Tablo 3.1: Veri Seti İçerisindeki Sınıflardan Örnekler.

						
Telefon ile konuşma	Alkışlama	Bisiklete binme	Dans etme	İçeçek içme	Yemek yeme	Kavga etme

Bu veri seti içerisinde telefon ile konuşma, alkışlama, bisiklete binme, dans etme, içecek içme, yemek yeme ve kavga etme gibi sınıflara ait insan eylem, jest ve etkinlikleri yer almaktadır. Resimler renkli (RGB) olup boyutları farklılık göstermektedir. Resimler günlük hayattan alınmış olup sabit bir arka plana sahip değildir.

3.1 VERİ ÖNİŞLEME

Action-Net veri seti içerisinde yer alan insan eylemi resimleri üç kanallı renkli görüntülerdir. Bu kanallar kırmızı, yeşil ve mavi olarak yer almaktadır. Her kanalda 0 ile 255 arasında piksel değerleri yer almaktadır. Bu sebeple görüntüler derin öğrenme modelleri tarafından hesaplama maliyeti getirmektedir. Bundan dolayı sistemin öğrenmesi zorlaşmaktadır. Modelin öğrenememesinin temel sebebi ağın ağırlıklı girdi toplamı kullanmasıdır. Ağın etkin bir şekilde kullanılması için toplamın küçük tutulması gerekmektedir. Bu problemin üstesinden gelmenin

başlıca yöntemi piksel değerlerini ölçeklendirmektir. Piksel değerlerini ölçeklendirmenin başlıca yöntemleri normalleştirme (Normalization), merkezileştirme (Centering) ve standart hale getirmektir (Standardization).

Normalleştirme işlemi piksel değerlerinin 0 ile 1 arasına yerleştirilmesidir. Görüntü üzerinde yer alan en küçük piksel değerinin sıfıra eşitlenmesi en büyük piksel değerinin bire eşitlenmesi ile yapılmaktadır. Merkezileştirme, görüntü üzerinde yer alan piksel değerlerinin ortalama piksel değeri hesaplanır. Her piksel değerinden çıkarılarak ortalama sıfır üzerinde ortalanmış bir dağılım ile sonuçlanmaktadır. Standardizasyon yöntemi ile görüntüye ait piksel değerleri Gauss eğrisi oluşturacak şekilde ölçeklenmektedir. Dolayısıyla piksellerin ortalaması sıfır standart sapması bir olan bir dağılım elde edilmektedir.

3.2 İNSAN EYLEMİ TANIMA İÇİN YIĞMA TOPLULUK ÖĞRENME MODELİ

Gerçek dünya problemi olarak karşımıza çıkan özellikle güvenlik ve gözetleme gibi alanlarda insan eylemlerinin izlenerek takip edilmesi bu gözlemlerden kararlar çıkarılması gibi iş akışlarına artan oranlarda ihtiyaç duyulmaktadır. Bu talepler nedeniyle araştırmacılar tarafından çok çeşitli çalışmalar yapılmakta çözüm önerileri getirilmektedir.

Yapılan çalışmada Yığma Topluluk Öğrenme Modeli (Stacking Ensemble Learning Model) ile insan eylemi tanıma için çözüm önerileri getirilmiştir. Topluluk öğrenmenin temelinde birden farklı modelin bir araya getirilerek karar verme sürecini birlikte işletmelerini sağlamaktır. Temel öğrenenler açısından topluluk öğrenmesi içerisinde bulunan modeller heterojen ya da homojen olabilmektedir [13]. Topluluk öğrenme modelleri içerisinde etkili yöntemlerden birisi de Yığma öğrenme modelidir. Bu modelin temel fikri temel öğrenicilerin kararlarını alarak meta öğreniciye beslemek ve meta öğrenicinin gelen kararlardan eğitim yapılarak öğrenmesini sağlamaktır. Böylece temel öğrenicilerin yaptığı hataları aza indirerek modelin genelleme hatasını minimize etmektedir [33].

Yapılan çalışmada iki adet mimari kullanılmış ve sonuçlar karşılaştırılmıştır. Birinci mimaride Yığma Topluluk Öğrenme modelinin temel öğreniciler katmanı için 4 adet evrişimli sinir ağı tercih edilmiştir. Üst katman olan meta öğrenici katmanında ise Lojistik Regresyon yöntemi kullanılmıştır. İkinci mimarimizde ise Yığma Topluluk Öğrenme modelimizin ilk katmanında

temel öğreniciler olarak 4 adet evrişimli sinir ağı tercih edilmiş olup ikinci katmanda yer alan meta öğrenici için ise çok katmanlı sinir ağı tercih edilmiştir.

Birinci mimaride 4 farklı evrişimli sinir ağı kullanılmıştır. Bunlardan ilki olan model A, dört kademeli bir mimariye sahiptir. İlk kademede iki adet 16 ve 32 olmak üzere iki filtre kullanılmıştır. Çekirdek boyutu (Kernel Size) olarak 3x3 çekirdek kullanılmıştır. Bir sonraki adımda havuzlama katmanı (Pool Size) olarak (2,2) kullanılmıştır. Çıkan piksel değerleri bir sonraki kademeye aktarılmıştır. Model A'nın ikinci kademesinde 64, 32 ve 64 olmak üzere 3 adet filtre kullanılmıştır. Sırasıyla (1,1), (3,3), (1,1) çekirdek boyutları kullanılmıştır. Ardından (2,2) boyutlarında havuzlama katmanı yer almaktadır. Model A'nın üçüncü kademesinde sırasıyla 64, 128, 64 filtreleri yer almaktadır. Son kademede ise 128 ve 256 boyutlarında filtreler konulmuştur. Sırasıyla her bir filtre için (3,3) ve (1,1) çekirdek boyutları yer almaktadır. Daha sonra bütün veriler 64 nörondan oluşan düzleştirme katmanına beslenmektedir. Tablo 3.2'de temel öğrenici mimarisinde kullanılan evrişimli sinir ağları modellerinin özellikleri belirtilmiştir. Yığma Topluluk Öğrenme modelimizin meta öğrenici kısmında Lojistik Regresyon yöntemi kullanılmıştır.

