



# Yapay Öğrenme ile Farklı Akıllı Ulaşım Senaryoları Altında Araçtan Her Şeye Haberleşme Standardı Seçimi

Hakan Alp Eren<sup>1\*</sup>, Nihat Adar<sup>2</sup>, Ahmet Yazar<sup>3</sup>

<sup>1,3</sup> Eskişehir Osmangazi Üniversitesi, Yazılım Mühendisliği Bölümü, Eskişehir, Türkiye

<sup>2</sup> Eskişehir Osmangazi Üniversitesi, Bilgisayar Mühendisliği Bölümü, Eskişehir, Türkiye

erenhalp@gmail.com, nadar@ogu.edu.tr, ahmet.yazar@ogu.edu.tr

## Öz

Akıllı ulaşım sistemlerine yönelik çalışmaların son yıllarda artmasıyla birlikte araçtan her şeye (V2X) haberleşme konsepti için farklı standartların geliştirilmesi önem kazanmıştır. Bu doğrultuda 5. Nesil (5G) haberleşmesine yön veren 3GPP ve Wi-Fi haberleşmesine yön veren IEEE gibi organizasyonlar farklı V2X standartları geliştirmişlerdir. Farklı senaryolarda bu iki kritik standardın birbirlerine karşı üstünlükleri olabileceğini gösteren çeşitli çalışmalar bulunmaktadır. Önerilen yöntem ile birlikte farklı şartlar altında 3GPP ve IEEE standartlarından hangisinin kullanılmasının daha avantajlı olacağı yapay öğrenme teknikleri ile belirlenmekte ve uygun V2X standardı otomatik olarak seçilmektedir. Bu kapsamda araçta ve çevre sistemlerinde her iki standartla ilişkili donanımların bulunduğu varsayılmaktadır. Bu amaca yönelik yeni bir yapay veri seti oluşturulmuş ve K-en yakın komşu, karar ağacı, yapay sinir ağı ile TabNet sınıflandırıcıları kullanılarak çeşitli yapay öğrenme modelleri eğitilmiştir. Ayrıca çapraz doğrulama ile hiperparametre optimizasyonu gerçekleştirilmiştir. TabNet sınıflandırıcısı ile doğruluk değeri ve ağırlıklı F1 skoru 0.88 olarak elde edilmiştir. Tüm bu çalışmalar beraber ele alındığında, V2X haberleşmesine yönelik özgün bir çalışma yapılarak literatüre önemli bir katkı sağlandığı görülmüştür. Geliştirilen yapay öğrenme tabanlı V2X standardı seçirme yönteminin akıllı ulaşım sistemleri altındaki araçlara entegre edilebileceği düşünülmektedir.

**Anahtar kelimeler:** 5G, Akıllı Ulaşım, Araçtan Her Şeye Haberleşme, TabNet, Yapay Öğrenme.

## Vehicle-to-Everything Communications Standard Selection Under Different Intelligent Transportation Scenarios with Artificial Learning

### Abstract

It has become important to develop different standards for vehicle-to-everything (V2X) communications concept with the increase in studies on intelligent transportation systems in recent years. In this direction, organizations such as 3GPP, which leads 5th Generation (5G) communications, and IEEE, which leads Wi-Fi communications, have developed different V2X standards. There are various studies showing that these two critical standards may have advantages over each other in different scenarios. With the proposed method, which of 3GPP and IEEE standards will be more advantageous under different conditions is determined by artificial learning techniques and appropriate V2X standard is selected automatically. In this context, it is assumed that there is related equipment for each of the two standards in the vehicle and its environmental systems. For this purpose, a new artificial dataset was created, and various artificial learning models were trained using K-nearest neighbor, decision tree, artificial neural network and TabNet classifiers. In addition, hyperparameter optimization was performed with cross validation. With the TabNet classifier, the accuracy value and the weighted F1 score were obtained as 0.88. When all these studies are considered together, it has been seen that a significant contribution to the literature has been made by conducting a novel study on V2X communications. It is thought that the developed artificial learning based V2X standard selection method can be integrated into vehicles under intelligent transportation systems.

**Keywords:** 5G, Intelligent Transportation, Vehicle-to-Everything Communications, TabNet, Artificial Learning.

\* Sorumlu yazar.  
E-posta adresi: erenhalp@gmail.com

Alındı : 21 Ekim 2022  
Revizyon : 5 Şubat 2023  
Kabul : 23 Şubat 2023

## 1. Giriş (Introduction)

Nesnelerin interneti (IoT) ve akıllı sistemler gibi kavramların 5. Nesil (5G) hücreli haberleşme sistemleri ile birlikte uygulamada daha çok yaygınlaşacak olmasıyla, akıllı ulaşım ve araçtan her şeye (V2X) haberleşme sistemleri de giderek daha fazla önem kazanmaya başlamıştır. 2010'lu yılların başında standartlaştırma çalışmaları hız kazanan V2X sistemleri günümüzde kritik bir haberleşme altyapısına dönüşmüştür (Filippi vd., 2017). 2030'lu yıllarda kullanılmaya başlanacağı öngörülen 6. Nesil (6G) kablosuz haberleşme sistemlerinde ise V2X konseptinin rolünün çok daha büyük olacağı düşünülmektedir (Yazar vd., 2020; Yazar, 2021). Tüm bu gelişmelere bakıldığında, V2X haberleşme standartlarına yönelik yapılan çalışmaların kritik öneme sahip olduğu değerlendirilmektedir.

Yaygın olarak kullanılmaya başlanan V2X standartları incelendiğinde, 5G haberleşmesine yön veren 3GPP ve Wi-Fi haberleşmesine yön veren IEEE organizasyonları tarafından iki farklı standardın geliştirildiği görülmektedir (MacHardy vd., 2018). Sırasıyla C-V2X ve IEEE 802.11p standartları çerçevesinde sürekli olarak yeni iyileştirmeler yapılmaktadır. Bu çerçevede iki ana standardın birbirlerine karşı farklı üstünlükleri oluşabilmektedir (Naik vd., 2019). Mevcut literatüre bakıldığında bu üstünlüklerin incelendiği çalışmalara rastlanmaktadır (Zhao vd., 2018; Ullah vd., 2019; Naik vd., 2019; Shrestha vd., 2020). Bununla birlikte, hangi standardın hangi şartlar altında kullanılmasının daha çok fayda sağlayacağını otomatik belirlenmesine yönelik yöntem geliştirme çalışmalarının bulunmadığı görülmektedir. Örneğin, bir akıllı araçta ve haberleşeceği çevre sistemlerinde her iki standartla ilişkili donanımların bulunduğu varsayımda, anlık olarak hangi standart ile haberleşilmesi gerektiğinin kararına ihtiyaç duyulacaktır.

