

YAPAY SİNİR AĞLARI VE DESTEK VEKTÖR MAKİNELERİ YÖNTEMLERİ İLE BÖLGESEL TRAFİK YOĞUNLUK TAHMİNİ

İclal ÇETİN TAŞ^{1*}, Ahmet Anıl MÜNGEN²

¹⁻²OSTİM Teknik Üniversitesi, Mühendislik Fakültesi, Yazılım Mühendisliği Bölümü, Ankara, Türkiye

Geliş Tarihi/Received Date: 14.07.2021 Kabul Tarihi/Accepted Date: 20.10.2021 DOI: 10.54365/adyumbd.971461

ÖZET

Büyük şehirlerde kilometre başına düşen insan yoğunluğu arttıkça trafik sıkışıklığı artmakta ve yolcuların daha fazla sürelerini trafikte harcamaktadırlar. Trafik sıkışıklığı nedeni ile harcanan ekstra zaman ve yakıt hem kullanıcılar hem de ülkeler için büyük bir gider kalemidir. Büyükşehirlerde yaşayan vatandaşların trafik yoğunluğunun zaman bazlı değişimini tahmin etmek ve buna göre planlama yapmaları bir zorunluluk haline dönüşmüştür. Trafik sıkışıklıkları genelde tüm şehirde aynı anda gerçekleşmez. Bölgesel olarak yaşanan trafik sıkışıklıkları diğer yolları da etkilemesi ile yaygınlaşır. Bu çalışma da yapay sinir ağları (YSA) kullanılarak önerilen yöntem ile geçmiş trafik verileri kullanarak bölgesel yoğunluklar tahmin edilmeye çalışılacaktır. Çalışma birçok benzer çalışmadan farklı olarak hava durumu gibi çevresel etkenleri de alarak tahmin modellemesinin başarısını arttırılmıştır. İstanbul Büyük Şehir Belediyesi Açık Veri Portalından toplanan 75 farklı noktaya ait 150.000 veri kullanarak önerilen model test edilmiş ve yaklaşık %90 başarı ile bölgesel trafik yoğunluğu tespit edilebilmiştir.

Anahtar Kelimeler: Trafik Tahmini, YSA, SVM, Ortalama Hız, Büyük Veri.

PREDICTION OF REGIONAL TRAFFIC INTENSITY WITH ARTIFICIAL NEURAL NETWORKS AND SUPPORT VECTOR MACHINES

ABSTRACT

As the density of people per kilometer increases in big cities, traffic congestion increases and passengers spend more time in traffic. The extra time and fuel spent due to traffic congestion is a big expense item for both users and countries. It has become a necessity to predict the time-based change in the traffic density of citizens living in metropolises and to plan accordingly. Traffic jams don't usually happen in the whole city at once. Regional traffic jams become widespread as they affect other roads. In this study, it will be tried to predict regional congestions by using historical traffic data with the proposed method using artificial neural networks (ANN). The study increases the success of forecasting modeling by taking environmental factors such as weather conditions apart from many equivalent studies. Using 150,000 data from 75 different points collected from the Istanbul Metropolitan Municipality Open Data Portal, the proposed model was tested and the regional traffic density could be determined with 90% success.

Keywords: Traffic Prediction, ANN, SVM, Average Speed, Big Data .

1. Giriş

Bir ulusal medya ajansı olan Ajans Press'in medyadan ve dijital uygulamalardan derlediği bilgilere göre İstanbul'da çalışanların insanların üçte biri her iş günü yaklaşık 2 saatten fazla süreyi trafikte zaman kaybetmektedir [1]. Ankara'da çalışan insanlar için ise bu oran 1 saatin biraz üzerindedir.

*¹ e-mail: iclal.cetintas@ostimteknik.edu.tr ORCID ID: <https://orcid.org/0000-0002-1101-9773> (Sorumlu Yazar),

² e-mail: ahmetanil.mungen@ostimteknik.edu.tr ORCID ID: <https://orcid.org/0000-0002-5691-6507>

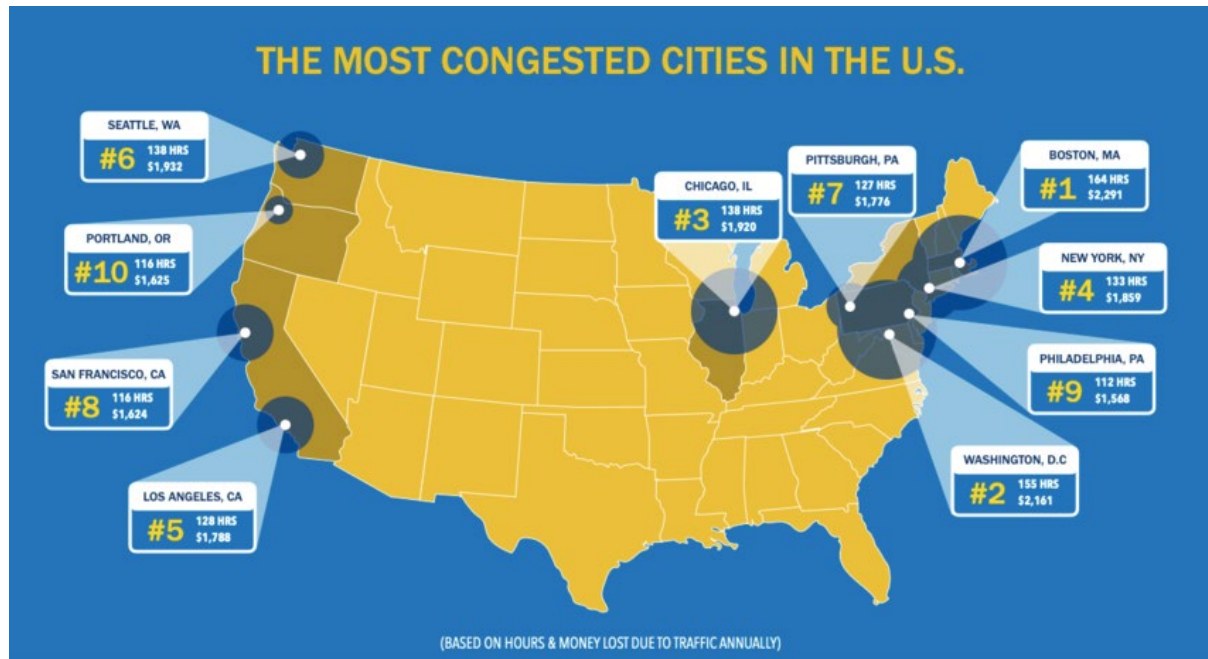
CuHaftada altı gün çalışan bir kişi 2021 yılı içerisinde 299 gün mesaiye gitmektedir. İstanbul'da çalışan bir kullanıcı 1 yıl içerisinde yaklaşık 600 saatini yani 25 gününü trafikte harcamaktadır. Bundan dolayı özellikle trafiğin hızlandırılması ile ilgili yapılacak çalışmalar büyük şehirde yaşayan insanların hayatını kolaylaştırması adına arz etmektedir.

Makinalar gibi sistemlerinde sahip oldukları işlem gücü diğer bir deyişle kapasitenin belirli bir sınırı vardır. Kısıtlı kaynaklarla sonsuz ihtiyaçları giderme ekonominin temel maddelerinden biridir. Diğer bir deyiş ile kapasite ve sınırlamalar günlük hayatın temel kurallardan biridir. Hiçbir kaynak sınırsız olmadığı için hizmet veren altyapıların veya makinaların kapasitelerinin optimize edilmesi hem hizmetinin verimliliği hem de ekonomik faydanın maksimize edilmesi önemlidir.