Tablo 3.2: Evrişimli Sinir Ağları Modellerinin İç Yapıları.

Temel Öğrenici Katmanda Yer Alan CNN Modellerinin Özellikleri				
	Model A	**Model B**	**Model C**	**Model D**
Katman Sayısı	4 katmanlı	5 katmanlı	5 katmanlı	6 katmanlı
Filtre Sayısı	16 x 32 64 x 32 x 64 64 x 128 x 64 128 x 256	16 32 x 64 x 96 64 x 96 x 128 32 x 64 256 x 512	16 32 x 64 x 32 64 x 64 x 96 32 x 64 96 x 128 x 256	16 32 x 64 x 96 128 32 x 64 64 x 96 128 x 512
Çekirdek Boyutu	(3,3) (1,1)	(3,3) (1,1)	(3,3) (1,1)	(3,3) (1,1)

Yığma topluluk öğrenme mimarimizde ikinci olarak temel öğrenici katmanımızda 4 adet farklı evrişimli sinir ağı kullanılmıştır. Evrişimli sinir ağı modellerinden ilki olan model I, 4 katmanlı bir mimariye sahiptir. Model II ise 5 katmanlı bir mimariye sahiptir. Temel öğrenicilerden bir diğeri olan Model III, 5 katmanlı bir mimariden oluşmaktadır. Model IV temel öğrenici mimarisi ise 6 katmanlı bir yapıya sahiptir. Tablo 3.3'te temel öğrenici olarak kullanılan 4 modelin içi mimarisi verilmiştir.

Tablo 3.3: Evrişimli Sinir Ağı ve Çok Katmanlı Sinir Ağı İç Yapısı.

İlk Aşamada Yer Alan CNN Modellerinin Özellikleri					
	Model I	**Model II**	**Model III**	**Model IV**	**Çok Katmanlı Sinir Ağı**
Katman Sayısı	4 katmanlı	5 katmanlı	5 katmanlı	6 katmanlı	4 katmanlı
Filtre Sayısı	16 x 32 64 x 32 x 64 64 x 128 x 64 128 x 256	16 32 x 64 x 96 64 x 96 x 128 32 x 64 256 x 512	16 32 x 64 x 32 64 x 64 x 96 32 x 64 96 x 128 x 256	16 32 x 64 x 96 128 32 x 64 64 x 96 128 x 512	128 64 32 16 7
Çekirdek Boyutu	(3,3) (1,1)	(3,3) (1,1)	(3,3) (1,1)	(3,3)(1,1)	

Meta öğrenici katmanında ise Lojistik Regresyon yerine çok katmanlı sinir ağı mimarisi önerilmiştir. Çok katmanlı sinir ağı mimarisi 4 katmandan oluşturulmuştur. İlk katmanda 128

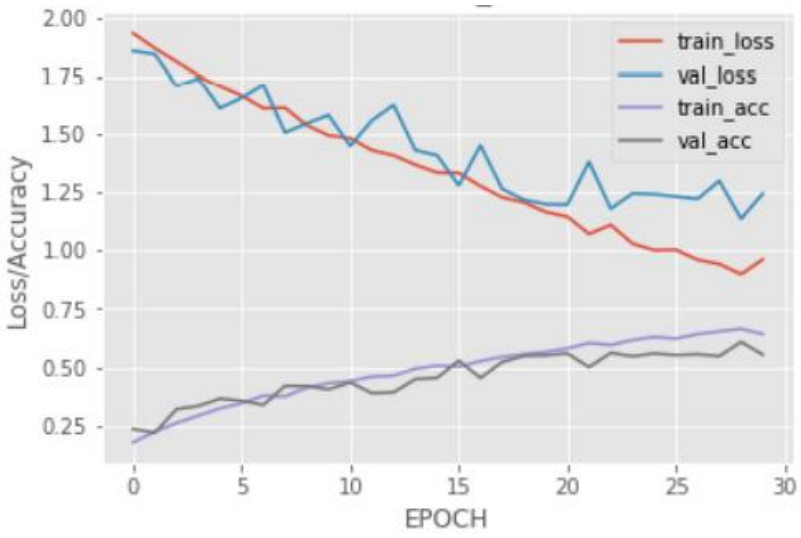
adet nöron bulunmakta olup aktivasyon fonksiyonu olarak relu seçilmiştir. Çok katmanlı sinir ağımızın ikinci katmanı ve üçüncü katmanında sırasıyla 64 ve 32 nörondan oluşmaktadır. Her iki katmanda aktivasyon fonksiyonu olarak relu kullanılmıştır. Son katmanda 7 nöron kullanılmış ve aktivasyon fonksiyonu olarak softmax tercih edilmiştir.

4 BULGULAR

Yapılan çalışmada insan eylemi tanıma için renkli resim görüntüleri iki adet yığma topluluk öğrenme mimarisine beslenmiştir. İlk mimaride temel öğrenici katmanında 4 farklı evrişimli sinir ağı kullanılmıştır. Meta öğrenici katmanında çoklu sınıflandırmayı desteklemek için lojistik regresyon yöntemi tercih edilmiştir. İlk mimarimizin doğruluk oranı 0.650 olarak çıkmıştır. İkinci mimaride ise 4 adet farklı katmanlardan oluşan CNN modeli tercih edilmiştir. Meta öğrenici katmanı için ise çok katmanlı sinir ağı modeli tercih edilmiştir. İkinci mimarimizde ise doğruluk oranı 0.743 ile daha yüksek bir başarım elde edilmiştir.

