



Short term traffic speed prediction with RNN method for roads characterized by density-based clustering method

Murat Akın^{1,2*}, Şeref Sağıroğlu¹

¹Department of Computer Engineering, Gazi University, Ankara, 06570, Turkey

²Department of Research, Başarsoft Information Technologies Inc., Ankara, 06520, Turkey

Highlights:

- A novel traffic speed prediction model is presented
- Road characterization is created by density-based clustering
- Short term traffic speed prediction is carried out considering traffic patterns in near time

Keywords:

- Short term traffic speed prediction
- Traffic prediction model
- Intelligent transportation system
- Recurrent neural networks
- Density based clustering method

Article Info:

Research Article
Received: 19.04.2021
Accepted: 04.06.2021

DOI:

10.17341/gazimmfd.921035

Acknowledgement:

Authors thank TÜBİTAK, Başarsoft Inc. and Gazi BIDISEC for the invaluable support.

Correspondence:

Author: Murat Akın
e-mail:
muratakin@gazi.edu.tr
phone: +90 505 664 3810

Graphical/Tabular Abstract

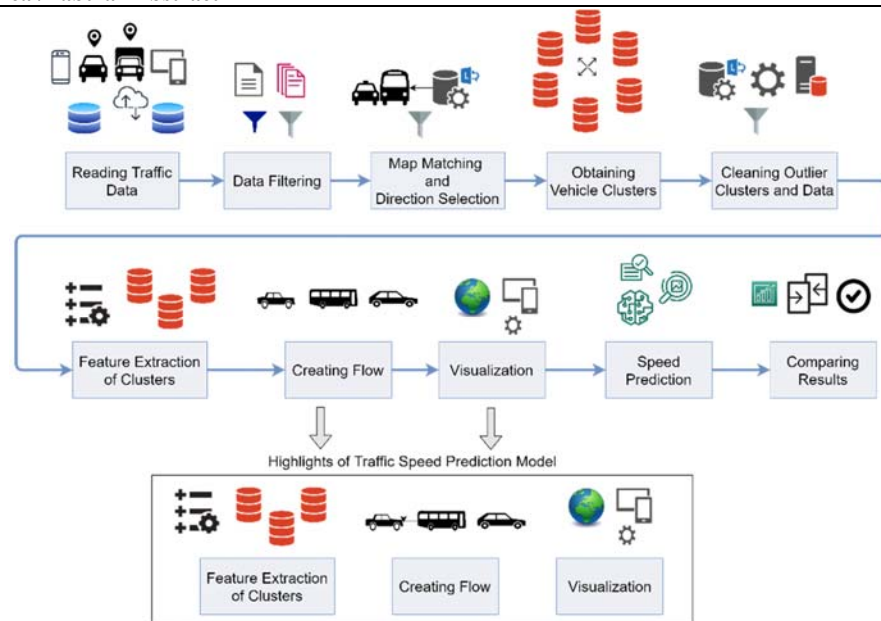


Figure A. General flow of traffic speed prediction model

Purpose: Different traffic parameters such as flow, density and speed are used for traffic management and these parameters are predicted within a road network. Studies focused on prediction with stationary data in a determined road, or distributed data to make long term prediction taking into account historical data. The purpose of study is to develop a short-term traffic speed prediction model that characterizes roads by density-based clustering methods and considers near time patterns of traffic movement.

Theory and Methods: This study presents a novel prediction model for traffic speed prediction consisting of nine stages. In the traffic model, data filtering and map matching methods were used to emplace GPS data into desired roads, density-based clustering method was applied to characterize roads and recurrent neural networks were employed for traffic speed prediction.

Results:

The proposed model using LSTM and GRU which of RNN variants was tested on Eskisehir and Istanbul Roads belonging to Ankara province. The error rates were measured as 8,595-8,656 km. from Eskisehir road and 7,331-7,955 km. from Istanbul road, respectively.