Yukarıda bahsedilen V2X haberleşme standartlarının farklı özellikler açısından güçlü ve zayıf yönleri bulunduğundan dolayı, bu standartlar arasından kullanılacak en uygun olanın tespiti için haberleşme gereksinimlerinin ve sistem özelliklerinin analiz edilmesi gerekmektedir. Sonrasında bu analiz temel alınarak hangi senaryolar altında hangi standardın kullanılmasının daha avantajlı olacağını kararı verdirilebilecektir. Bu çalışmada önerilen yöntem ile birlikte bu karar verme sürecine yapay öğrenme modellerinin dahil edilmesi sağlanarak uygun V2X haberleşme standardının bulunmasının zaman ve insan gücü bakımından verimli hale getirilmesi sağlanmıştır. Bu yöntemeye dayalı olarak geliştirilen uygulamanın akıllı araç üreticilerine, son kullanıcılara ve akıllı ulaşım alanında çalışan araştırmacılara faydalı olacağı düşünülmektedir. Geliştirilen uygulama akıllı araç üreticileri tarafından daha verimli çalışan bir V2X haberleşme sistemi oluşturulması noktasında fayda

getirirken, araç sahiplerinin ise bu kapsamda daha iyi bir hizmet alabilmesi sağlanacaktır. Ayrıca araç üreticileri kendi çözümlerini bu çalışmada geliştirilen yöntem üzerine kurgulayabileceklerdir.

Makalenin 2. bölümünde V2X haberleşme sistemleri ile ilgili olarak yapılan kaynak araştırmasına yer verilmiştir. 3. bölüm kapsamında veri seti ve kullanılan yapay öğrenme modelleri tanıtılmıştır. Modellerden elde edilen sonuçlara ait performans metrikleri 4. bölümde sunulmuştur. Son olarak 5. bölümde sonuç ve tartışmalar yer almaktadır.

## 2. Kaynak Araştırması (Literature Survey)

Literatürdeki benzer çalışmalar incelendiğinde hizmet kalitesi tahmini, araç haberleşme ağlarında hatalı davranış tespiti, V2X ağlarında performans tahmini, 5G hücreli ağ analizi, gecikme tahmini, zaman-frekans analizi ile V2X teknolojisinin tespit edilmesi gibi konuların ele alındığı görülmektedir. Fakat bu çalışmaların hiçbirinde farklı V2X haberleşme standartları arasında seçim yaptırılmasına yönelik bir yöntemin geliştirilmediği görülmektedir.

Moreira vd. (2020) bir V2X haberleşme senaryosu üzerinde makine öğrenmesi ile bir paketin teslim edilmesindeki hizmet kalitesini tahmin etmeye çalışmışlardır. İşlemin istenen gecikme süresi içerisinde gerçekleştirilip gerçekleştirilemeyeceğini gösterecek şekilde ikili sınıflandırma yapılmaktadır. Simülasyon aracılığıyla elde edilen sonuçlar incelendiğinde hizmet kalitesi seviyesinin iki saniye önceden %85 güvenilirlikle tahmin edilebildiği sonucuna ulaşılmaktadır. Ancak çalışmada ele alınan senaryonun hücreli haberleşme üzerine olması nedeniyle sadece 3GPP standardına odaklanılmaktadır.

Abdellah vd. (2021) V2X trafiğinin performansını tahmin amacıyla 3GPP standardı altındaki 5G ağları için derin öğrenme modeli oluşturmuşlardır. Modelin aykırı değerlerden daha az etkilenmesi ve eğitimin iyileştirilmesi için M-tahmin ediciler kullanılarak yeni kayıp fonksiyonları önerilmektedir. Ortalama kare hatası (MSE) fonksiyonu ile karşılaştırıldığında önerilen fonksiyonlar ile sağlaştıran modelin daha başarılı sonuçlar sağladığı gözlemlenmektedir.

Gyawali ve Qian (2019), araç iletişim ağlarındaki hataların tespit edilmesini sağlayan bir sistem önermişlerdir. Veri seti oluşturmak için araç ağı ortamı kapsamlı olarak simüle edilmiştir. Farklı makine öğrenmesi modelleri ile yapılan testlerde yüksek kesinlik, duyarlılık ve F1 skoru değerlerine ulaşılmıştır.

Sevgican vd. (2020) 5G hücreli ağları analiz etmeyi sağlayan yeni bir sistem sunmuşlardır. Oluşturulan modeller ağ yükünü tahmin etme ve anomalileri tespit etme üzere iki farklı problemi çözmek için kullanılmaktadır. Ağ yükünü tahmin etmek için zaman serisi analizi ile doğrusal regresyon, uzun kısa süreli bellek (LSTM) ve özyinelemeli sinir ağı (RNN) modelleri eğitilmiş ve MSE azaltılmaya çalışılmıştır. İkinci problem olan anomali tespitinde ise doğrusal

regresyon ve ekstrem gradyan artırma (XGBoost) ile alıcı işlem karakteristiği (ROC) eğrisinin altında kalan alanın artırılması hedeflenmiştir. Model eğitimi için sentetik 5G hücresele ağ veri seti üretilmiştir. Deneyler sonucunda her iki problem için de diğer modellerin doğrusal regresyondan daha fazla başarı gösterdiği tespit edilmiştir.