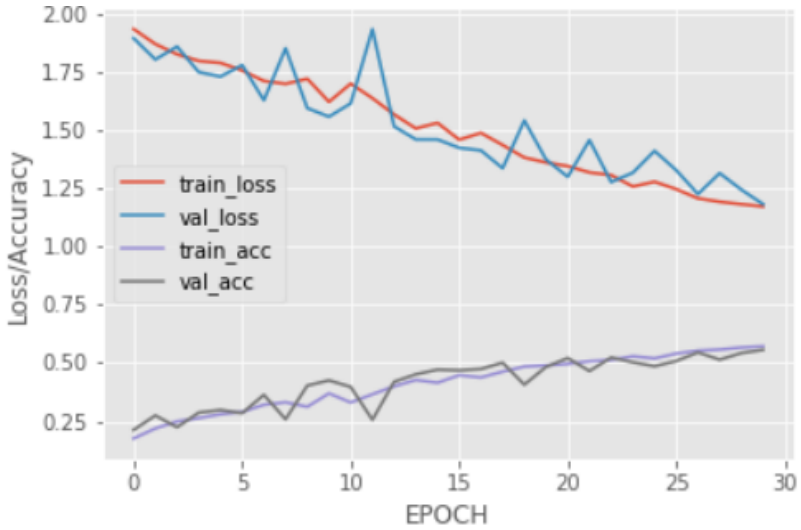
4.1 LOJİSTİK REGRESYON META ÖĞRENİCİYE SAHİP YIĞMA TOPLULUK ÖĞRENME MODELİ

İlk mimarimizin temel öğrenici katmanında 4 adet farklı parametrelere sahip CNN modeli bulunmaktadır. Temel öğrenici katmanlarından ilki olan Model A'ya ait eğitim doğruluğu ve eğitim kaybı (Train Accuracy – TA, Train Loss - TL), doğrulama kaybı ve doğruluk değerleri (Validation Loss – VL, Validation Accuracy - VA) Şekil 4.1'de gösterilmiştir.



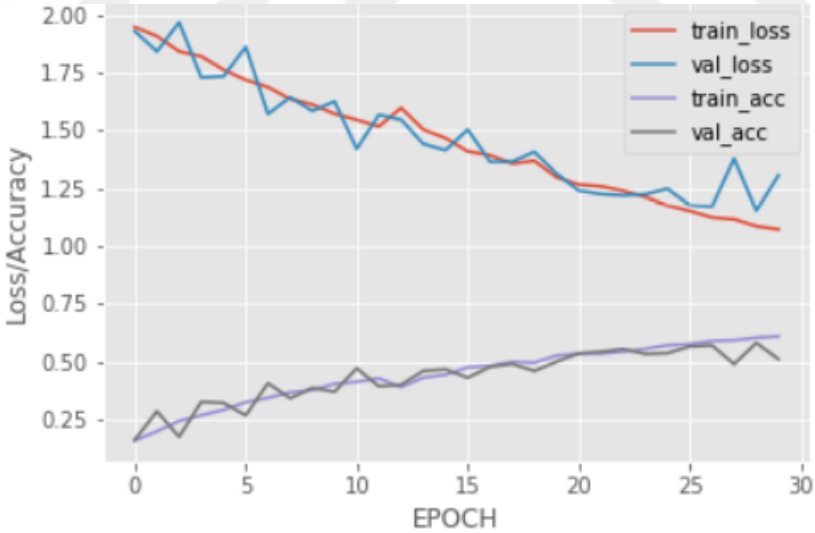
Şekil 4.1: Model A Doğruluk ve Kayıp Değerleri Grafiği.

Model A'nın VL değeri 1.136 olarak bulunmuştur. VA değeri ise 0.609 olarak bulunmuştur. İkinci temel öğrenici modelimizde ise veri seti ile eğitim gerçekleştirildikten sonra Şekil 4.2'de gösterilen grafik elde edilmiştir.



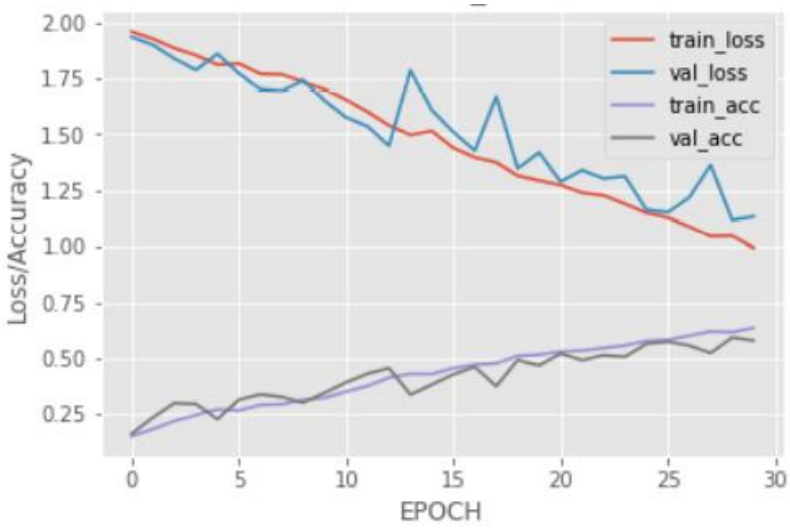
Şekil 4.2: Model B Doğruluk ve Kayıp Değerleri Grafiği.