Conclusion:

The results have shown that proposed model yielded successful and satisfying outcomes in terms of stochastic and changeable traffic nature and this study offers new and alternative horizons to prediction studies in traffic management.



Yoğunluk tabanlı kümeleme yöntemiyle karakteristiği oluşturulan yollar için RNN yöntemi ile kısa zamanlı trafik hız tahmini

Murat Akın^{1,2*}, Şeref Sağıroğlu¹

¹Gazi Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği, 06570, Ankara, Türkiye

²Başarsoft Bilgi Teknolojileri A.Ş., AR-GE Birimi, 06520 Balgat Ankara, Türkiye

Ö N E Ç İ K A N L A R

- Yeni bir trafik hız tahmin modeli sunulmuştur
- Yoğunluk tabanlı kümeleme ile yol karakteristiği oluşturulmuştur
- Yakın zamandaki trafik örüntüleri dikkate alınarak kısa zamanlı trafik tahmini gerçekleştirilmiştir

Makale Bilgileri

Araştırma Makalesi
Geliş: 19.04.2021
Kabul: 04.06.2021

DOI:

10.17341/gazimmfd.921035

Anahtar Kelimeler:

Kısa zamanlı trafik hız tahmini,
trafik tahmin modeli,
akıllı ulaşım sistemleri,
tekrarlı sinir ağları,
yoğunluk tabanlı kümeleme yöntemi

ÖZ

Akıllı ulaşım sistemleri trafik akış, yoğunluk ve hız gibi parametreleri şehir trafiğini yönetmek için kullanılmaktadır. Bu çalışma, trafik hız tahmini yapan ve 9 aşamadan oluşan yeni bir tahmin modeli sunmaktadır. Sunulan modelde gerçek araç verileri, veri filtreleme ve harita eşleme işlemlerinden geçirilmiş, yoğunluk tabanlı kümeler oluşturulmuş, küme öznitelikleri üretilmiş, anlık trafik gösterimi yapılmış ve hız tahmini yapay sinir ağı RNN modeliyle gerçekleştirilmiştir. Daha önce yapılan çalışmalarda trafik hız tahmini sabit veri kaynaklarıyla belirli bir yolda veya dağıtık GPS kayıtlarıyla farklı günlerde yapılabilmekte iken, geliştirilen modelde yoğunluk tabanlı araç küme öznitelikleriyle ilgilenilen yola ait karakteristik özellikler oluşturulmuş, trafiğin kendi değişken yapısında kısa zamanlı ve veri odaklı hız tahmini yapılmıştır. Hız tahmini Ankara iline ait Eskişehir ve İstanbul yolu güzergâhlarında test edilmiş, tahmin için RNN modeli varyantı olan LSTM ve GRU modelleri kullanılarak hata oranları tespit edilmiş, Eskişehir yolu güzergâhında LSTM-GRU modelleri hata oranları sırasıyla 8,595-8,656, İstanbul yolu güzergâhında model hata oranları 7,331-7,955 olarak ölçülmüştür. Trafiğin değişken yapısı için geliştirilen modelle yakın gerçek zaman için başarılı sonuçlar elde edilmiştir. Önerilen modelin, gelecekte yapılacak olan trafik parametrelerinin tahmininde farklı ve yeni çözümler sunacağı, süreçleri hızlandıracağı ve kullanıcılara daha doğru ve hızlı hizmet verilmesine katkılar sağlayacağı değerlendirilmektedir.