Zhang vd. (2020) gecikmeye duyarlı V2X uygulamaları için gecikmenin tahmin edilmesini sağlayan bir çatı geliştirmişlerdir. Bu çalışmada kullanılan veriler, LTE tabanlı bir ağda çeşitli hareket senaryolarını içeren sürüşlerden gecikme değerlerinin ölçülmesi ile elde edilmiştir. Veriler işlenerek zaman içindeki gecikme eğilimini gösteren ve rastgele gürültü gibi davranan olmak üzere iki bileşene ayrılmıştır. Bu bileşenler için LSTM ve K-medoids algoritmalarına ek olarak Epanechnikov çekirdeği ve hareketli ortalama fonksiyonlarının birleşiminden oluşan bir yöntem önerilmiştir. Sonuçlar incelendiğinde önerilen yöntemin tahminlerdeki hatayı azaltmada etkili olduğu görülmüştür.

Zhang vd. (2019) araçtan altyapıya (V2I) kullanıcıları için haberleşme kapasitesini maksimize etmek amacıyla V2X hücresele haberleşmede uygun iletim modunu seçme ve kaynak tahsisini gerçekleştirme üzerine çalışmışlardır. Problem bir Markov karar süreci (MDP) modeli ile ifade edilmiş ve çözüm için derin pekiştirmeli öğrenme (DRL) kullanılmıştır. Simülasyon sonuçları, önerilen yöntemin gecikme ve güvenilirlik gereksinimleri doğrultusunda problemin optimizasyonu için etkili olduğunu doğrulamıştır.

Skiribou ve Elbahhar (2021), makine öğrenmesi ile V2X haberleşme sinyallerinin tanınması için anlık frekans (IF) öznelik vektörünü kullanmışlardır. Zaman-frekans analizinden elde edilen bu vektörün boyutu, tekil değer ayrıştırması (SVD) ile küçültülmüştür. Rastgele orman modeli ile ITS-G5 sinyalleri için %100 doğruluk elde edilmiştir. LTE-V2X ve NR-V2X sinyalleri için kesinlik, duyarlılık ve F1 skoru değerleri 0.99 olarak gözlemlenmiştir.

### 3. Materyal ve Yöntem (Material and Method)

#### 3.1. Veri seti (Dataset)

V2X haberleşmesi için 3GPP ve IEEE standartlarının haberleşme gereksinimleri açısından karşılaştırıldığı çalışmalar (Zhao vd., 2018; Ullah vd., 2019; Naik vd., 2019; Shrestha vd., 2020) incelendiğinde hem şebeke mimarisi hem de fiziksel katman özelliklerinden dolayı IEEE standardının iletim gecikmeleri noktasında daha başarılı olduğu görülmektedir. Buna karşılık, 3GPP standardının haberleşme kapasitesi ve kapsama alanı ölçütleri açısından daha ön plana çıktığı tespit edilmiştir. Yapılan çalışma kapsamında farklı performans ölçütleri ile akıllı ulaşım senaryoları arasında Tablo 1 üzerindeki ilk sütunda verilen çeşitli senaryo bilgileri üzerinden bağ kurulmuştur. Arka planda, farklı akıllı ulaşım

senaryolarındaki kablosuz haberleşme kanalının genel olarak değişimi incelenmiştir. Dolayısıyla kablosuz haberleşme kanalına doğrudan bakmadan, ona etki edebilecek karakteristik senaryo bilgileri temel alınarak çalışmaya dahil edilmiştir.

Literatür taraması sonucunda V2X haberleşmesinde uygun standardın belirlenmesine yönelik bir veri setinin mevcut olmadığı görülmektedir. Bu problemi çözmek için veri üretme amacıyla farklı yöntemler karşımıza çıkmaktadır. Benzer konularda çalışma yapmış araştırmacılar, amaçlarına yönelik verileri yapay olarak üretmişlerdir.

Yapay veri setinin oluşturulması amacıyla MATLAB platformu üzerinde bir komut dizisi hazırlanmıştır. Bu komut dizisi, Tablo 1 üzerinde verilen öznelikleri birbirleriyle ilişkili olabilecek şekilde rastgele olarak 1-10 değerleri arasında ve ölçüt sütununda verilen iki genel tanım arasında değişiklik gösterecek şekilde üretmektedir. Veri etiketleme için ayrı bir mekanizma oluşturulmuş, otomatik ve elle olacak şekilde hibrit bir etiketleme süreci sürdürülmüştür. Bu hibrit mekanizmanın otomatik çalışan kısmının ilk aşamasında öznelikler oluşturulmadan önce sınıf etiketleri üretilmektedir. Öznelikler ise bu sınıf etiketleri ile uyumlu ve rastgele olacak şekilde yukarıda bahsedilen komut dizisi aracılığıyla oluşturulmaktadır. Hibrit veri etiketleme mekanizmasının otomatik çalışan kısmının ikinci aşaması öznelikler oluşturulduktan sonra devreye girmektedir. Bu aşamadaki amaç, sınıf etiketleri ile ilişkili olarak rastgele üretilen özneliklere bakarak, bazı özel tanımlı kurallar dikkate alınarak gerçekleşmesi mümkün olmayan durumlara karşılık gelen örneklemelerin veri setinden çıkarılmasıdır. Otomatik çalışan kısmın devamında ise eğer belirsiz olabilecek durumlar tespit edilmişse, bu durumların elle düzeltilmesine imkân tanımak için uyarı oluşturulmaktadır. Nihai veri setine bakıldığında sınıf etiketi, verilen öznelik değerlerine göre, hangi haberleşme standardının daha iyi bir çözüm sunacağı noktasında ulaşılan kararı göstermektedir. 5957 adet 3GPP ve 4043 adet IEEE sınıfına ait örnek içeren veri setinde toplamda 10000 örnek bulunmaktadır.