Sanayi devriminin ardından teknoloji geliştikçe makine ve araçların hayatımızdaki yeri artmıştır. Günlük hayatta artık vazgeçilmez olarak kabul edilen araçlardan biri de kişisel kullanımdaki otomobillerdir. Teknolojik gelişmeler ile araç üretiminin artması ve maliyetlerinin düşmesi araçları daha alt ekonomik düzeydeki kullanıcıların ulaşmasını sağlamıştır. Ayrıca altyapılardaki gelişmeler bireysel taşıtların kullanım oranını önemli düzeyde arttırmıştır. Yöneticilerin bireysel taşıt sayısı artan kullanıcılardan gelen talepler doğrultusunda daha fazla ulaşım altyapısına yatırım yapması da yine bireysel taşıtların sayısını artıran etkenlerden biridir.

Her sınırlı sistem de arz talebi karşılamadığı zaman yoğunluk yaşanabilir. Arz/Talep dengesi olarak adlandırılan bu olgu kapasitenin üzerinde hizmet talebi olduğunda görülür ve hizmet vermede sıkışma ve beklemler olur. Bazı durumlarda talep zamana homojen yayılmayacağı için günün belli saatleri veya belli dönemler için hizmette yoğunluk ve sıkışıklık ortaya çıkabilir. Hem dijital dünyada hem de fiziksel dünyada bu tür sıkışıklıkları sıklıkla görmekteyiz.

Talep ve altyapı dengesizliğinden dolayı sürecin düzgün işlemediği ve kullanıcıların beklediği alanlardan biri de trafiktir. Trafik tanım olarak araçların aynı altyapıyı yani yolları trafiğin işleyişi nedeniyle araçlar belli bir sırayla yer alır ve bu sırayla yolları yani altyapıyı kullanır. Eğer hizmet sağlayıcı olan yolların kapasitesi talep edenlerin yani hizmeti kullanmak isteyen araçların sayısından daha düşükse araçlar aynı anda hizmet alamayacağı için bir sıra oluşur. Bu durumda araçların birbirini beklemesine trafik sıkışıklığı olarak isimlendirilir. Trafik kullanıcılar için hem zaman hem de yakıt kaybı yarattığı için ekonomik sonuçları olan bir problemdir.



Şekil 1. ABD için en çok zaman ve para kaybına yol açan trafik sıkışıklıklarının haritası [2]

Tüm dünyada büyük şehirlerin sayısında ve bu şehirlerin nüfuslarında artma eğilimi vardır [3]. Ayrıca dikey mimarinin artması nedeni ile küçük bir bölgede yaşayan insan yoğunluğu da artmaktadır. Böylece bölgede yer alan insan sayısı ve bölgede kullanılan araç sayısı da artmaktadır. Bireysel araç sahibi sayısının da artmasıyla oluşan sıkışıklık hemen hemen tüm dünyada görülen ortak bir problemdir. Bu trafik sıkışıklığından dolayı her sene milyonlarca insan milyarca dakika zaman kaybetmektedir. Yapılan bir araştırmaya göre trafik sıkışıklığının Amerika Birleşik Devletleri ekonomisine maliyeti 2018 yılı için 87 Milyar dolar olmuştur [2]. Şekil 1’de ABD’de en çok trafik sıkışıklığı olan bölgelerin harita üzerinde konumları gösterilmiştir.

Araştırmacılar trafik sıkışıklıklarını düzenli tekrarlayan ve tekrarlamayan olmak üzere 2 ana gruba ayırmışlardır. Toplu etkinlikler, yol bakım çalışması, ani hava durumu değişiklikleri ve kaza yapan araçların yarattığı durumlar da dâhil olmak üzere trafik yoğunluğundaki olağandışı değişimler tekrarlanmayan trafik sıkışıklığı kategorisindedir [4]. Trafik sıkışıklarının yarısından fazlası tekrarlayan sıkışıklardır [5]. Tekrarlayan trafik sıkışıklığı incelendiğinde sıkışıklığının nedeninin daha çok belli bölgede yolların sıkışması ve o yolların yoğunlaşmasından kaynaklı tüm yollarda yoğunlaşma olduğu gözlemlenmektedir. Trafik problemi aslında genellikle bir bölgedeki tüm yollar ile ilintili olmayıp bazı kesişim noktalarında ortaya çıkmaktadır.

Araç yolları konusunda yapılan altyapı çalışmaları hem mali hem de insan kaynakları olarak ciddi bir planlama ve bütçe gerektirmektedir. Bundan dolayı da şehir planlayıcıları ve yöneticiler ihtiyaçtan çok daha geniş yollar yapmak yerine kısa ya da orta vade de şehir büyüme planlarına göre yollar yapmaktadır. Altyapı çalışmaları dışında yerel yöneticiler trafiğin akışını düzenlemeleri için de bazı yetkileri vardır. Trafik lambalarının yönleri ve sürelerini değiştirerek ve yollar ile ilgili akış kurallarını gün içinde farklı zamanlarda değiştirerek kısmi olarak trafiğin sıkışıklığının önüne geçilmesi sağlanabilir. Yerel yöneticilerin alacağı bu tip önlemler düşük maliyet ile kolayca uygulanabilir. Bu çözümler fiyat/performans metriğine göre değerlendirildiğinde trafik sıkışıklığında azaltılmasında önemli bir faydaya sahiptir.

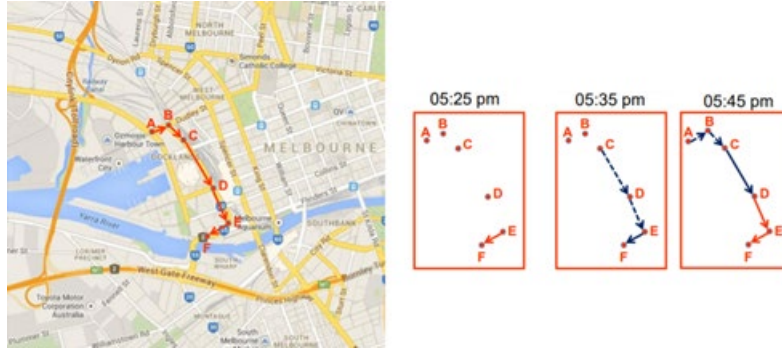
İstanbul içerisindeki bir bölgede yer alan ve araçların geçiş hızlarını ölçen sensörler kullanılarak trafikten alınan verilerin geçmişe yönelik analiz edilmesi ve bölgesel olarak hava durumu verilerinin ile birleştirilerek gelecekteki trafik sıkışıklığının zaman ve yoğunluk olarak tespitinin yapılabileceği bir yöntem sunar. Çalışma sonucunda önerilen yöntem ile belli bir tarihte seçilen bölgedeki trafik yoğunluğunu tahmin edilmeye çalışılmıştır. Böylece hava durumunun etkisiyle trafik sıkışıklığı oluşacak bölgeleri tespit edip yerel yöneticilere raporlayabilecek bir model ortaya koyulmuştur. Böylece yoğunluk olacak bölgeler için trafik kuralı düzenlemeleri yapılabilir. Yeni yapılacak yol planlamalarında bu çalışma ile elde edilen sonuçlar analiz edilebilir. Ayrıca belediye görevli araçlarının görev saati ve yerinin belirlenmesinde planlama amaçlı kullanılabilir.