Short term traffic speed prediction with RNN method for roads characterized by density-based clustering method

H I G H L I G H T S

- A novel traffic speed prediction model is presented
- Road characterization is created by density based clustering
- Short term traffic speed prediction is carried out considering traffic patterns in near time

Article Info

Research Article
Received: 19.04.2021
Accepted: 04.06.2021

DOI:

10.17341/gazimmfd.921035

Keywords:

Short term traffic speed prediction,
traffic prediction model,
intelligent transportation system,
recurrent neural networks,
density-based clustering method

ABSTRACT

Intelligent transportation systems use parameters such as traffic flow, density and speed to manage city traffic. This paper presents a novel prediction model for traffic speed prediction consisting of nine stages. In the presented model, real vehicle data were passed through data filtering and map matching processes, density-based clusters were created, cluster features were generated, instant traffic state was displayed, and traffic speed prediction was performed using with the artificial neural network RNN model. In previous studies, while traffic speed prediction can be performed on a specific road with stationary data sources or on different days with distributed GPS records, in the developed model, characteristics of the interested road are created by obtaining density-based vehicle cluster features, short-term and data-driven speed prediction is made within the changeable structure of traffic. Speed prediction was tested on Eskişehir and İstanbul roads belonging to Ankara province, the error rates were determined for speed prediction using the RNN variant LSTM and GRU methods, Eskişehir road LSTM-GRU error rates were measured as 8,595-8,656 and İstanbul road error rates as 7,331-7,955, respectively. The developed model for the changeable nature of traffic has yielded successful results in near real time. It is considered that the proposed model will offer different and new solutions in the prediction of traffic parameters, accelerate the processes and assist to the users more accurate and faster services.

1. GİRİŞ (INTRODUCTION)

Sanayileşmenin ve kentsel yaşam kalitesinin artmasından dolayı şehir merkezlerinde nüfus artışı hızlanmakta, büyük şehirler gelişmekte, yerleşim alanlarında kullanılan araç sayıları da artmakta, bu artışta doğal olarak beraberinde trafik ve ulaşım problemlerini getirmektedir. Dünya üzerindeki araç sayısı yaklaşık bir milyar civarındadır ve bu sayının önümüzdeki yirmi yılda iki katına çıkması beklenmektedir [1]. Artan araç sayısına bağlı olarak mevcut ulaşım altyapısı aynı oranda gelişmemektedir ve şehir trafiğinde araçlar sınırlı ulaşım ağı kullanmaktadır. Devlet ve kamu kurumları şehir içi ve şehirlerarası karayolu ulaşım ağını iyileştirmekte, gereklilik halinde yeni yollar yapmakta fakat ulaşım ağının geliştirilmesi artan araç sayısındaki talepleri karşılayamamaktadır. Doğal olarak, şehirlerdeki trafik yoğunluğu her geçen gün artmaktadır. Bunun sonucunda, ulaşım yönetimi ile ilgilenen kurumlar trafik sıkışıklığının kontrolünden, sürücü güvenliği ve çevresel etkiye kadar olan konulara çözüm üretmeye ve iyileştirmeye çalışmaktadır [2]. Trafik ekosisteminin iyileştirilmesi için güncel teknolojiler ulaşım kaynaklarının daha iyi organize edilmesi, araç trafiğinin yollara uygun şekilde dağıtımının ve planlanmasını sağlayacak sistemler oluşturulmasında kullanılmaktadır. Bu kapsamda geliştirilen sistemler farklı şekilde isimlendirilmekte olup bunlar içinde en öne çıkan ifade akıllı ulaşım sistemleridir. Akıllı ulaşım sistemleri (AUS) trafiğin etkin kullanımını sağlarken, trafik tıkanıklığı oluşmaması için çözümler üretmekte, yol altyapısını genişletmenin planlamasını yapmakta, trafik kazalarının oluşma olasılığını ve çevre kirliliğini azaltmaktadır [3]. Güncel teknolojileri kullanan AUS, şehri yöneten kurumların yönetim etkinliğini arttırmaktadır. Belediye ve devlet kurumları, bilgi ve iletişim araçları ile şehir sakinlerine hayatı kolaylaştıracak servisleri sağlarken aynı zamanda şehirleri izlemekte, istatistiki bilgi tutmakta, şehirle ilgili çıkarımlar yapabilmek için veriler toplamaktadır. Akıllı şehir uygulamaları için farklı veri toplama araçları mevcuttur ve veri kaynakları döngü detektörü, mikrodalga sensörü, basınç tüpleri, radar cihazları, video kamera sistemleri olarak öne çıkmaktadır [4]. Tüm bu araçlar etkin olarak trafik izlemekte kullanılırken bir yandan da maliyet getirmektedir. Trafiği izlemek için daha az maliyetli RFID teknolojisi ve kablosuz sensörler de kullanılırken [5, 6] bunun yanı sıra internetin oldukça yaygınlaşması ile katılımcılardan veri elde etme yöntemi ile mobil ve IoT cihazlardan da faydalanılmaktadır [7]. Gelişen teknolojiyle bahsedilen trafik kaynaklarının artması sayesinde trafik araştırmaları ve mühendisliği, ekonomik koşullar dikkate alınarak sorunsuz, güvenli trafik uygulamaları geliştirmek için trafik davranışlarını incelemektedir [8]. Trafiğin davranışlarını ve karakteristiğini anlamak, bir trafik ağını tanımlamak için trafik parametreleri tanımlanmış, temel parametreler olarak akış, yoğunluk ve hız olarak ifade edilmiştir [9, 10]. Trafik parametrelerinin tahmini için model tabanlı ve veri odaklı yöntemlerde, yol ağının bir örneği oluşturularak araçların ve sürücülerin davranışları analiz edilmektedir. Veri odaklı