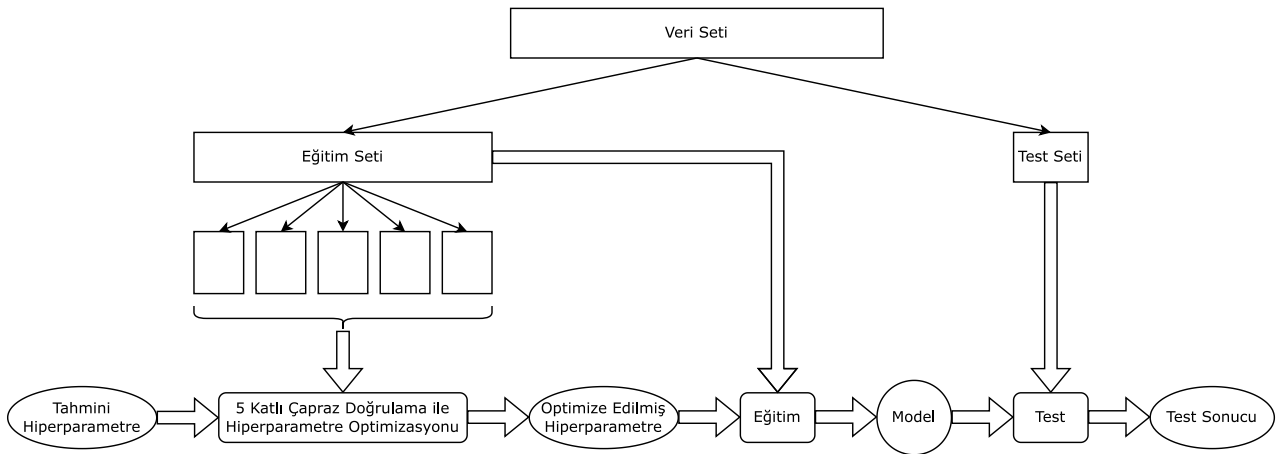
**Tablo 1.** Veri setindeki öznelikler ve sınıf etiketi (Features and class label in the dataset)

Öznelikler ve Sınıf Etiketleri	Ölçüt	Değer
Coğrafya özellikleri	Zorlu – Basit	1 – 10
Yerleşim özellikleri	Kentsel – Kırsal	1 – 10
Araç yoğunluğu	Yoğun – Seyrek	1 – 10
Yaya yoğunluğu	Yoğun – Seyrek	1 – 10
Veri hızı gereksinimi	Kritik – Normal	1 – 10
Güvenilirlik gereksinimi	Kritik – Normal	1 – 10
Gecikme gereksinimi	Kritik – Normal	1 – 10
Kapsama gereksinimi	Kritik – Normal	1 – 10
Standart	3GPP, IEEE	0, 1

### 3.2. Deney tasarımı (Experiment design)

Deney başlangıcında veri seti karıştırılarak 8000 adet eğitim ve 2000 adet test verisi olacak şekilde alt setlere ayrılmıştır. Ayırma işlemi tabakalı olarak gerçekleştirilerek orijinal veri setindeki sınıf etiketi dağılım oranının alt setlerde de korunması hedeflenmiştir. Sınıf etiketlerine ait örnek sayıları 3GPP için 5957 ve IEEE için 4043 olduğundan dolayı alt setlerde de yaklaşık olarak 6/4 oranı bulunmaktadır. Eğitim veri seti, çapraz doğrulamada kullanılmak üzere kendi içinde beş parçaya bölünmüştür. Bu aşamada da karıştırma ve tabakalı ayırma yapılarak dengesiz dağılımların önüne geçilmiştir. Çapraz doğrulama sonuçları karşılaştırılarak en başarılı sonuca sahip

hiperparametre değerleri tespit edilmiştir. Son aşamada modeller optimize edilmiş hiperparametre değerleri ve eğitim seti ile eğitilerek, test seti üzerinde sınanmıştır. Bu işlemler, Şekil 1'de verilen diyagramda özetlenmiştir. Veri setinin parçalara bölünmesi, sınıflandırma metriklerinin hesaplanması ve makine öğrenmesi modellerinin oluşturulmasında Scikit-Learn (Pedregosa vd., 2011) modülünden faydalanılmıştır. Ayrıca dizi ve veri çerçevesi yapıları üzerinde gerçekleştirilecek işlemler için NumPy (Harris vd., 2020) ve Pandas (McKinney, 2010) kütüphaneleri kullanılmıştır. Proje kapsamında kullanılan veri seti, deney kodları ve model dosyalarına <https://github.com/haeren/vehxcom> adresinden erişilebilmektedir.



Şekil 1. Deney tasarımı (Experiment design)

### 3.3. Sınıflandırmada kullanılan modeller (Models used for classification)

Deneylerde makine öğrenmesi ve derin öğrenme algoritmalarını kullanan dört adet sınıflandırma modeli yer almaktadır. İlk olarak en basit makine öğrenmesi algoritmalarından biri olan K-en yakın komşu kullanılmıştır. Ardından tablo halindeki verilerde başarılı olmasıyla bilinen karar ağacı modeli oluşturulmuştur. Son olarak yapay sinir ağı sınıflandırıcısı ve tablo şeklindeki veriler ile birlikte kullanılmak üzere tasarlanan bir derin öğrenme yöntemi olan TabNet modeline yer verilmiştir.

K-en yakın komşu algoritması, test edilen örnek ve eğitim kümesindeki veriler arasındaki uzaklıkları hesaplayarak sınıflandırma işlemini gerçekleştirmektedir.  $\mathbb{R}^q$ 'da uzaklık hesaplamak için Eşitlik (1)'de verilen Minkowski metriği tercih edilmektedir (Kramer, 2013). Eşitlikteki p değeri değiştirilerek sınıflandırma sırasında farklı türdeki uzunluklar kullanılabilir. Örneğin p = 1 için Manhattan ve p = 2 için Öklid uzaklıkları elde edilmektedir. Bu çalışmada oluşturulan modellerde Öklid uzunluğu ile sınıflandırma yapılmaktadır. Test edilen örneğin x olarak kabul edilmesi durumunda y, eğitim kümesindeki bir örneği temsil etmektedir. Bu iki

veri arasındaki uzaklık, q adet özneliğin mutlak farklarının toplanması ile hesaplanmaktadır.

$$\|x - y\|^p = \left( \sum_{i=1}^q |x_i - y_i|^p \right)^{1/p} \quad (1)$$

K-en yakın komşu algoritmasındaki en önemli hiperparametrelerden biri olan komşu sayısı, k ile ifade edilmektedir. Uzunluklar hesaplandıktan sonra test edilen örneğe en yakın k adet komşu belirlenmektedir. Belirlenen komşular içerisinde çoğunluğu oluşturan sınıf etiketi, tahmin sonucunu vermektedir. Dolayısıyla oylamada beraberliği önlemek için genellikle tek sayıda komşu seçilmesi önerilmektedir. Gerçekleştirilen deneylerde çapraz doğrulama ile modelin farklı sayıdaki komşular için başarısı gözlemlenmiştir.