Bu çalışma dört bölüm ile devam etmektedir. İkinci bölüm trafikle ilgili yapılan önemli çalışmalar sıralanmış ayrıca İstanbul trafiği ile ilgili yapılan çalışmalar için ayrı bir paragraf ayrılmıştır. Üçüncü bölümde toplanan veri ve uygulanan yöntem ile yöntemin sonuçları detaylı şekilde açıklanmıştır. Son bölüm ise tartışma ve sonuç kısmıdır.

2. Literatür Taraması

Trafik sıkışıklarının nedenleri ve çözümleri ile ilgili yapılan çalışmalar akademik yayınlar ve raporlar olmak üzere ikiye ayrılabilir. Amerikan Ulusal Ulaştırma Dairesi tarafından akademisyenlere ve araştırma kuruluşlarına yaptırılan birçok rapor mevcuttur. Mazzenga ve Demetsky yaptıkları araştırma raporunda önerdikleri yöntem ile trafik sorununu çözmek için otoyollar ile ilgili yapılabilecek değişikliklerin ilk yıl getirisinin yarım milyon dolar olduğu ve sadece onların önerdiği çözümden toplam 8 milyon dolara yakın tasarruf yapılabileceğini göstermiştir [6]. Amerika Birleşik Devletleri Ulusal Ulaşım Kütüphanesinden [7] de bu konuda yapılmış çokça çalışma mevcuttur. Çalışmalar alınan kısa vadeli önerilerin de trafiğinin akışını olumlu etkilediği göstermişlerdir. Amerikan Ulusal Ulaştırma

Dairesi raporları daha çok otoyollar ve bu otoyollara bağlı yoğunlaşmalar ile ilgilenmiştir. Bu çalışmada ise şehir içi ve çoğu otoyol olmayan ana ve tali yollardaki yoğunluklara odaklanılmıştır.



Şekil 2. Zaman periyodu boyunca futbol maçlarından sonra yoğunluğun hesaplanması ve modellenmesi [8]

Derin sinir ağları birçok çalışmada daha önce trafik sıkışıklarının tespiti amacı ile kullanılmıştır. Sun ve arkadaşları [9] yaptıkları çalışmada derin sinir ağları kullanarak tekrarlamayan ve futbol maçlarından kaynaklı oluşan trafik yoğunluğunu tespit eden ve çözüm öneren bir sistem önermişlerdir. Sun ve arkadaşlarının önerisi tekrarlamayan sıkışıklıklara odaklanmış olmasıyla bizim çalışmamızdan farklı bir alana odaklanmıştır. Buna karşın sıralı kontrol noktalarını takip etmesi ve noktaların yoğunluklarını tespit etmesi bu çalışmada yer alan aynı caddenin sırayla birden fazla bölümüne odaklanmamız ile benzerlikler taşır. Farrak ve arkadaşları [10] karayollarının sıkışıklığını hafifletmek için kullanılan Dynamic hard shoulder running ve ramp closure metotlarının başarısını simülasyonlar ile ölçmeye çalışmışlardır. Uyguladıkları testlerde önerilen yöntemlerin bazı durumlarda %50'ye yakın tasarruf sağladığını ortaya koymuşlardır. Brennan ve arkadaşları [11] eyaletler arası trafik sıkışıklığının tespiti ve analizi için yöntem önermişlerdir. Lan ve arkadaşları [12] tekrarlayan ve tekrarlamayan trafik sıkışıklıklarını ABD otoyollarında modelleyebilecek bir model önermişlerdir. Kidando ve arkadaşları [13] Random Variation Markov Structured Model kullanarak yinelenen trafik sıkışıklığının dinamik evrimini analiz etmiştir. Nguyen [14] ve arkadaşları nedensellik kavramı içinde birbirine neden olan sıkışıklıkların zamansal yayılımını araştırmış ve bu bilgiler ile gelecekteki tıkanıklıkları tahmin etmeye çalışmışlardır. Anjum ve arkadaşları [15] hem hava kalitesi hem trafik akış modellemesini kullanan yuğünsal bir yaklaşım ile trafik tıkanıklığını modellemeye çalışmıştır.

Amini ve arkadaşları [16] yaptıkları çalışmayla toplu taşıma ve toplu taşımanın kullandığı yolları modellemeye çalışmışlardır. Önerdikleri model ile şerit kapatma ve diğer durumlarda trafikteki artışı da öngörmeye çalışmışlardır. Her ne kadar bu çalışma sürekli tekrarlayan sıkışıklıklar odaklanılmasa da Sun ve arkadaşların [9] Şekil 2'de gösterilen yöntemlerinde olduğu gibi tekrarlamayan sıkışıklıkları çalışan araştırmacıların önerdikleri yöntemlerden yararlanılmıştır.

Bu çalışmalar dışında trafik akışı ve trafik akışındaki problemleri analiz etmek ile ilgili de farklı çalışmalar yapılmıştır. Wu ve arkadaşları [17] hibrit bir derin öğrenme yöntemi ile trafik akışını tahmin etmeye ve anlamaya çalışmışlardır. Fu ve arkadaşları [18] LSTM kullanarak trafik akış modellemesi önermişlerdir. Heejing ve arkadaşları [19] TensorFlow kullandıkları derin öğrenmeli bir model önermiş ve 5 dakikalık aralıklı gerçek veri kullanarak trafik tahmini yapan bir model önermişlerdir. Sharma ve arkadaşları [20] Hindistan'da iki şeritli bölünmemiş otoyol için yapay sinir ağını kullanarak kısa vadeli bir trafik tahmin modeli önermişlerdir. Her iki çalışmada da MAE minimum 15km/h olarak elde edilmiştir.

İstanbul'da yer alan trafik sorunu konusunda yapılmış akademik çalışmalarda mevcuttur. Çetiner ve arkadaşları [21] yapay sinir ağları kullanarak geçmiş veri ile İstanbul'daki büyük kesişim noktalarındaki değişen trafik boyutunu tespit etmeye çalışmışlardır. Canitez ve arkadaşları [22] İstanbul'un sürekli değişen ulaşım politikaları ve gelecek stratejilerini analiz etmiştir. Yaman ve arkadaşları [23] yaptıkları çalışma ile 366 noktadan toplanan trafik verisi ile RMSE ve MAPE