yöntemlerde, elde edilen trafik verisi dikkate alınıp buna uygun yöntem uygulanarak trafik tahmini yapılmaktadır. Model tabanlı veya veri odaklı yöntemler, ayrıca tahmin yapılacak aralığın ifade edildiği kısa ve uzun zaman diliminde uygulanmaktadır [11, 12]. Farklı yöntemler ve sınıflandırmalar olsa da tüm trafik tahmin çalışmalarının odağı ve hedefi akıllı ulaşım sistemleri ve kullanıcılar için doğru tahmin sonuçları üretmektir.

Trafik tahmini, hava durumu, olaylar, yol kazaları, veri toplama gibi birçok dış sorundan etkilenen oldukça zorlu süreçleri içerisinde barındırır [13]. Günümüzde trafik tahmin uygulamaları geliştiren araştırmacılar için bol miktarda veri bulunsa da verinin lineer olmayan yapısı tahmin yapmayı zorlaştırmaktadır. Tahmin sürecini kolaylaştırmak ve doğruluğunu arttırmak için üretilecek trafik tahmin parametresine özgü veri türünü etkin yöntemlerle birleştirmek gerekmektedir. Mikrodalga sensör [14] ile döngü detektörü [15] genellikle trafik akışı için kullanılırken Küresel Konumlama Sistemi (Global Positioning System-GPS) verisi anomali [16] ve trafik tıkanıklığı tespitinde [17] kullanılmaktadır. Veri kaynaklarından her yola yerleştirilmesi maliyetli olan sensör ve detektörlerden elde edilen veriler ile belirli bir bölgede, sadece o yola ait tahmin yapılabilmektedir. Literatürde yapılan birçok çalışma sadece bir yol segmenti veya birbirini takip eden yol segmentleri üzerine odaklanmıştır [18]. Trafik bileşen tahmininde segment kavramı trafik yol ağında bulunan bir yol parçasını ifade derken, tek bir yol segmenti ve birbirini takip eden yollar üzerinde yapılan çalışmalar sadece o segmentten veri elde edilebilmesindedir. Bu çalışmada ise dağıtılmış GPS verisi kullanarak, trafik tahmini yola bağımlı olmayan, istenilen ve seçilen bir yol segmenti için yeni bir tahmin modeli geliştirilmiştir. Geliştirdiğimiz model veri odaklı, kısa zamanlı bir modeldir. Çalışmada, Başarsoft Bilgi Teknolojileri AŞ'den elde edilen Ankara iline ait GPS araç verileri kullanılmıştır. Araç konum verileri Ankara ili sınırları içerisinde öncelikle istenilen bölge için filtrelenmiş, geometrik şekillerden oluşan yol ağ bilgisi ile eşleştirilerek seçilen yola ait veriler elde edilmiştir. Filtreleme ve harita eşleme işlemlerinden sonra açılal yön bilgisi kullanarak aynı yönde hareket araç verileri ayıklanmış, araç verileri yoğunluk tabanlı kümelendirilmiştir. Kümeleme yapmaktaki temel amaç, konum verileri birbirine yakın araçları gruplayabilmek ve gruplanan araç bilgileri ile yola ait bir karakteristik oluşturabilmektir. Yol karakteristiğini oluşturulan kümeler, bir başlangıç noktası seçilerek döllemsel sıralanmış, Tekrarlı Sinir Ağı (Recurrent Neural Network-RNN) yöntemi ile yine belirlenen bölge için hız tahmini yapılmıştır. Böylelikle trafiğin kendi olasılığı içerisinde, yol karakteristiği oluşturularak, yüksek doğrulukta tahmin yapılması amaçlanmıştır. Bu çalışma daha önce yapmış olduğumuz çalışmanın da genişletilmiş bir çalışmasıdır [19].