Model B'nin VL değeri 1.181 olarak bulunmuştur. VA değeri ise 0.556 olarak bulunmuştur. Diğer bir temel öğrenici model olarak kullanılan Model C, 4 katmanlı bir CNN mimarisine sahiptir. Şekil 4.3'te doğruluk ve kayıp değerleri grafiği verilmiştir. Model C'nin VL değeri 1.152 olarak bulunmuştur. VA değeri ise 0.588 olarak bulunmuştur.



Şekil 4.3: Model C Doğruluk ve Kayıp Değerleri Grafiği.

Temel öğrenici model olarak kullanılan son CNN mimarisinin doğruluk ve kayıp değerleri Şekil 4.4'te verilmiştir.



Şekil 4.4: Model D Doğruluk ve Kayıp Değerleri Grafiği.

4 adet CNN modeli temel öğrenici olarak insan eylemi içeren görüntüler üzerinde eğitilmiştir. Elde edilen tahminler meta öğrenici model olan lojistik regresyonun eğitimi için kullanılmıştır. Meta öğrenici modelimizin VL değeri 1.090 olarak bulunmuştur. VA değerimiz ise 0.650 olarak bulunmuştur. Yığma topluluk öğrenme modelimizin ayrı ayrı eğitilen 4 farklı CNN modelimizden daha yüksek bir başarı gösterdiği Tablo 4.1'de ortaya konulmuştur.

Tablo 4.1: A, B, C, D ve Topluluk Öğrenme Modellerine ait VA ve VL Oranları.

Modeller	*VL(Validation Loss)*	*VA(Validation Accuracy)*
Model A	1.1362	0.609
Model B	1.181	0.556
Model C	1.152	0.581
Model D	1.119	0.593
Topluluk Öğrenme	**1.090**	**0.650**

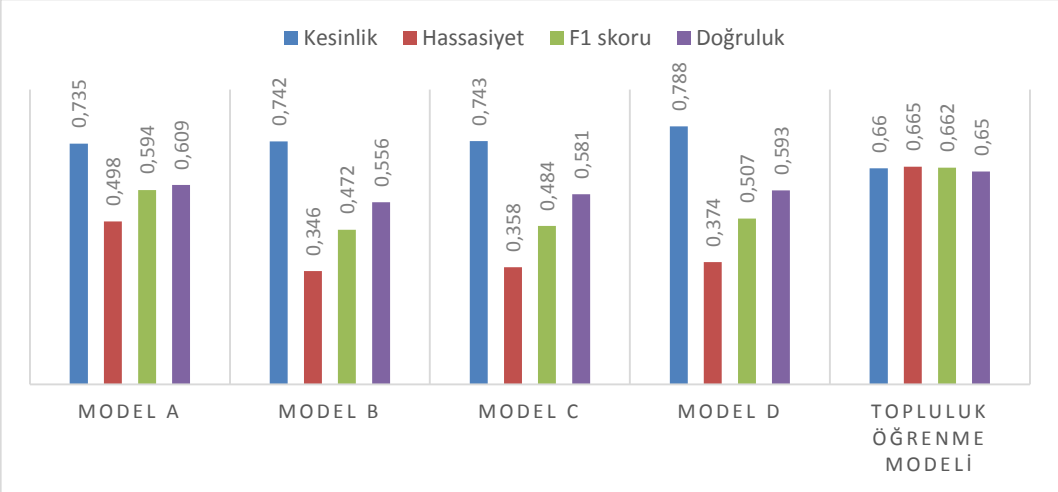
Tablo 4.1'de görüldüğü gibi Yığma topluluk öğrenme modelimiz Model A'ya göre yaklaşık %5 performans iyileştirmesi sağlamıştır. Model B'ye göre ise %10'a yakın daha yüksek bir başarım oranı göstermiştir.

Modelimiz pozitif değerleri tahmin etmede ne kadar başarılı olabildiğini 0.838 kesinlik değeri ile göstermiştir. Modelimizin F1 değeri 0.781 ile ne kadar güçlü olduğunu belirli ölçüde ortaya koymuştur. Meta öğrenici katmanında çok katmanlı sinir ağının bulunduğu topluluk öğrenme mimarimize ait performans metrikleri Tablo 4.2'de gösterilmiştir.

Tablo 4.2: Lojistik Regresyon Meta Öğreniciye Sahip Topluluk Öğrenme Modeli Performans Metrikleri.

Performans Metriği	**Değer (%)**
Doğruluk (Accuracy)	0.650
Kesinlik (Precision)	0.660
Hassasiyet (Recall)	0.665
F1 Değeri	0.662

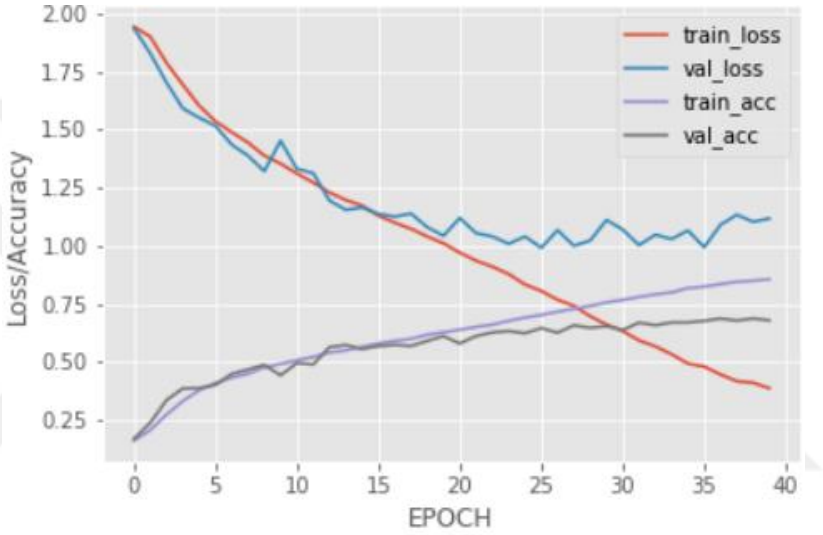
Yığma topluluk öğrenme modelimize ait temel öğrenici katmanında yer alan 4 farklı evrişimli sinir ağına ait kesinlik, hassasiyet, F1 ve doğruluk değerleri ile yığma topluluk öğrenme modelimize ait değerler Şekil 4.5'te gösterilmiştir. Topluluk öğrenme modelimiz 0,665 ile en yüksek hassasiyet değerini elde etmiştir. Ayrıca 0.662 ile en yüksek F1 değerini elde etmiştir.