Karar ağacı algoritması ile sınıflandırma, eğitim verilerinin alt kümelerine ayrılacak şekilde bölünmesi sonucunda oluşturulan ağaç yapısı üzerinden ilerleyerek gerçekleştirilmektedir. Bu yapı, veri setindeki herhangi bir özneliğin belirli bir koşulu sağlayıp sağlamadığına bağlı olarak dallanan düğümlerden oluşmaktadır. Bölünme sırasında hedef, sınıf etiketlerinin karışmasını minimize eden özneliği ve koşulu belirleyerek olabildiğince saf alt kümeler elde etmektir (Tangirala, 2020). Bu amaç doğrultusunda bölünmenin kalitesini

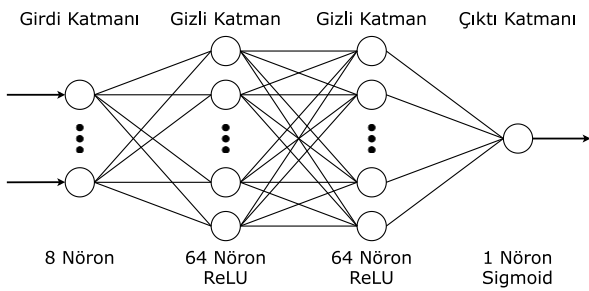
değerlendirebilmek için yaygın olarak kullanılan kriterler olan entropi ve Gini Eşitlik (2) ve (3)'te gösterilmektedir.

$$\text{Entropi} = -\sum_{i=1}^c p_i \log_2 p_i \quad (2)$$

$$\text{Gini} = 1 - \sum_{i=1}^c p_i^2 \quad (3)$$

Herhangi bir değişkenle ilişkili belirsizliğin ölçülmesinde entropi kullanılmaktadır. Belirsizlikle orantılı bir şekilde artan veya azalan entropinin minimum değeri 0 iken maksimum değeri sınıf sayısına bağlı olarak değişmektedir. Üretilen veri setine benzer bir şekilde iki sınıf bulunması durumunda entropinin değer aralığı [0, 1] olmaktadır. Eşitlik (2)'de  $p_i$ , veri setinden rastgele seçilen bir elemanın  $i$  sınıfına ait olma olasılığını ifade etmektedir (Rahmati vd., 2020). Gini indeksi ise safsızlığı ölçen bir öznitelik ölçüsüdür. Minimum Gini indeksine sahip öznitelik, bölme işleminde kullanılmaktadır (Gupta vd., 2017). Her iki kriter için de çapraz doğrulama yapılmış ve sonuçlar karşılaştırılmıştır.

Yapay sinir ağı, nöronlardan oluşan katmanların birbirlerine bağlanması ile elde edilen bir model mimarisidir. Eğitim boyunca nöronlardaki ağırlıkların güncellenerek modelin daha başarılı tahminler yapması hedeflenmektedir. Basit bir sinir ağı yapısında birer adet girdi ve çıktı katmanı ile bu katmanlar arasındaki gizli katmanlar yer almaktadır. Oluşturulan sinir ağına 8 nörona sahip girdi katmanı ve 1 nörona sahip çıktı katmanı bulunmaktadır. Ek olarak 64 nörondan oluşan iki adet gizli katman kullanılmıştır. Gizli katmanlara ReLU ve çıktı katmanına sigmoid aktivasyon fonksiyonu eklenmiştir. Sinir ağının mimarisi Şekil 2'de gösterilmektedir. Model, Keras API (Chollet, 2018) kullanılarak oluşturulmuştur.



Şekil 2. Yapay sinir ağı mimarisi (Artificial neural network architecture)

Ele alınan problemin ikili sınıflandırma içermesinden dolayı modelde kayıp fonksiyonu olarak ikili çapraz entropi kullanılmıştır. Eniyileştirici (optimizer) için Adam ve değerlendirme ölçütü için doğruluk tercih edilmiştir. Eğitimde maksimum evre sayısı (epoch) 100 seçilmiş ve erken durdurma mekanizması ile aşırı öğrenmenin engellenmesi hedeflenmiştir. Bu sayede doğruluk değerinin 15 evre boyunca iyileşmemesi sonucunda eğitim

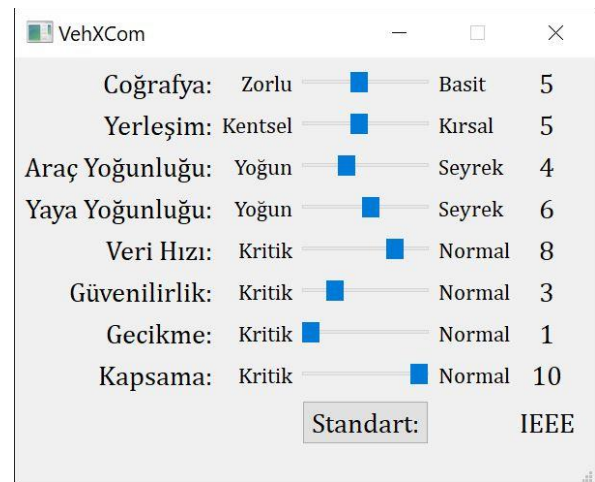
durdurulmaktadır. Farklı yığın boyutları (batch size) için çapraz doğrulama yapılarak daha başarılı sonuçlar veren boyutun bulunması hedeflenmiştir.

TabNet (Arik ve Pfister, 2021), tablo şeklindeki veriler için önerilmiş bir derin öğrenme mimarisidir. Ardışık dikkat (sequential attention) mekanizması ile karar adımlarının her birinde kullanılacak öznitelikleri belirlemektedir. Öznitelik seçiminin örnek tabanlı olarak gerçekleştirilmesi, öğrenme kapasitesinin en dikkat çekici öznitelikler için kullanılmasını sağlamaktadır. Bu sayede yorumlanabilirlik ve daha başarılı bir öğrenme elde edilmektedir. Farklı veri setleri ile yapılan testlerde TabNet modellerinin, XGBoost (Chen vd., 2015) ve LightGBM (Ke vd., 2017) yöntemlerini geride bıraktığı görülmektedir (Arik ve Pfister, 2021).