Yapılan çalışmanın yenilikçi yönleri ve trafik hız tahmini alanına yapacağı katkılar aşağıda maddeler halinde sıralanmıştır.

- Trafik hız tahmini için dokuz aşamalı ve sıralı yeni bir model geliştirilmiştir. Mevcut verinin tahmin için hazırlanma aşamalarında veri ölçekleme, harita eşleme, yoğunluk tabanlı kümeler oluşturma, aykırı verilerinin temizleme işlemleri yapılırken, sonrasında elde edilen kümelerden öznitelik çıkarımı, akış oluşturma, görselleştirme ve tahmin aşamaları gerçekleştirilmektedir.
- Geçmiş çalışmalarda, tahmin yapılacak yollara yerleştirilen sabit veri kaynakları hem maliyet getirmekte hem de bu kaynaklarından elde edilen veriler ile sadece o yola ait tahminler üretilebilmektedir. Geliştirilen modelde büyük bir maliyet getirmeden, dağıtık GPS verisi kullanılarak istenilen yol için tahmin yapılmaktadır.
- Trafik hız tahmini segment uzunluğuna bağımlı olmadan, istenilen yol bölümünde, istenilen başlangıç-bitiş noktaları arasında yapılmaktadır.
- Belirlenen hedef yollar için araçlar yoğunluk tabanlı kümelenmiş, küme öznitelikleri oluşturularak anlık trafik hız renklendirmeleri yapılmış, yoğunluk tabanlı kümeler ile karakteristiği oluşturulan yollar için trafik hız tahmin doğruluğu artırılmıştır.

Makalenin 2. Bölümünde trafik hız tahmini üzerine literatürde yapılan çalışmalarda kullanılan yöntemler sunulmuştur. 3. Bölümde kullanılan veri kümesi ve uygulanan yöntemler açıklanmıştır. Yaptığımız çalışmanın modeli ve süreci 4. Bölümde anlatılmış, elde edilen sonuçlar 5. Bölümde açıklanmış ve tartışılmıştır. Son bölümde ise çalışma sonucu verilmiş ve kısa bir değerlendirme yapılmıştır.

2. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Trafik tahmini, üzerine uzun süredir çalışılan bir araştırma konusudur ve araştırmalar kısa zamanlı ve veri odaklı tahminler üzerine yoğunlaşmıştır [20]. Kısa zamanlı trafik uygulamaları birkaç dakikadan birkaç saate uzanan zaman dilimi içerisinde, mevcut ve geçmiş verilere dayanarak trafik akış, yoğunluk, hız gibi trafik parametreleri için tahmin üretmektedirler. Trafik parametrelerinin tahmininde trafik için toplanan veri kaynakları sabit ve hareketli veri olarak ayrılarak trafik uygulamalarına yön vermiştir [21]. Mevcut çalışmalar genellikle trafik akışı üzerinedir ve bu çalışmaların yapılmasının etkisi veri türünün farklılığıdır. Vloagianni vd. yaptıkları derleme çalışmasında 2014-2016 yılları arasında incelenen 63 çalışmanın 40 tanesi trafik akışı üzerine olup yoğunluk, hız ve seyahat süresi üzerine çalışmalar daha azınlıktadır [22]. Fakat trafik hızı üzerine yapılan çalışmaların yaygınlaşması önemlidir ve bu çalışmaların artmasıyla AUS, istenilen ve belirlenen yollar için etkin izleme gerçekleştirebilecek, kullanıcılara daha iyi yönlendirmeler yapabilecektir.