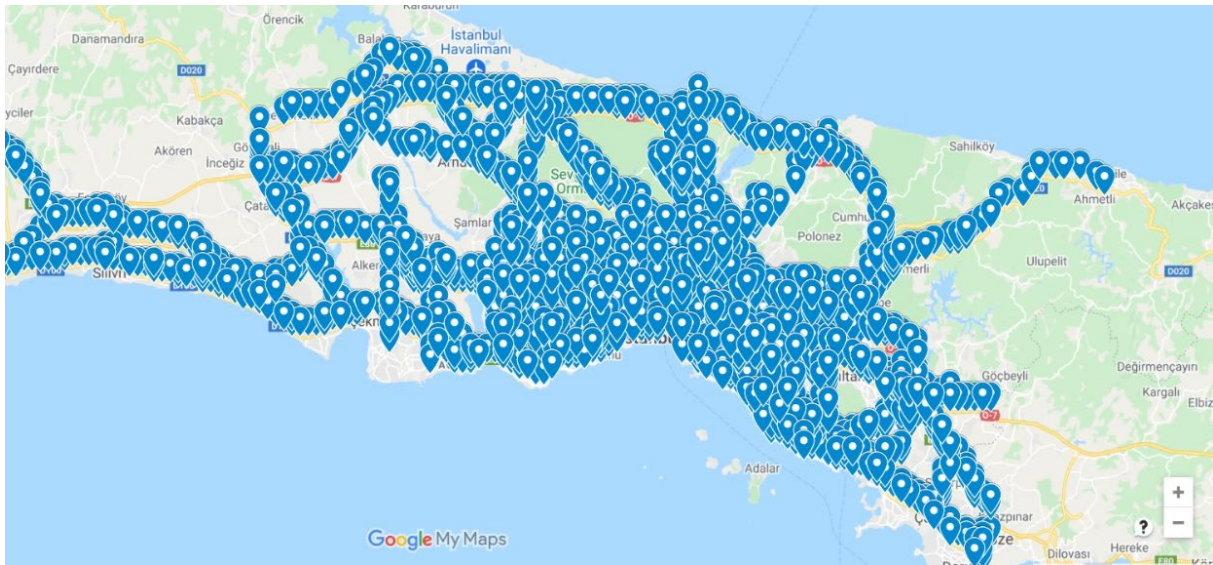
performans göstergelerini kullanarak trafik yoğunluğunun tespitine çalışmışlardır. Canitez ve Deveci [24] toplu taşıma sistemlerine bütünleşmiş araç paylaşım uygulamalarının önemini ve bu çerçevede bir model önermiştir.

3. Materyal ve Metod

Bu bölümde, çalışma için önerilen YSA algoritmasının yapısı ve uygulama adımları açıklanmıştır. Ayrıca çalışmada toplanan gerçek veriler ve bu verilerin özellikleri bu kısımda açıklanmıştır. YSA'lar biyolojik sinir ağlarından ilham alınarak tasarlanmıştır. YSA, tahmin uygulaması için sıklıkla yöntemler kullanmıştır. Destek vektör makineleri (SVM), denetimli bir makine öğrenimi yöntemidir. Bu yöntem YSA gibi birçok farklı alanda uygulanmaktadır. Bu çalışmada lineer destek vektör makineleri (SVM Lineer), radyal tabanlı fonksiyonları kullanan destek vektör makineleri (SVM RBF), çok katmanlı yapay sinir ağları (MLP) ve genel regresyon sinir ağı (GRNN) yöntemleri kullanılmıştır.

3.1. Veri Toplama

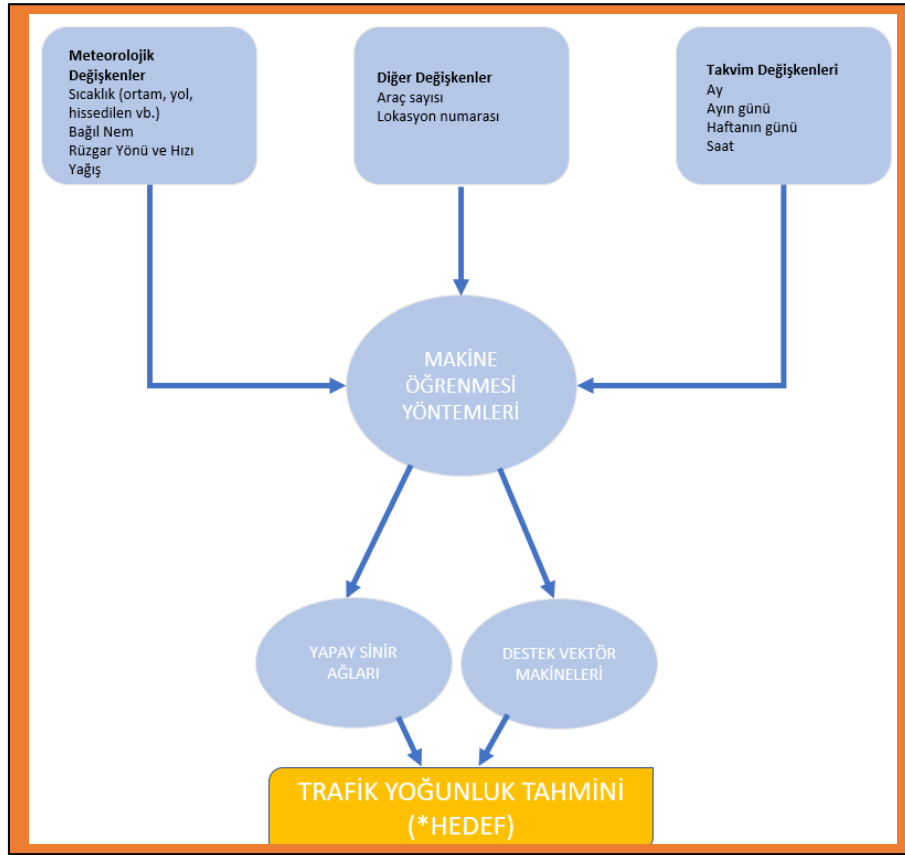
İstanbul Büyükşehir Belediyesi tarafından İstanbul ile ilgili elde edilen verilerin akademik ve diğer amaçlarla ücretsiz olarak erişebilmesi için IBB Açık Veri Portalı [25] kurulmuştur. Bu portal üzerine çevre, sağlık, yer bilimi de dahil olmak üzere birçok farklı disiplini ilgilendiren 100'den fazla çeşit veri paylaşımı yapılmıştır. İstanbul Büyükşehir Belediyesi İtfaiye Daire Başkanlığı tarafından paylaşılan veri seti 1 Ocak 2020 - 1 Ocak 2021 tarihleri arasında İstanbul bölgesinde bulunan 30 dan fazla istasyondan alınan saatlik hava durumu veriler mevcuttur. Bu servis aracılığı ile trafik yoğunluğu ölçüm istasyonundan kullanılan bölgeye en yakın hava durumu bilgisi alınmıştır.



Şekil 3. Çalışma kapsamında toplanan veri ve verinin harita üzerinde gösterimi.

İstanbul Büyükşehir Belediyesi Ulaşım Daire Başkanlığı tarafından paylaşılan veri seti 1 Ocak 2020 - 1 Ocak 2021 tarihleri arasında İstanbul bölgesinde bulunan 500 dan fazla trafik yoğunluk istasyonundan alınan saatlik hava durumu veriler mevcuttur. Bu çalışma kapsamında trafik yoğunluk ölçüm istasyonlarından 15'i, hava durumu istasyonlarından ise 10 tanesi kullanılmıştır. Toplam 21000 trafik yoğunluk durum verisi servis aracılığı ile alınmış en yakın hava durumu istasyonu verisi ile birleştirilerek sistemi eğitmek ve test etmek için veri tabanına kaydedilmiştir. Toplanan veri ve toplandığı alanın konumları Şekil 3'te gösterilmiştir. API ile elde edilen tüm veriler ilişkisel bir

veritabanı olan MySQL 'de tutulmuştur. Bu veriler model eğitimi yapılacağı belirli filtreler uygulanarak CSV formatında model eğitimi yapılacak platforma aktarılmıştır.