Deneylerde kullanılan TabNet modeline ait parametrelerin değerleri, yapay sinir ağındaki değerlere benzer şekilde seçilmiştir. Kayıp fonksiyonu olarak çapraz entropi ve eniyileştirici olarak Adam kullanılan modelde değerlendirme ölçütü için doğruluk tercih edilmiştir. Maksimum evre sayısı 100 ve tahammül (patience) değeri 15 olarak belirlenerek farklı yığın boyutları için çapraz doğrulama gerçekleştirilmiştir.

### 3.4. Kullanıcı arayüzü tasarımı (User interface design)

Geliştirilen sistemin araçlara entegre edilmesi durumunda kullanıcıların da bilgilendirilmesi için minimum sayıda bileşen kullanılarak sade bir kullanıcı arayüzü tasarlanmıştır. Gerçekleştirilen eğitimler sonucunda elde edilen en başarılı model arka planda kullanılarak çeşitli V2X haberleşme senaryoları için en uygun haberleşme standardının bulunabilmesi hedeflenmektedir. Tasarlanan arayüz ile örnek bir senaryonun sınıflandırılması Şekil 3'te gösterilmektedir.



Şekil 3. Kullanıcı arayüzü ile sınıflandırma (Classification with user interface)

#### 4. Deneysel Sonuçlar (Experimental Results)

Modellerin farklı hiperparametre değerleri için çapraz doğrulama sonuçları incelenerek en yüksek başarıyı sağlayan değerler tespit edilmiştir. Test aşamasında doğruluğa ek olarak karışıklık matrisinden türeyen diğer metrikler olan kesinlik, duyarlılık ve F1 skoru değerleri gösterilmektedir. Modellerin performansları hem sınıf bazında hem de makro ortalama ve ağırlıklı ortalama olarak analiz edilmiştir.

K-en yakın komşu sınıflandırıcısı ile yapılan çapraz doğrulamada en başarılı sonuç  $K = 17$  komşu için elde edilmiştir. Tablo 2'deki sonuçlar ardışık olarak artan komşu sayılarına göre incelendiğinde tek sayıdaki komşular için daha başarılı olduğu gözlemlenmektedir. Buna ek olarak komşu sayısı artırıldığında 17 komşuya kadar olan tek ve çift sayıdaki komşuların kendi içlerinde doğruluk değerlerinin yükseldiği görülmektedir.

Tablo 6'da görülebileceği üzere K-en yakın komşu modeli, test setinde 1031 gerçek pozitif ve 575 gerçek negatif olmak üzere toplamda 1606 adet örneği doğru olarak sınıflandırarak 0.803 doğruluk değerine ulaşmıştır. Tablo 7'deki diğer metrikler sınıf bazında incelendiğinde 3GPP sınıfının daha yüksek değerlere sahip olduğu görülmektedir. 3GPP sınıfı için duyarlılık daha fazla iken IEEE sınıfı için ise kesinlik değerinin daha yüksek olması dikkat çekmektedir. IEEE sınıfını tespit edebilmenin önemli olduğu durumlarda bu modelin kullanılması tavsiye edilmemektedir.

**Tablo 2.** Farklı komşu sayıları için doğrulama sonuçları (Validation results for different numbers of neighbors)

Komşu	Doğruluk	Komşu	Doğruluk
1	0.753625	11	0.811125
2	0.751000	12	0.808250
3	0.784125	13	0.812375
4	0.778500	14	0.810625
5	0.797750	15	0.815750
6	0.797000	16	0.811750
7	0.802875	17	0.816250
8	0.799250	18	0.811375
9	0.804625	19	0.813125
10	0.804625	20	0.812375

Farklı kriterlerin kullanıldığı karar ağacı sınıflandırıcılarında birbirine oldukça yakın sonuçlar elde edilmiştir. Tablo 3'te verilen doğrulama sonuçlarına bakıldığında entropi kriteri ile bir miktar daha yüksek doğruluğa ulaşıldığı gözlemlenmektedir.

Karar ağacı modelinin test verisi üzerindeki sonuçlarına ait karışıklık matrisi Tablo 6'da gösterilmektedir. Test seti üzerinde 0.8325 doğruluk değerine sahip modelin kesinlik, duyarlılık ve F1 skoru değerleri Tablo 7'de detaylı olarak verilmektedir. Modelin 3GPP sınıfı için daha başarılı sonuçlar verdiği görülmektedir. K-en yakın komşu sınıflandırıcısından farklı olarak sınıf içinde birbirine benzer kesinlik ve duyarlılık değerleri elde edilmiştir.

**Tablo 3.** Farklı kriterlere sahip ağaçlar için doğrulama sonuçları (Validation results for trees with different criteria)

Kriter	Doğruluk
Gini	0.846375
Entropi	0.847625

Çapraz doğrulamada 128, 256, 512 ve 1024 olmak üzere farklı yığın boyutları ile eğitilen yapay sinir ağı modellerinin doğruluk değerleri Tablo 4'tedir. En başarılı modelin yığın boyutu için 128 seçildiğinde elde edildiği ve boyut arttırıldıkça başarının düştüğü gözlemlenmektedir.

Yapay sinir ağı sınıflandırıcısı ile test veri setindeki örneklerin sınıflandırılması sonucu elde edilen Tablo 6'daki karışıklık matrisi, 0.819 doğruluk değerine ulaşıldığını ifade etmektedir. Tablo 7'deki sınıflandırma metriklerine göre model, 3GPP sınıfını tahmin etmede daha başarılı olmaktadır. K-en yakın komşu sınıflandırıcısındaki benzer şekilde 3GPP sınıfı için duyarlılık ve IEEE sınıfı için kesinlik daha yüksektir.