Trafik için hız tahmini yapılan çalışmalar kapsamında;

- Yol segmentlerinin ortalama hızını sinir ağı ile tahmini ve sonrasında Hidden Markov Model kullanarak her aracın yol/segment ortalama ilişkisinin istatistiksel gösterimi [23],

- Farklı yolların hız, zaman ve sürüş yönü bilgilerinin hava durumu bilgisiyle birleştirilerek LSTM, GRU ve SAE sinir ağları ile hafta içi ve hafta sonu durumlarının modellenmesi [24],
- Büyük ölçekte bir yol ağında trafik hız tahmini yaparken mekân-zaman örüntülerini çıkarmak için K-means ile PCA, tahmin için SVR yöntemini uygulama [25],
- Uygun veri kümesi kullanımı için yeni bir istatistiksel yöntem, Monte Carlo yöntemine dayalı üç aşamalı filtreleme işlemi ve karar faktörü ile her bir yol için günlük hız tahminlerini zaman serisi kullanarak hesaplama [26],
- İki mikrodalga sensör arasında ölçülen ve tek bir yol segmenti üzerinde elde edilen verilerin LSTM tabanlı modellenmesiyle yol segmentinde hız tahmini [27],
- Araç hızının sürücü davranışı, hız limiti, kavis, trafik akışı ve hava durumundan etkilediği göz önünde bulundurularak Yapay Sinir Ağı-Bulanık Mantık bir model ile farklı REST API'lerden elde edilen 94 araç verisi kullanarak beş farklı yol segmenti için hız tahmini [28],
- Araçlara ait GPS verileriyle L-U-Net ismi verilen YSA modeli ile hız tahmini ve geliştirilen model kullanarak araç hızına bağlı mekânsal ve zamansal özelliklerin çıkarımı [29], ve
- Pekin şehrini çevreleyen otoyolda iki yol segmenti üzerinde hareket eden lojistik araçların GPS verisini kullanarak ilgili segmentlere ait farklı zaman dilimlerinde hız tahmini [30] yapılmıştır.

Yoğunluk tabanlı kümelerin kullanıldığı çalışmalarda;

- Trafik tıkanıklığını tespit etmek adına DBSCAN yöntemi ile trafik yoğunluğu [31], ve
- FN-DBSCAN yöntemi kullanarak izlenen yolun GPS verileri ile tıkanıklık seviyesini ölçme [17] bunlardan bazılarıdır.

İncelenen çalışmalarda görülebileceği üzere mikrodalga sensörü için zorunlu segmentlerde çalışmalar yapılmış, GPS verisi kullanan çalışmalarda ise belirli bir yol segmenti seçilerek bu segmentin farklı günlerinde aynı zaman dilimi ele alınarak hız tahmini yapılmıştır. Bu makale kapsamında önerilen modelde ise istenilen yol segmentinde yol karakteristiği oluşturmak için yoğunluk tabanlı kümeler kullanılmıştır ve sonrasında RNN ile aynı gün içerisinde ve kısa bir zaman dilimi için hız tahmini yapılmıştır. Trafik hız tahmini için geliştirilen modelde kullanılan yöntemler takip eden başlıkta anlatılmıştır.