Şekil 4. Veri seti değişkenlerinin, yöntemlerinin ve hedefin akış şeması

3.2. GRNN

GRNN, önemli YSA modellerinden biridir. GRNN, sürekli değişkenler sağlar ve temel durgunluk düzlemine yakınsar. Bu yöntemin ana prensibi, olasılık yoğunluk fonksiyonlarının tahminidir. GRNN en önemli özellikleri hızlı eğitim sürelerinin olması ve doğrusal olmayan fonksiyonlar modellenenbilmeleridir. GRNN, doğrulanmamış bir doğrusallık varsayımı için kullanılabilir. GRNN, sadece bir eğitim seti vererek, x ve y 'nin birleşik olasılık yoğunluk fonksiyonunu tahmin etme prensibine dayanır. $Y(x)$, x girişinin tahmin değeridir ve şu şekilde tanımlanır[26] [27]:

$$Y(x) = \frac{\sum_{k=1}^N y_k K(x, x_k)}{\sum_{k=1}^N K(x, x_k)} \quad (1)$$

Yoğunluk fonksiyonu $K(x, x_k)$ radyal tabanlı fonksiyondur:

$$K(x, x_k) = e^{-\frac{D^2}{2\sigma^2}} \quad (2)$$

D^2 skaler fonksiyondur ve öklid mesafesini temsil etmektedir. Bu mesafenin hesaplanması x_k ve x girişler cinsinden Denklem 3 'teki gibi ifade edilebilir.

$$D^2 = (x - x_j)^T (x - x_j) \quad (3)$$

3.3. MLP

MLP, bir tür ileri beslemeli sinir ağıdır. Birincisi, işlenecek giriş sinyalini alan giriş katmanıdır. Çıkış katmanında tahmin ve sınıflandırma yapılır. Giriş ve çıkış katmanı arasına yerleştirilen isteğe bağlı sayıda gizli katman, MLP'nin hesaplama parçalarıdır. Veri, girişten çıkış katmanına ileri yönde akar. MLP'ler, herhangi bir sürekli fonksiyona yaklaşmak için tasarlanmıştır ve doğrusal ve ayrılabilir doğrusal olmayan sorunları çözebilir. MLP'nin başlıca kullanımı, örüntü sınıflandırma, tanıma, tahmin ve yaklaşım uygulamalarıdır.

$$a = \Phi \left(\sum_j w_j x_j + b \right) \quad (4)$$

$$o(x) = G(b(2) + W(2)h(x)) \quad (5)$$

$$h(x) = \Phi(x) = s(b(1) + W(1)x) \quad (6)$$

a birimin aktivasyonudur. $b(1), b(2)$ sapma vektörleridir; ağırlık matrisleri W , G ve s aktivasyon fonksiyonlarıdır [28][29]. s fonksiyonu tipik seçimler olarak belirlenebilir örneğin \tanh veya lojistik sigmoid seçilebilir.

3.4. SVM

SVM modelleri MLP'ye çok benzemektedirler. Genel olarak, sınıflandırıcıları eğitmek için kullanılır. Farklı kernel fonksiyonları kullanılarak uygulanabilir. Bu uygulama için RBF ve doğrusal kernel seçilmiştir. Doğrusal kısıtlamalarla ilgili ikinci dereceden bir programlama problemini çözmek, çekirdek işlevindeki ağ yapısının ağırlıklarını belirler.

SVM'ler doğrusal olmayan bir sınıflandırma uygulayabilir. SVM, hiper düzleme yakın veri noktalarını özetler. Genel olarak, orijinal maksimum marj hiperdüzlem algoritması olarak doğrusal bir sınıflandırıcı kullanılır. Bununla birlikte, doğrusal olmayan sınıflandırıcılar kullanarak maksimum marjlı hiperdüzlemler için çekirdek hilesinden farklı yöntemler kullanılır.

SVM yöntemi için kullanılan RBF kernel fonksiyonu Denklem 7'de [30] verilmiştir.

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (7)$$

$$\gamma = \frac{1}{2\sigma^2} \quad (8)$$

Burada x ve x' , RBF çekirdeğinin iki örneğini temsil eder. σ serbest parametre temsilidir.

3.5. GMDH

GMDH algoritması, normalde özel bir yapay sinir ağı türü olarak sınıflandırılan bir tür ileri beslemeli ağıdır. Model doğal seçim ilkesine bağlı olduğu için hayvan evrimi veya bitki yetişmesi ilkesine dayanmaktadır. Çok katmanlı kriterler, birbirini izleyen nesiller için üstün ağları korur ve sonunda optimal bir ağ sağlar.

GMDH algoritmasında, girdi ve çıktı değişkenleri arasındaki genel bağlantı, Kolmogorov-Gabor polinomunun Volterra fonksiyonel serisi ile ifade edilir[27] [31][32].

$$y(t) = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k + \dots \dots \quad (9)$$

$y(t)$ çıkış değişkenini, $X(x_1, x_2, \dots, x_m)$ giriş vektörlerini ve $A(a_1, a_2, \dots, a_m)$ ise ağırlıkları temsil etmektedir. Genellikle, denklem iki değişkenin ikinci dereceden formunda kullanılır. GMDH ağlarında, giriş katmanı, x ile gösterilen her giriş değişkeni için nöronlar içerir. Birinci katmandaki her bir nöron, girdilerini girdi katmanındaki iki nörondan alır. İkinci ve üçüncü katmandaki nöronlar, girdilerini bir önceki katmandaki nöronlardan ikisinden alır ve bu işlem çıktı katmanına kadar devam eder. Çıktı katmanı, girdilerinden ikisini önceki katmandan alır ve girdi ve çıktı değişkenleri arasındaki ilişkiyi sağlamada en uygun analitik ifadeyi gösteren nihai sonucu üretir.

GMDH ağlarında bir katmandaki nöron sayısı n ise, bir sonraki katmanda aday nöron sayısı hesabı 2 değişkenli polinomlar için Denklem 10'daki şekilde yapılır.

$$\binom{n}{2} = \frac{n(n-1)}{2} \quad (10)$$

Eğitim sürecinde, girişler ana eğitim verileri ve kontrol verileri olmak üzere iki farklı şekilde kullanılırlar. Kontrol verileri genellikle ana eğitim verilerinin yaklaşık %20'si kadar veri içerir. Eğitim algoritması sırasında, her bir nöron için ortalama hata karesi değeri hesaplanır ve kontrol verilerine de uygulanır. Geçerli katmandaki en iyi nöronun ortalama hata karesi, kontrol verisi bir önceki katmandaki en iyi nöronun ortalama hata karesinden daha düşüktür ve maksimum katman sayısı henüz elde edilmemişse eğitim süreci yeni bir katman oluşturmaya devam eder. Aksi takdirde eğitim süreci durur. Aşırı uyum gösterme başladığında, kontrol verileriyle ölçülen hatanın artacağı, dolayısıyla eğitim sürecinin duracağı unutulmamalıdır.