Şekil 4.5: Lojistik Regresyon Meta Öğreniciye Sahip Topluluk Öğrenme Modeli Sınıflandırıcıları.

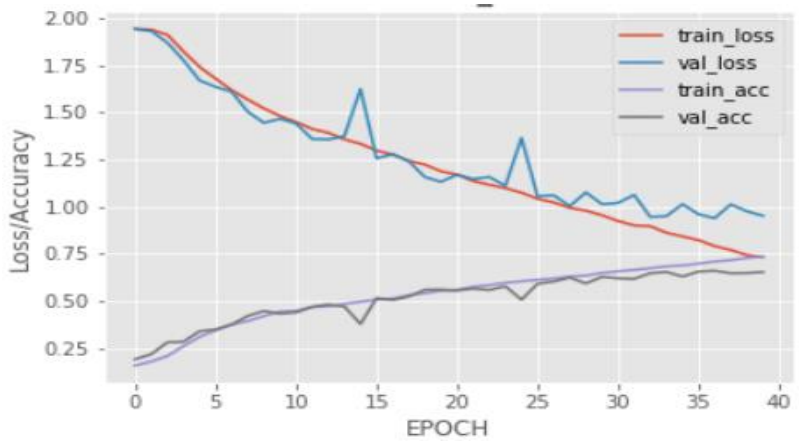
4.2 ÇOK KATMANLI SİNİR AĞI META ÖĞRENİCİYE SAHİP YIĞMA TOPLULUK ÖĞRENME MODELİ

İkinci mimari olarak kullandığımız yığma topluluk öğrenme modelimiz temel öğrenici bölümünde 4 adet evrişimli sinir ağı modeli bulunmaktadır. Meta öğrenici kısmında ise bir adet çok katmanlı sinir ağı yer almaktadır. Meta öğrenici katmanında bulunan ilk evrişimli sinir ağı modelimiz Şekil 4.6'da doğruluk ve kayıp değerleri grafiği ile gösterilmiştir. Model I'e ait VL değeri 0.991 olarak bulunmuştur. VA değeri ise 0.646 olarak bulunmuştur.



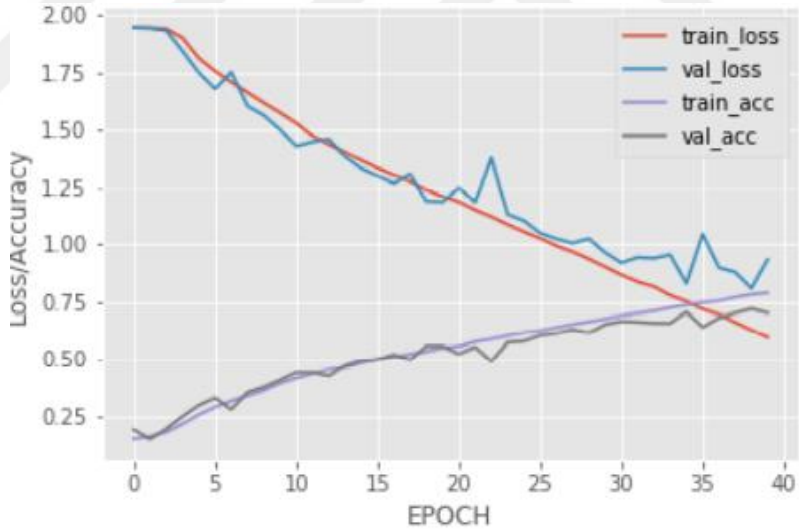
Şekil 4.6: Model I Doğruluk ve Kayıp Değerleri Grafiği.

Yığma topluluk öğrenme modelimizin ikinci katmanında yer alan Model II'ye ait doğruluk ve kayıp oranları şekil 4.7'de gösterilmiştir. 5 katmanlı evrişimli sinir ağımız VL değeri olarak 0.938 elde etmiştir. VA değeri olarak 0.660 elde edilmiştir.



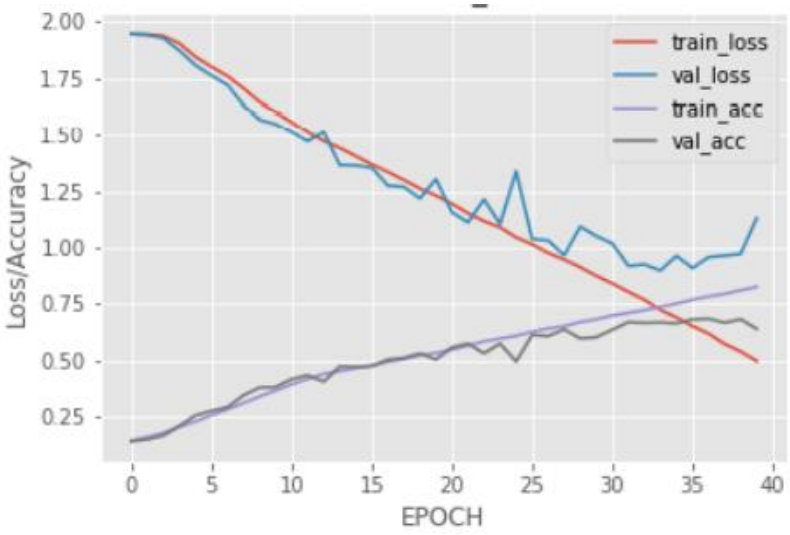
Şekil 4.7: Model II Doğruluk ve Kayıp Değerleri Grafiği.