3. MATERYAL VE METOT (MATERIAL AND METHOD)

Kısa zamanlı ve veri odaklı tahmin model açıklamadan önce kullanılan verinin tanımlanması ve verilerin seçilmeden önce geçirilmesi gereken işlemler ve kullanılan metotlar bu bölümde açıklanmıştır.

3.1. Kullanılan Veri Kümesi (Used Data Set)

Bu çalışmada, Ankara şehrine ait 15 Ocak 2020/15:00-16:00 tarih-zaman aralığında bir saatlik "gerçek araç verileri"

kullanılmıştır. GPS sinyali içeren gerçek araç verileri herhangi bir filtreleme işlemi uygulanmadan 231 MB boyutunda olup yaklaşık 3 milyon 800 bin kayıt içermektedir. GPS sinyalleri farklı periyotlarda kaydedilmektedir ve kaydedilen verilerin belirli bir frekansı yoktur. Araçların ürettiği sinyaller için altı öznitelik mevcuttur. Araç tarafından üretilen öznitelikler ile bunların açıklamaları Tablo 1’de ve araçlara ait GPS kayıtlarının örnek bir gösterimi Tablo 2’de verilmiştir.

3.2. OPTICS (Ordering points to identify the clustering structure)

OPTICS genişletilmiş bir DBSCAN algoritması olarak çalışmaktadır [32]. DBSCAN algoritmasında rastgele seçilen bir p noktası çekirdek nesne olarak seçilmektedir ve girdi parametresi olarak çekirdek nesnenin komşu nesnelere yakınlık yarıçapı ε (epsilon) ve kaç üye komşuluğuna bakılacağı $MinPts$ değerlerini kullanılmaktadır. OPTICS algoritmasında ise “üretilen uzaklık” ε değerinden küçük ε_i değerleri dikkate alınmaktadır. OPTICS algoritmasının DBSCAN algoritmasından farkı küme üyeliklerinin ilk etapta belirlenmemesidir. Küme üyelerinin hemen belirlenmesi yerine her bir nesnenin sırası oluşturulmakta, çekirdek mesafe ve uygun bir erişilebilirlik mesafe bilgisi kaydedilmektedir. Veri sırası elde edilen nesnelere çekirdek nesne olarak belirlenen nesnelere için en az sayıda yakın komşuluğuna bakılarak kümeleme işlemi yapılmaktadır. Eş. 1 ile çekirdek mesafesi belirleme eşitliği görülmektedir. Bir p noktasının ε uzaklığındaki komşuluğu $N_\varepsilon(p)$ ile ifade edilirken, $MinPts$ değerinden küçük ise tanımsız kabul edilmiştir. Eğer bu uzaklık en az noktaya eşit ve büyük ise komşu noktalara uzaklıklar $MinPts$ -mesafe olarak ele alınmış ve an yakın komşu eleman sonraki çekirdek nokta seçilmiştir [33].

$$\begin{cases} \text{Eğer yakınlık } (N_\varepsilon(p)) < MinPts, & \text{Tanımsız} \\ \text{Diğer durumlarda} & , MinPts - mesafe(p) \end{cases} \quad (1)$$

3.3. RNN (Recurrent Neural Network)

RNN yöntemi sıralı veriler için dahili hafızayı kullanarak bir tekrar oluşturan güçlü bir derin sinir ağıdır [34]. RNN yönteminde hafıza hücreleri bir zincir oluşturmaktadır ve giriş dizisindeki her eleman için aynı görevi gerçekleştirmektedir. Yöntemin güçlü olması giriş ve çıkış değerleri arasındaki eşleşmeyi doğrudan öğrenmesinden ileri gelmektedir. Fakat RNN yöntemi eğitimin yok olması (vanishing gradient) ve çok yüksek değerlere çıkması (exploding gradient) problemlerinden etkilenmektedir ve sorunu çözmek için Uzun Kısa Zamanlı Hafıza (Long Short Term Memory-LSTM) yöntemi önerilmiştir [35, 36]. LSTM yönteminde hücreler kendi durumlarını zamana karşı korurken uzun zaman bağımlılıklarını da öğrenmektedir. Giriş elemanları $x = (x_1, x_2, x_3, \dots, x_n)$ olan modelde t zaman aralığında LSTM input(i), forget(f) ve output(o) kapılarında oluşmaktadır ve kapılara ait matematiksel ifadeler Eş. 2-4 ile görülmektedir. Eşitliklerde σ aktivasyon fonksiyonunu, W ağırlık matrislerini, b ilgili bias değerini ifade etmektedir. Eş. 5 ile hücre durumu hesaplanırken, sonraki hücreye giriş değeri olacak ve öğrenmeyi sağlayacak gizli katman Eş. 6’da tanımlanmıştır.