Yöntemin Pozitif Yönlerini;

- Verilen probleme göre uyarlanabilir ağ topolojileri sunar,
- Güvenilirlik sayesinde yerel olarak iyi ağırlıklar bulurlar,
- Seyrek bağlantı yapısı ile hızla eğitilebilirler,

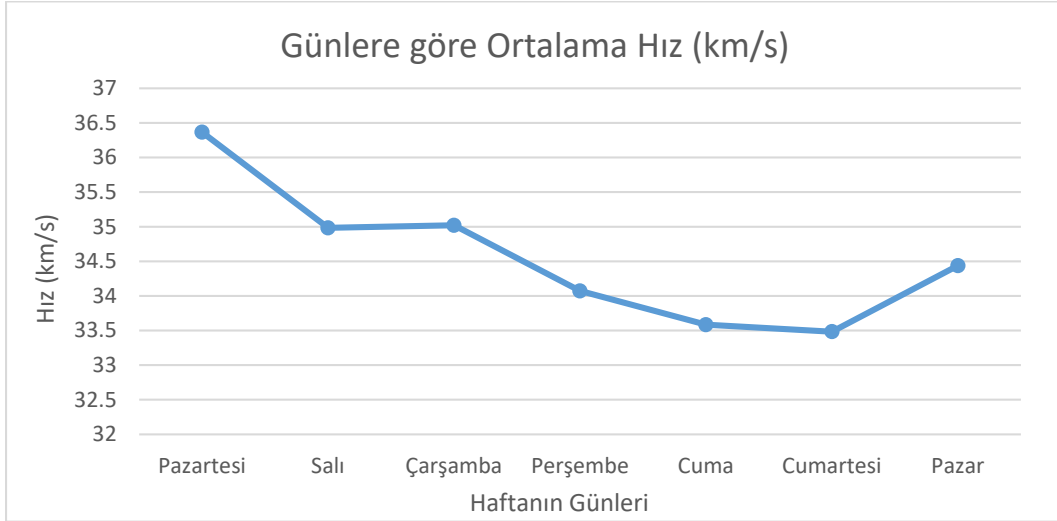
Negatif Yönlerini;

- Basit sistemler için oldukça karmaşık polinomlar üretme eğilimindedirler,
- Gerçek yapıyı oluşturmayı garanti etmezler,
- En küçük kareler yöntemi ile katsayıları yanlış tahmin edebilirler,

olarak sıralayabiliriz.

4. Deneysel Sonuçlar ve Tartışma

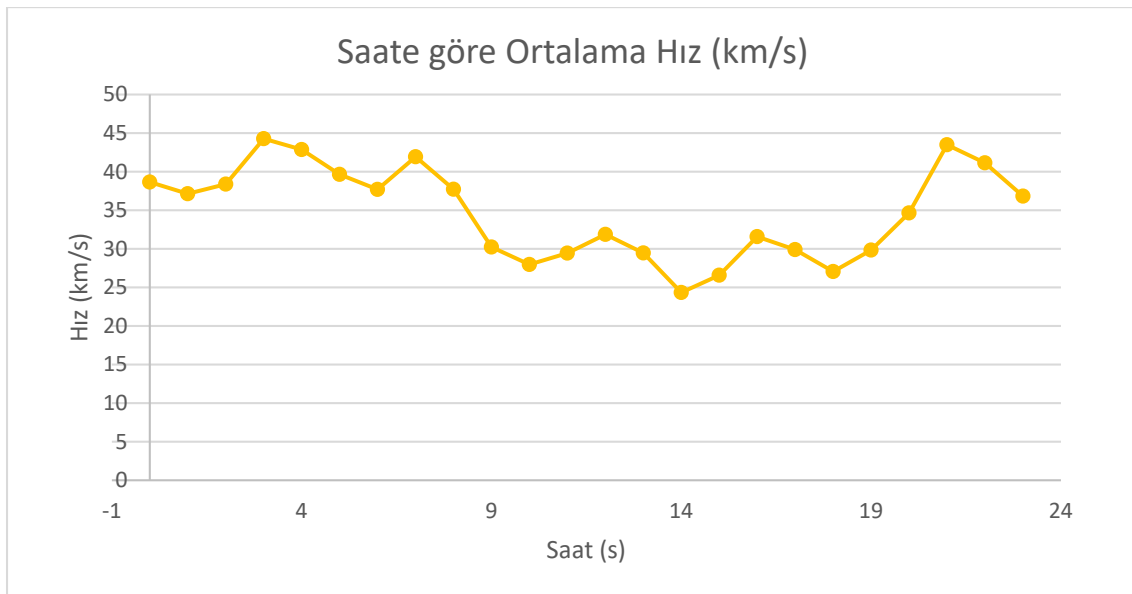
Trafik verisi çok hareketlilik gösteren bir veridir haftanın günü, saat, hava durumu, bölgedeki özel bir durum ya da etkinliğe (kongre, miting, maç, konser vb.) göre de anlık olarak değişmektedir. Şekil 5' de güne göre ortalama hız değişimlerinin grafikleri verilmiştir. Ortalama hız verisi IBB servisinden alınmıştır.



Şekil 5. Haftanın Günlerine göre Ortalama Hız Değeri Grafiği

Grafik incelendiğinde haftanın ilk iş günü olan pazartesi gününde başlamak üzere ortalama hızın azaldığı, pazar günü ise iş yoğunluğu minimum seviyede olup sosyal yoğunluktan kaynaklı trafik sıkışıklığı bulunduğundan ortalama hızın bir miktar arttığı görülmüştür.

Şekil 6'da saate göre ortalama hız değişimlerinin grafikleri verilmiştir. Benzer şekilde burada da mesai saatleri içerisinde hızın en minimum seviyelerde olduğu gözlemlenmiştir.



Şekil 6. Günün Saatine göre Ortalama Hız Değeri Grafiği

MySQL veritabanında yer alan hava durumu ve yol verileri çalışmanın kapsamını daraltmak için belirli bir bölgedeki verileri almıştır. Bu ön işleme adımı JAVA ile gerçekleştirilmiştir. Önışlemeden geçen veriler MATLAB'a aktarılacak için ortak CSV dosyalarına yazılmıştır. Çalışmada gerçekleştirilen uygulamalar MATLAB kullanılarak eğitilmiş ve test edilmiştir.

	1	2	3	4	5	6	7	8	9	10	11	12
	id	DOW	HOURS	MONTH	DOM	DATETT	AVERAGE_SPEED	AVERAGE_TEMPERATURE	NEM	MINIMUM_SPEED	MAXIMUM_SPEED	NUMBER_OF_VEH
1	261898	4	0	1	1	1'01 January'	33	6	77	6	75	
2	295686	4	1	1	1	1'01 January'	32	5	77	6	96	
3	670242	4	2	1	1	1'01 January'	29	5	67	6	90	
4	14466	4	3	1	1	1'01 January'	34	4	62	6	107	
5	191923	4	4	1	1	1'01 January'	33	5	65	7	68	
6	288919	4	5	1	1	1'01 January'	38	5	75	7	78	
7	177042	4	6	1	1	1'01 January'	38	5	88	6	85	
8	70475	4	7	1	1	1'01 January'	36	5	93	6	110	
9	233395	4	8	1	1	1'01 January'	15	6	95	6	30	
10	265544	4	9	1	1	1'01 January'	37	6	89	7	75	
11	36215	4	10	1	1	1'01 January'	31	8	68	6	74	
12	54783	4	11	1	1	1'01 January'	29	8	65	6	64	
13	169835	4	12	1	1	1'01 January'	32	8	62	6	78	
14	151223	4	13	1	1	1'01 January'	29	7	66	6	73	
15	166276	4	14	1	1	1'01 January'	31	6	67	6	79	
16	415923	4	15	1	1	1'01 January'	13	6	66	6	27	
17	51712	4	16	1	1	1'01 January'	28	6	64	6	93	
18	149605	4	17	1	1	1'01 January'	22	6	62	6	71	

Şekil 7. Matlab ortamında verilerin aktarımı

Bu bölümde, çalışmada uygulanan yöntemlerde göre kullanılan konfigürasyon bilgileri sırasıyla açıklanmıştır. MLP'de, en uygun sayıda gizli katman nöronunun aranması için 2 ila 20 nöron arası seçilmiştir. Aktivasyon fonksiyonları, gizli katman için lojistik, çıkış katmanı için doğrusal olarak kullanılır. GRNN, Gauss çekirdek işleviyle birlikte kullanılır. σ değer aralığı 0.001 ile 10 arasında belirlenir. GMDH uygulaması için maksimum ağ katmanı sayısı 20, maksimum polinom derecesi 16, her katmandaki maksimum nöron sayısı 20 olarak belirlenmiştir.