Yığma topluluk öğrenme modelimizin temel öğrenici modellerinden üçüncü olarak 5 katmanlı evrişimli sinir ağı kullanılmıştır. Model III'e ait doğruluk ve kayıp değerleri Şekil 4.8'de gösterilmiştir. Temel öğrenici olarak kullanılan model III'te, VL değeri olarak 0.806 elde edilmiştir. Model III'e ait VA değeri, 0.719 elde edilmiştir.



Şekil 4.8: Model III Doğruluk ve Kayıp Değerleri Grafiği.

Yığma topluluk öğrenme modelimizin temel öğrenici katmanında bulunan son evrişimli sinir ağımız 6 katmandan inşa edilmiştir. Temel öğrenicilerden sonuncusu olan Model IV için, VL değeri 0.835 elde edilmiştir. Model IV'e ait VA değeri ise 0.669 olarak bulunmuştur. Model IV'e ait doğruluk ve kayıp sayılarına ait grafik Şekil 4.9'da gösterilmiştir.



Şekil 4.9: Model IV Doğruluk ve Kayıp Değerleri Grafiği.

Topluluk öğrenme modelimizde 4 adet farklı katmanları olan evrişimli sinir ağı modeli temel öğrenici olarak insan eylemi içeren görüntüler üzerinde eğitilmiştir. Elde edilen tahminler meta öğrenici model olan çok katmanlı sinir ağı eğitimi için kullanılmıştır. Meta öğrenici modelimizin VL değeri 0.594 olarak bulunmuştur. VA değerimiz ise 0.743 olarak bulunmuştur. Yığma topluluk öğrenme modelimizin, ayrı ayrı eğitilen 4 farklı CNN modelimizden daha yüksek bir başarı gösterdiği Tablo 4.3'te ortaya konulmuştur.

Tablo 4.3: I, II, III, IV ve Topluluk Öğrenme Modellerine ait VA ve VL Oranları.

Modeller	*Validasyon Kaybı*	*Validasyon Değeri*
Model I	0.991	0.646
Model II	0.938	0.660
Model III	0.806	0.719
Model IV	0.8354	0.669
Topluluk Öğrenme	**0.594**	**0.743**

Tablo 4.3'te görüldüğü üzere yığma topluluk öğrenme modelimiz Model I'e göre yaklaşık %10 performans iyileştirmesi sağlamıştır. Model III'e göre ise %3'e yakın daha iyi bir başarım oranı göstermiştir.

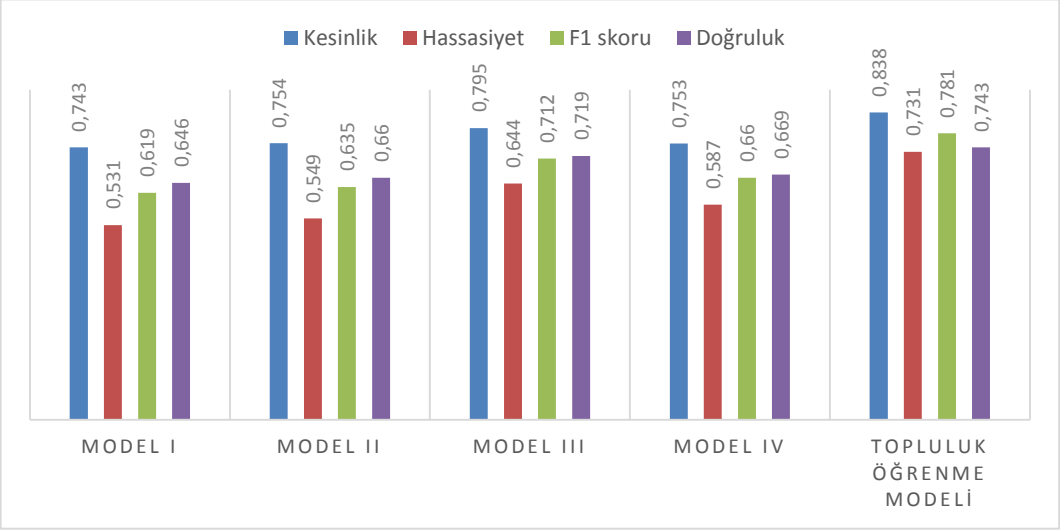
Yapılan çalışmada iki farklı yığma topluluk öğrenme modeli test edilmiştir. İlk modelimizin meta öğrenici katmanında lojistik regresyon kullanılmış ve VA değeri olarak 0.650 bulunmuştur. İkinci modelimizin meta öğrenici katmanında çok katmanlı sinir ağı kullanılmış ve VA değeri olarak 0.743 bulunmuştur.

Modelimiz pozitif değerleri tahmin etmede ne kadar başarılı olabildiğini 0.838 kesinlik değeri ile göstermiştir. Böylece sezgisel olarak sınıflandırıcının negatif olan bir örneği pozitif olarak etiketleme yeteneği 0.838 olduğu bulunmuştur. Modelimizin F1 değeri 0.781 ile ne kadar güçlü olduğunu belirli ölçüde ortaya koymuştur. Meta öğrenici katmanında çok katmanlı sinir ağının bulunduğu topluluk öğrenme mimarimize ait performans metrikleri Tablo 4.4'te gösterilmiştir.