**Tablo 4.** Yapay sinir ağına farklı yığın boyutları için doğrulama sonuçları (Validation results for different batch sizes in artificial neural network)

Yığın boyutu	Doğruluk
128	0.8177500009536743
256	0.8142499923706055
512	0.8091249942779541
1024	0.8024999856948852

Farklı büyüklükteki yığınlar ile eğitilmiş TabNet modellerine ait çapraz doğrulama sonuçları Tablo 5'te gösterilmektedir. Yapay sinir ağındaki duruma benzer olarak en başarılı modelin 128 boyutundaki yığınlar ile eğitim gerçekleştirildiğinde elde edildiği ve bu sayının arttırıldıkça başarının düştüğü görülmektedir.

TabNet sınıflandırıcısı kullanılarak test seti üzerinde 0.8795 doğruluk değeri elde edildiği gösteren ilgili karışıklık matrisi Tablo 6'da verilmektedir. Tablo 7'deki performans metrikleri sınıf bazında incelendiğinde kesinlik, duyarlılık ve F1 skoru değerlerinin, kullanılan dört sınıflandırıcıda da 3GPP sınıfı için daha yüksek olduğu anlaşılmaktadır. TabNet ile karar ağacı sınıflandırıcısına paralel olarak sınıf içinde birbirine yakın kesinlik ve duyarlılık değerlerine ulaşılmıştır. K-en yakın komşu ve yapay sinir ağı sınıflandırıcılarında ise duyarlılığın 3GPP sınıfı için daha yüksek iken kesinliğin IEEE sınıfı için yüksek olması dikkat çekmektedir. Performans metriklerinin makro ortalama ve ağırlıklı ortalama değerlerine bakıldığında TabNet sınıflandırıcısının diğer modelleri geride bıraktığı görülmektedir. TabNet mimarisinin yüksek başarı göstermesinin sebebi kullanılan ardışık dikkat mekanizması ile öznetelik seçiminin örnek bazında gerçekleştirilmesidir. Modelin öğrenme sürecinde her bir örnek için en önemli öznetelikler seçilerek verim arttırılmakta ve yorumlanabilirlik sağlanmaktadır (Arık ve Pfister, 2021).

**Tablo 5.** TabNet'te farklı yığın boyutları için doğrulama sonuçları (Validation results for different batch sizes in TabNet)

Yığın boyutu	Doğruluk
128	0.875250
256	0.874250
512	0.866000
1024	0.863625

**Tablo 6.** Karışıklık matrisleri (Confusion matrices)

		Tahmin		
K-en yakın komşu (K = 17)	Gerçek	3GPP	3GPP	IEEE
		IEEE	1031	160
		IEEE	234	575
Karar ağacı (Kriter = Entropi)	Gerçek	3GPP	3GPP	IEEE
		IEEE	1016	175
		IEEE	160	649
Yapay sinir ağı (Yığın boyutu = 128)	Gerçek	3GPP	3GPP	IEEE
		IEEE	1050	141
		IEEE	221	588
TabNet (Yığın boyutu = 128)	Gerçek	3GPP	3GPP	IEEE
		IEEE	1073	118
		IEEE	123	686

**Tablo 7.** Sınıflandırma metrikleri (Classification metrics)

Model			Kesinlik	Duyarlılık	F1 skoru	Örnek sayısı
K-en yakın komşu	Sınıf bazında	3GPP	0.82	0.87	0.84	1191
		IEEE	0.78	0.71	0.74	809
	Ağırlıklı ortalama		0.80	0.79	0.79	2000
			0.80	0.80	0.80	2000
Karar ağacı	Sınıf bazında	3GPP	0.86	0.85	0.86	1191
		IEEE	0.79	0.80	0.79	809
	Ağırlıklı ortalama		0.83	0.83	0.83	2000
			0.83	0.83	0.83	2000
Yapay sinir ağı	Sınıf bazında	3GPP	0.83	0.88	0.85	1191
		IEEE	0.81	0.73	0.76	809
	Ağırlıklı ortalama		0.82	0.80	0.81	2000
			0.82	0.82	0.82	2000
TabNet	Sınıf bazında	3GPP	0.90	0.90	0.90	1191
		IEEE	0.85	0.85	0.85	809
	Ağırlıklı ortalama		0.88	0.87	0.87	2000
			0.88	0.88	0.88	2000

## 5. Sonuçlar (Conclusions)

Akıllı ulaşım konseptinin son yıllarda oldukça önem kazanmasıyla birlikte farklı V2X haberleşme standartları geliştirilmiştir. Bu standartların haberleşme gereksinimlerine göre güçlü ve zayıf yanları bulunmaktadır. Dolayısıyla tasarımlarda kullanılacak uygun V2X standardının seçilmesi kritik öneme sahip olacaktır. Bu doğrultuda 3GPP ve IEEE standartları analiz edilerek yapay bir veri seti oluşturulmuştur. Deneysel çeşitli sınıflandırma modelleri kullanılmış ve hiperparametre optimizasyonu gerçekleştirilmiştir. En basit makine öğrenmesi algoritmalarından biri olan K-en yakın komşu sınıflandırıcısı ile 0.80 doğruluk elde edilirken, tablo halindeki veriler üzerinde yüksek performans göstermesiyle bilinen karar ağacı sınıflandırıcısıyla ise 0.83 doğruluk değerine ulaşılmıştır. Bu değer, yapay sinir ağı sınıflandırıcısı kullanıldığında 0.82 olarak gözlemlenmiştir. Son olarak, tablo şeklindeki veriler için geliştirilmiş bir derin öğrenme mimarisi olan TabNet modeli eğitilmiş ve test aşamasında 0.88 doğruluk oranı elde edilmiştir. Sonuçlar karşılaştırıldığında TabNet mimarisinin diğer yapay öğrenme algoritmalarını geride bıraktığı ve sınıf bazında başarımın 3GPP standardı için daha yüksek

olduğu görülmektedir. Sınıfların kesinlik ve duyarlılık değerlerinin birbirine yakın olduğu iki sınıflandırıcı olan karar ağacı ve TabNet modelleri, diğer modellere göre daha güvenilir sonuçlar vermektedirler. 3GPP sınıfını tespit etmede bütün modeller 0.84 ve üzeri F1 skorlarına ulaşmışlardır. Fakat K-en yakın komşu ve yapay sinir ağı sınıflandırıcılarının diğer modellere göre daha düşük IEEE sınıfı duyarlılığına sahip olması, bu sınıfı tespit edebilmenin önemli olduğu durumlarda kullanılmamaları gerektiğini ifade etmektedir.