SVM uygulaması için çekirdek fonksiyonu olarak RBF ve lineer olmak üzere iki ayrı fonksiyon seçilmiştir. ε -SVR regresyon modeli kullanılmış ve parametre aralıkları C için 0.1-5000, ε için 0.0001-100 ve γ için 10-3 ve 0.001-50 arasında belirlenmiştir.

Tüm modeller için rastgele örnekleme uygulanır. Eğitim verilerini oluşturmak için veri setinin %80'i oranında belirlenmiş bir oran değeri kullanılır ve veri setinin %20'si doğrulama verilerini oluşturmak üzere ayrılır.

Yöntem performanslarını karşılaştırmak için ortalama mutlak hata (MAE) (Denklem10) ortalama mutlak yüzde hata (MAPE) (Denklem11) ve R-kare (R^2) (Denklem12) metrikleri kullanılır. Literatürde trafik ile ilgili araştırmalarda genel olarak söz konusu metriklerin kullanıldığı görülmüştür. Bu nedenle, bu çalışmada bunları karşılaştırmak için aynı hata hesaplama türleri seçilmiştir.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(y_i - \hat{y}_i)^2} \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (11)$$

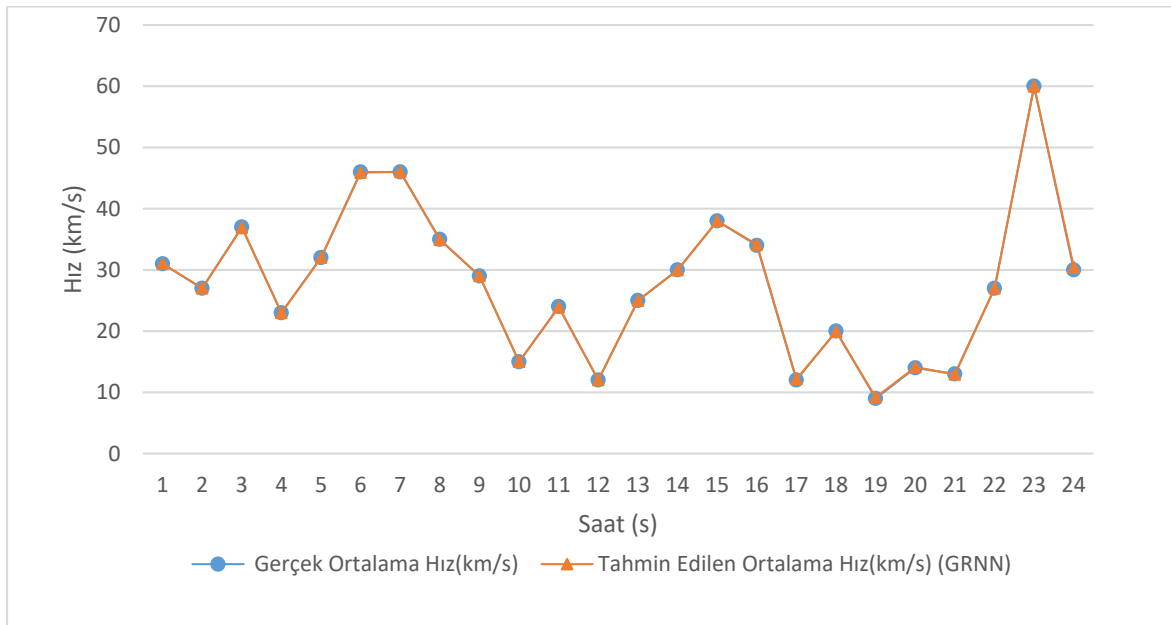
$$(12)$$

$$MAPE(\%) = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i|}$$

Çizelge 1. Uygulanan Yöntemler Performanslarının Metrik Değerler Üzerinden Karşılaştırılması

Konum 1		Hesaplanan Hata Oranı		
Yöntem	Tür veya kernel çeşidi	MAE (km/h)	MAPE (%)	R ² (%)
YSA	MLP	4.272	13.508	82.762
	GRNN	3.078	9.372	89.888
	RBF	3.912	12.645	84.885
	GMDH	4.089	13.496	83.343
SVM	SVM RBF	3.876	12.622	84.928
	SVM Lineer	3.999	12.901	85.272

Şekil 8’de rastgele seçilen 1 gün için gerçek ve GRNN ile tahminlenen ortalama hız değerlerine ait grafik görülmektedir. En düşük hata oranına sahip olan bu yöntem ile yapılan tahminlerin gerçek değerler ile örtüştü gözlemlenmektedir.



Şekil 8. Gerçek ve Tahminlenen Ortalama Hız Değeri Grafiği

5. Sonuçlar

Bu çalışma sonucunda yapay sinir ağları ve kara destek makineleri kullanılarak belirli bir zaman aralığı ve konum için trafik yoğunluğu tahmin modeli önerilmiştir. Literatürde makine öğrenmesi metotları kullanılarak yapılan bu uygulamalara ilgi duyulmaktadır. Birbirinden farklı kurgulara sahip yapay sinir ağı modelleri yapılmış ve sonuçları metrik olarak karşılaştırılmıştır. Önerilen modeli geliştirmek için 75 noktadan toplanan 150.000 trafik verisi kullanılmıştır. Ayrıca veri setine bölgesel hava tahmini raporları (sıcaklık, nem rüzgâr, yağış v.b) gibi çevresel raporlar ile trafik yoğunluğu tahminindeki başarı geliştirilmiştir. Yapay sinir ağları yöntemlerinin trafik yoğunluğu tahminleme

uygulamalarında kullanılabileceği görülmüştür. Önerilen model %90 başarı oranı ile trafik yoğunluğunu doğru tespit ettiği görülmüştür. Böylece İstanbul'da yaşayan yerel yöneticilerin ve araç sahiplerinin yollar ile ilgili yoğunluk tahmini veren bir model önerilmiştir.