Tablo 4.4: Çok Katmanlı Meta Öğrenici Modele Sahip Yığma Topluluk Öğrenme Modeli Performans Metrikleri.

Performans Metriği	**Değer (%)**
Doğruluk (Accuracy)	0.743
Kesinlik (Precision)	0.838
Hassasiyet (Recall)	0.731
F1 Değeri	0.781
AUC	0.853

Yığma topluluk öğrenme modelimize ait temel öğrenici katmanında yer alan 4 farklı evrişimli sinir ağına ait kesinlik, hassasiyet, F1 ve doğruluk değerleri ile yığma topluluk öğrenme modelimize ait değerler şekil 4.10'da gösterilmiştir. Topluluk öğrenme modelimiz 0.838 ile en yüksek kesinlik değerini elde etmiştir. Ayrıca 0.781 ile en yüksek F1 değerini elde etmiştir.



Şekil 4.10: Çok Katmanlı Sinir Ağı Meta Öğreniciye Sahip Topluluk Öğrenme Modeli Sınıflandırıcıları.

Çok katmanlı sinir ağı meta öğreniciye sahip yığma topluluk öğrenme modelimiz 0.731 Hassasiyet değeri elde etmiştir. Lojistik regresyon meta öğrenici modeline sahip yığma topluluk öğrenme modelimiz ise 0.665 hassasiyet değeri elde etmiştir. İki mimarinin sırasıyla F1 değerleri 0.781 ve 0.662 olarak elde edilmiştir.

5 TARTIŞMA VE SONUÇ

Yapılan çalışmada insan eylemi tanıma için derin öğrenme modelleri kullanılmıştır. İnsan eylemleri tanıma için yayınlanmış açık kaynak erişimli Action-Net isimli veri seti üzerinde derin öğrenme modelleri ile testler gerçekleştirilmiştir. Action-Net, insan eylemi tanıma çalışmalara katkı sunmak amacıyla insan eylem, jest ve etkinliklerini anlamak üzere toplanmış bir veri kümesidir. Action-Net içerisinde 16 sınıf barındırmakta, 19.200 adet 3 kanallı renkli resimler yer almaktadır. Her sınıf için yaklaşık 1000 resim eğitim verisi için ayrılmış 200 resim ise test verisi için ayılmıştır. Resimler gerçek hayat görüntülerinden alınmıştır. Temelde iki ayrı yığma topluluk öğrenme modeli karşılaştırılmıştır. Ayrıca farklı katmanlara sahip evrişimli sinir ağı modelleri test edilmiş ve performans metrikleri karşılaştırması yapılmıştır.

Öncelikle insan eylemi tanıma için oluşturulmuş Action-Net isimli veri seti farklı CNN modelleri ile eğitilmiş ve modellerin test sonuçları elde edilmiştir. CNN modelleri katman sayıları her katmanda bulunan filtre sayıları değiştirilerek karşılaştırılmıştır. Elde edilen hassasiyet, kesinlik, F1 ve doğruluk değerleri analiz edilmiştir.

Veri seti ile yapılan bir diğer çalışmada iki farklı yığma topluluk öğrenme modeli oluşturularak performans metrikleri üzerinde karşılaştırmalı analizler yapılmış ve sonuçlar değerlendirilmiştir. İlk yapılan yığma topluluk öğrenme modeli iki aşamadan meydana gelmektedir. Bu aşamalar sırasıyla temel öğreniciler katmanı ve meta öğrenici katmanıdır. Temel öğrenici katmanında 4 adet CNN modeli veri seti üzerinde eğitilmiştir. Temel öğrenicilerin yaptığı tahminler bir sonraki aşamada yer alan meta öğrenici katmana beslenerek öğrenme gerçekleştirilmiştir. Meta öğrenici katmanda lojistik regresyon yöntemi tercih edilmiş çoklu sınıflandırma yapılmıştır.

Veri seti ile yapılan sonraki çalışma ise bir diğer yığma topluluk öğrenme modeli ile gerçekleştirilmiştir. Bir önceki topluluk öğrenme modelinden farklı olarak meta öğrenici katmanda çok katmanlı sinir ağı yöntemi tercih edilmiştir. Yaptığımız deneylerde yığma topluluk öğrenme modelimizin ilk aşamasında 4 farklı CNN modeli veri seti ile eğitim gerçekleştirilmiştir. Elde edilen tahminler bir sonraki meta öğrenici katmana beslenerek çok katmanlı sinir ağı modeli ile öğrenme gerçekleştirilmiştir.

Yapılan deneylerin sonucunda insan eylemi tanıma için bir adet CNN modelinden farklı olarak yığma topluluk öğrenme modelinin başarılı olduğu gösterilmiştir. Hem temel öğrenici katmanında hem de meta öğrenici katmanında derin öğrenme yöntemlerinin kullanılması performans metriklerinde iyileştirmelere yol açtığı değerlendirilmektedir.

Gelecekte insan eylemi tanıma alanında yapılacak çalışmalarda topluluk öğrenme modeli içerisinde yer alan temel öğreniciler için kullanılan derin öğrenme yöntemleri heterojen olarak kullanılması düşünülmektedir. Ayrıca meta öğrenici katmanında bir adet derin öğrenme yerine iki veya daha fazla kullanılmasının fayda sağlayacağı düşünülmektedir.

6 KAYNAKLAR

[1]. Aggarwal, J. K., & Ryoo, M. S., 2011, Human activity analysis: A review. ACM Computing Surveys (CSUR), 43(3), 1-43.

[2]. Duckworth, P., Alomari, M., Gatsoulis, Y., Hogg, D. C., & Cohn, A. G., 2016, Unsupervised activity recognition using latent semantic analysis on a mobile robot, 22nd European Conference on Artificial Intelligence, 29 Aug - 02 Sep 2016 The Hague, Netherlands, ISBN: 978-1-61499-671-2, 1062-1070.