%80'in üzerinde başarıya ulaşıldığını gösteren sonuçlar, önerilen yöntemin akıllı ulaşım sistemleri altındaki araçlara entegre edilmesinin fayda getireceğini göstermektedir. Bu sayede bir akıllı araç üzerinde yer alan her iki haberleşme standardı ile ilgili modülden hangisinin ne zaman daha etkin şekilde kullanılabileceği otomatik olarak belirlenebilecek ve araç sahibi daha yüksek kalitede bir hizmet alabilecektir. Bir başka bakış açısı olarak, araç üreticisi firmalar hangi haberleşme modüllerini nasıl kullanabileceklerine karar verme noktasında bu çalışmadaki yöntemden yararlanabileceklerdir. Diğer bir deyişle araç üreticileri kendi çözümlerini bu çalışmada geliştirilen yöntem üzerine inşa edebileceklerdir. İlerleyen süreçte yapılacak çalışmalarda, mevcut yöntemde sınıf etiketi olarak kullanılan haberleşme standartları için alt sürüm

sayısının artırılması ve farklı yapay öğrenme modellerinin sisteme eklenmesi hedeflenmektedir.

### Teşekkür (Acknowledgment)

Bu çalışma 122E400 no'lu Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) projesi kapsamında desteklenmiştir.

### Kaynaklar (References)

- Abdellah, A. R., Alshahrani, A., Muthanna, A., Koucheryavy, A., 2021. Performance Estimation in V2X Networks Using Deep Learning-Based M-Estimator Loss Functions in the Presence of Outliers. *Symmetry*, 13(11), 2207.
- Ank, S. Ö., Pfister, T., 2021. Tabnet: Attentive interpretable tabular learning. In *Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 8, pp. 6679-6687)*.
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., 2015. Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4.
- Chollet, F., 2018. Keras: The python deep learning library. *Astrophysics source code library*, ascl-1806.
- Filippi, A., Moerman, K., Martinez, V., Turley, A., Haran, O., Toledano, R., 2017. IEEE802. 11p ahead of LTE-V2V for safety applications. *Autotalks NXP*, 1, 1-19.
- Gupta, B., Rawat, A., Jain, A., Arora, A., Dhama, N., 2017. Analysis of various decision tree algorithms for classification in data mining. *International Journal of Computer Applications*, 163(8), 15-19.
- Gyawali, S., Qian, Y., 2019. Misbehavior detection using machine learning in vehicular communication networks. In *ICC 2019-2019 IEEE International Conference on Communications (ICC) (pp. 1-6)*. IEEE.
- Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E., 2020. Array programming with NumPy. *Nature*, 585(7825), 357-362.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T. Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- Kramer, O., 2013. Dimensionality reduction with unsupervised nearest neighbors (Vol. 51, pp. 13-23). Berlin: Springer.
- MacHardy, Z., Khan, A., Obana, K., Iwashina, S., 2018. V2X access technologies: Regulation, research, and remaining challenges. *IEEE Communications Surveys & Tutorials*, 20(3), 1858-1877.
- McKinney, W., 2010. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference (Vol. 445, No. 1, pp. 51-56)*.
- Moreira, D. C., Guerreiro, I. M., Sun, W., Cavalcante, C. C., Sousa, D. A., 2020. QoS predictability in V2X communication with machine learning. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring) (pp. 1-5)*. IEEE.
- Naik, G., Choudhury, B., Park, J. M., 2019. IEEE 802.11 bd & 5G NR V2X: Evolution of radio access technologies for V2X communications. *IEEE access*, 7, 70169-70184.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Rahmati, O., Avand, M., Yariyan, P., Tiefenbacher, J. P., Azareh, A., Bui, D. T., 2020. Assessment of Gini-, entropy-and ratio-based classification trees for groundwater potential modelling and prediction. *Geocarto International*, 1-20.
- Sevgican, S., Turan, M., Gökarslan, K., Yilmaz, H. B., Tugcu, T., 2020. Intelligent network data analytics function in 5G cellular networks using machine learning. *Journal of Communications and Networks*, 22(3), 269-280.
- Shrestha, R., Nam, S. Y., Bajracharya, R., Kim, S., 2020. Evolution of V2X communication and integration of blockchain for security enhancements. *Electronics*, 9(9), 1338.
- Skiribou, C., Elbahhar, F., 2021. V2X wireless technology identification using time-frequency analysis and random forest classifier. *Sensors*, 21(13), 4286.
- Tangirala, S., 2020. Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm. *International Journal of Advanced Computer Science and Applications*, 11(2), 612-619.
- Ullah, H., Nair, N. G., Moore, A., Nugent, C., Muschamp, P., Cuevas, M., 2019. 5G communication: an overview of vehicle-to-everything, drones, and healthcare use-cases. *IEEE Access*, 7, 37251-37268.
- Yazar, A., Dogan-Tusha, S., Arslan, H., 2020. 6G vision: An ultra-flexible perspective. *ITU Journal on Future and Evolving Technologies*, 1(1), 121-140.
- Yazar, A., 2021. Requirement Analysis and Clustering Study for Possible Service Types in 6G Communications. In *IEEE Sinyal İşleme ve İletişim Uygulamaları Kurultayı (pp. 1-4)*. IEEE.
- Zhang, W., Feng, M., Krunz, M., Volos, H., 2020. Latency prediction for delay-sensitive v2x applications in mobile cloud/edge computing systems. In *GLOBECOM 2020-2020 IEEE Global Communications Conference (pp. 1-6)*. IEEE.
- Zhang, X., Peng, M., Yan, S., Sun, Y., 2019. Deep-reinforcement-learning-based mode selection and resource allocation for cellular V2X communications. *IEEE Internet of Things Journal*, 7(7), 6380-6391.
- Zhao, L., Fang, J., Hu, J., Li, Y., Lin, L., Shi, Y., Li, C., 2018. The performance comparison of LTE-V2X and IEEE 802.11p. In *2018 IEEE 87th Vehicular Technology Conference (VTC Spring) (pp. 1-5)*. IEEE.