Kaynaklar

- [1] SputnikNews TR, "İstanbul'da trafikte ne kadar vakit kaybediyor? - Sputnik Türkiye." <https://tr.sputniknews.com/analiz/201712201031487131-istanbullular-trafikte-ne-kadar-vakit-kaybediyor/> (accessed Jun. 17, 2021).
- [2] INRIX, "Scorecard - INRIX," 2020. Accessed: Mar. 18, 2021. [Online]. Available: <https://inrix.com/scorecard/>.
- [3] E. Romanova, "Increase in Population Density and Aggravation of Social and Psychological Problems in Areas with High-Rise Construction," in *E3S Web of Conferences*, Mar. 2018, vol. 33, p. 03061, doi: 10.1051/e3sconf/20183303061.
- [4] M. E. Hallenbeck, J. M. Ishimaru, and J. Nee, "MEASUREMENT OF RECURRING VERSUS NON-RECURRING CONGESTION," Washington (State). Dept. of Transportation, Oct. 2003. Accessed: Mar. 18, 2021. [Online]. Available: <https://rosap.ntl.bts.gov/view/dot/38804>.
- [5] "FHWA Operations - Reducing Recurring Congestion." https://ops.fhwa.dot.gov/program_areas/reduce-recur-cong.htm (accessed Mar. 18, 2021).
- [6] N. J. Mazzenga Graduate, R. Assistant, and M. J. Demetsky, "Investigation of Solutions to Recurring Congestion on Freeways Virginia Transportation Research Council," Virginia Transportation Research Council, 2009. Accessed: Mar. 18, 2021. [Online]. Available: http://www.virginiadot.org/vtrc/main/online_reports/pdf/09-r10.pdf.
- [7] "Welcome to ROSA P | Welcome." <https://rosap.ntl.bts.gov/> (accessed Mar. 19, 2021).
- [8] F. Sun, A. Dubey, and J. White, "DxNAT - Deep neural networks for explaining non-recurring traffic congestion," in *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, Jul. 2017, vol. 2018-January, pp. 2141–2150, doi: 10.1109/BigData.2017.8258162.
- [9] F. Sun, A. Dubey, and J. White, "DxNAT - Deep neural networks for explaining non-recurring traffic congestion," in *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, Jul. 2017, vol. 2018-Janua, pp. 2141–2150, doi: 10.1109/BigData.2017.8258162.
- [10] S. G. Farrag, F. Outay, A. U.-H. Yasar, and M. Y. El-Hansali, "Evaluating Active Traffic Management (ATM) Strategies under Non-Recurring Congestion: Simulation-Based with Benefit Cost Analysis Case Study," *Sustainability*, vol. 12, no. 15, p. 6027, Jul. 2020, doi: 10.3390/su12156027.
- [11] T. M. Brennan, R. A. Gurriell, A. J. Bechtel, and M. M. Venigalla, "Visualizing and Evaluating Interdependent Regional Traffic Congestion and System Resiliency, a Case Study Using Big Data from Probe Vehicles," *J. Big Data Anal. Transp.*, vol. 1, no. 1, pp. 25–36, Jun. 2019, doi: 10.1007/s42421-019-00002-y.
- [12] C.-L. Lan, R. Venkatanarayana, and M. D. Fontaine, "Development of a Methodology for Determining Statewide Recurring and Nonrecurring Freeway Congestion: Virginia Case Study," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2673, no. 6, pp. 566–578, Jun. 2019, doi: 10.1177/0361198119850471.
- [13] E. Kidando, R. Moses, T. Sando, and E. E. Ozguven, "Evaluating Recurring Traffic Congestion using Change Point Regression and Random Variation Markov Structured Model," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2672, no. 20, pp. 63–74, Dec. 2018, doi: 10.1177/0361198118787987.
- [14] H. Nguyen, W. Liu, and F. Chen, "Discovering Congestion Propagation Patterns in Spatio-Temporal Traffic Data," *IEEE Trans. Big Data*, vol. 3, no. 2, pp. 169–180, Jul. 2016, doi: 10.1109/tbdata.2016.2587669.
- [15] S. S. Anjum *et al.*, "Modeling Traffic Congestion Based on Air Quality for Greener Environment: An Empirical Study," *IEEE Access*, vol. 7, pp. 1–24, 2019, doi: 10.1109/ACCESS.2019.2914672.

- [16] S. Amini, N. Motamedidehkordi, E. Papapanagiotou, and F. Busch, "Estimation of traversal speed on multi-lane urban arterial under non-recurring congestion," in *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings*, Aug. 2017, pp. 514–519, doi: 10.1109/MTITS.2017.8005726.
- [17] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transp. Res. Part C Emerg. Technol.*, 2018, doi: 10.1016/j.trc.2018.03.001.
- [18] R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," 2017, doi: 10.1109/YAC.2016.7804912.
- [19] H. Yi, J. Heejin, and S. Bae, "Deep Neural Networks for traffic flow prediction," 2017, doi: 10.1109/BIGCOMP.2017.7881687.
- [20] B. Sharma, S. Kumar, P. Tiwari, P. Yadav, and M. I. Nezhurina, "ANN based short-term traffic flow forecasting in undivided two lane highway," *J. Big Data*, vol. 5, no. 1, p. 48, Dec. 2018, doi: 10.1186/s40537-018-0157-0.
- [21] B. Gültekin Çetiner, M. Sari, and O. Borat, "A neural network based traffic-flow prediction model," *Math. Comput. Appl.*, vol. 15, no. 2, pp. 269–278, 2010, doi: 10.3390/mca15020269.
- [22] F. Canitez, P. Alpkokin, and S. T. Kiremitci, "Sustainable urban mobility in Istanbul: Challenges and prospects," *Case Stud. Transp. Policy*, vol. 8, no. 4, pp. 1148–1157, Dec. 2020, doi: 10.1016/j.cstp.2020.07.005.
- [23] T. T. Yaman, H. B. Sezer, and E. Sezer, "Modeling Urban Traffic by Means of Traffic Density Data: Istanbul Case," in *Advances in Intelligent Systems and Computing*, Jul. 2021, vol. 1197 AISC, pp. 867–874, doi: 10.1007/978-3-030-51156-2_100.
- [24] F. Canitez and M. Deveci, "An Integration Model for Car Sharing and Public Transport : Case of Istanbul," *Transist Istanbul Transp. Congr. Exhib.*, no. April, pp. 1–10, 2017, Accessed: Mar. 20, 2021. [Online]. Available: <https://www.researchgate.net/publication/324530842>.
- [25] "Istanbul Metropolitan Municipality Air Quality Station Information Web Service." <https://data.ibb.gov.tr/tr/dataset/hava-kalitesi-istasyon-bilgileri-web-servisi> (accessed Feb. 26, 2021).
- [26] D. F. Specht, "A General Regression Neural Network," *IEEE Trans. Neural Networks*, vol. 2, no. 6, pp. 568–576, 1991, doi: 10.1109/72.97934.
- [27] A. S. Ahmad *et al.*, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renewable and Sustainable Energy Reviews*, vol. 33. Elsevier Ltd, pp. 102–109, 2014, doi: 10.1016/j.rser.2014.01.069.
- [28] R. Grosse, "Lecture 5: Multilayer Perceptrons."
- [29] F. Wahid, R. Ghazali, A. S. Shah, and M. Fayaz, "Prediction of Energy Consumption in the Buildings Using Multi-Layer Perceptron and Random Forest," *Int. J. Adv. Sci. Technol.*, vol. 101, pp. 13–22, Apr. 2017, doi: 10.14257/IJAST.2017.101.02.
- [30] "Electrical load forecasting using support vector machines | IEEE Conference Publication | IEEE Xplore."
- [31] O. Nelles, "Classical Polynomial Approaches," *Nonlinear Syst. Identif.*, pp. 893–901, 2020, doi: 10.1007/978-3-030-47439-3_20.
- [32] I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, and F. Khoshbin, "GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs," *Eng. Sci. Technol. an Int. J.*, vol. 18, no. 4, pp. 746–757, Dec. 2015, doi: 10.1016/J.JESTCH.2015.04.012.