[3]. Jegham, I., & Khalifa, A. B., 2017, Pedestrian detection in poor weather conditions using moving camera, 14th International Conference on Computer Systems and Applications (AICCSA), 30 Oct.-3 Nov. 2017, Tunisia, IEEE, 358-362.

[4]. Chebli, K., & Khalifa, A. B., 2018, Pedestrian detection based on background compensation with block-matching algorithm, In 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD), 19-22 March 2018 Tunisia, IEEE, 497-501.

[5]. Zhang, J., Han, Y., Tang, J., Hu, Q., & Jiang, J., 2016, Semi-supervised image-to-video adaptation for video action recognition. IEEE transactions on cybernetics, 47(4), 960-973.

[6]. Gravina, R., Alinia, P., Ghasemzadeh, H., & Fortino, G., 2017, Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. Information Fusion, 35, 68-80.

[7]. Wang, S., & Zhou, G., 2015, A review on radio based activity recognition. Digital Communications and Networks, 1(1), 20-29.

[8]. Cheok, M. J., Omar, Z., & Jaward, M. H., 2019, A review of hand gesture and sign language recognition techniques. International Journal of Machine Learning and Cybernetics, 10(1), 131-153.

[9]. Rahmani, H., Mian, A., & Shah, M., 2017, Learning a deep model for human action recognition from novel viewpoints. IEEE transactions on pattern analysis and machine intelligence, 40(3), 667-681.

[10]. Anitha, G., & Priya, S. B., 2019, Posture based health monitoring and unusual behavior recognition system for elderly using dynamic Bayesian network, Cluster Computing, 22(6), 13583-13590.

[11]. Zhang, S., Wei, Z., Nie, J., Huang, L., Wang, S., & Li, Z., 2017, A review on human activity recognition using vision-based method, Journal of healthcare engineering, 2017, 31.

[12]. Nunez, J. C., Cabido, R., Pantrigo, J. J., Montemayor, A. S., & Velez, J. F., 2018, Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition. Pattern Recognition, 76, 80-94.

[13]. Bishop, C. M., 2006, Pattern recognition. Machine learning, 128(9).

[14]. Kotsiantis, S. B., Zaharakis, I., & Pintelas, P., 2007, Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160(1), 3-24.

[15]. Molinaro, A. M., Simon, R., & Pfeiffer, R. M., 2005, Prediction error estimation: a comparison of resampling methods. Bioinformatics, 21(15), 3301-3307.

[16]. Friedman, J., Hastie, T., & Tibshirani, R., 2001, The elements of statistical learning, 1(10), New York: Springer series in statistics.

[17]. Sutton, R. S., & Barto, A. G., 1998, Introduction to reinforcement learning (Vol. 135). Cambridge: MIT press.

[18]. Hebb, D. O., 1949, The organization of behavior; a neuropsycholocigal theory. A Wiley Book in Clinical Psychology, 62, 78.

[19]. Krizhevsky, A., Sutskever, I., & Hinton, G. E., 2012, Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25, 1097-1105.

[20]. LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D., 1989, Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), 541-551.

[21]. Kiros, R., Zhu, Y., Salakhutdinov, R. R., Zemel, R., Urtasun, R., Torralba, A., & Fidler, S., 2015, Skip-thought vectors, In Advances in neural information processing systems, 3294-3302.

[22]. Pascanu, R., Mikolov, T., & Bengio, Y., 2013, On the difficulty of training recurrent neural networks, Proceedings of the 30th International Conference on Machine Learning, 17-19 June 2013, Atlanta, Georgia, USA, 1310-1318.

[23]. Polikar R., 2012, Ensemble Learning. In: Zhang C., Ma Y. (eds) Ensemble Machine Learning. Springer, Boston, MA, ISBN: 978-1-4419-9325-0, 1-34.

[24]. Dietterich, T. G., 2002, Ensemble learning. The handbook of brain theory and neural networks, 2, 110-125.

[25]. Vatsavai, R. R., Shekhar, S., & Bhaduri, B., 2008, A semi-supervised learning algorithm for recognizing sub-classes. In 2008 IEEE International Conference on Data Mining Workshops, 15-19 Dec. 2008 Italy, IEEE, 458-467.

[26]. Krogh, A., & Vedelsby, J., 1995, Neural network ensembles, cross validation, and active learning, Advances in neural information processing systems, 7, 231-238.

[27]. Breiman, L., 1996, Bagging predictors. Machine learning, 24(2), 123-140.

[28]. Mendes-Moreira, J., Soares, C., Jorge, A. M., & Sousa, J. F. D., 2012, Ensemble approaches for regression: A survey. Acm computing surveys (csur), 45(1), 1-40.

[29]. Freund, Y., & Schapire, R. E., 1997, A decision-theoretic generalization of on-line learning and an application to boosting. Journal of computer and system sciences, 55(1), 119-139.

[30]. Zhou, Z. H., 2012, Ensemble methods: foundations and algorithms, Chapman & Hall/CRC, ISBN: 978-1-439-830031.

[31]. Bruce, P., Bruce, A., & Gedeck, P., 2020, Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python, Tokyo, O'Reilly Media.

[32]. Olafenwa, M., 2019, *Action-Net*, GitHub - OlafenwaMoses/Action-Net: A dataset of images for artificial intelligence models to recognize human actions., [Ziyaret Tarihi: 20 Ocak 2021].

[33]. Pernía-Espinoza, A., Fernández-Ceniceros, J., Antonanzas, J., Urraca, R., & Martinez-de-Pison, F. J., 2018, Stacking ensemble with parsimonious base models to improve generalization capability in the characterization of steel bolted components. Applied Soft Computing, 70, 737-750.