$$i_t = \sigma(W_{iX_t} + W_{ih_{t-1}} + b_i) \quad (2)$$

$$f_t = \sigma(W_{fX_t} + W_{fh_{t-1}} + b_f) \quad (3)$$

$$o_t = \sigma(W_{oX_t} + W_{oh_{t-1}} + b_o) \quad (4)$$

Tablo 1. Araçlar için üretilen öznitelikler ve açıklamaları (Extracted features and definitions for vehicles)

Öznitelik sayısı	Araç tarafından üretilen öznitelikleri	Açıklama
1	Araç ID	Navigasyon cihazını kullanan araçlara ait benzersiz bir numara
2	Zaman Damgası	Tarih ve saat formatında olan zaman bilgisi
3	Y	Boylam Koordinatı
4	X	Enlem Koordinatı
5	Araç Hızı	Hız bilgisi aracın anlık ölçülen hızı
6	Açısal Yön	Açısal Yön bilgisi 0^0 kuzey yönünü gösterecek şekilde aracın koordinat düzlemindeki yönü

Tablo 2. Araç GPS kayıtlarının örnek bir gösterimi (Sample of vehicle GPS records)

Araç ID	Zaman Damgası	Y	X	Araç Hızı	Açısal Yön
fce*****	15.01.2020 15:00	39,954361	32,571415	35	40
fcf*****	15.01.2020 15:00	39,902203	32,765202	50	171
84C*****	15.01.2020 15:00	39,920960	32,993813	101	236
129*****	15.01.2020 15:00	39,894039	32,803215	49	270
DA1*****	15.01.2020 15:01	39,949560	32,597176	33	279
cb4*****	15.01.2020 15:01	39,960018	32,686771	65	322
cb5*****	15.01.2020 15:01	39,802967	32,393925	57	202

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{cx_t} + W_{ic_{t-1}} + b_c) \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

Öğrenme süreci hızlı olan ve eğimin yok olması problemine karşı başarılı başka bir yöntem Tekrarlı Kapı Birimi (Gated Recurrent Unit - GRU) ayrı hafıza hücresi bulundurmayan LSTM farklı bir türüdür. GRU basitleştirilmiş LSTM olarak ifade edilebilir ve LSTM gibi üç kapıya değil güncelleme kapısı (z) ve reset kapısı (r) olarak iki kapıya sahiptir.

3.4. Hata Metrikleri (Error Metrics)

Trafik parametrelerinin tahmini için geliştirilen uygulamaların sonuçlarını değerlendirmek için geleneksel hata göstergeleri olan RMSE(Root Mean Square Error), MAPE(Mean Absolute Percentage Error) ve MAE(Mean Absolute Error) hata metrikleri kullanılmaktadır [37]. Hata oranları, n elemanlı bir dizi için her tahmin değerinden (\bar{y}_i) ile gerçek değer(y_i) arasındaki fark değerlerini dikkate alınarak hesaplanmaktadır. RMSE, MAPE ve MAE hesaplama eşitlikleri sırasıyla Eş. 7-9'da ifade edilmiştir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (7)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\bar{y}_i - y_i|}{y_i} \quad (8)$$

$$MAE = \sum_{i=1}^n \frac{|\bar{y}_i - y_i|}{y_i} \quad (9)$$

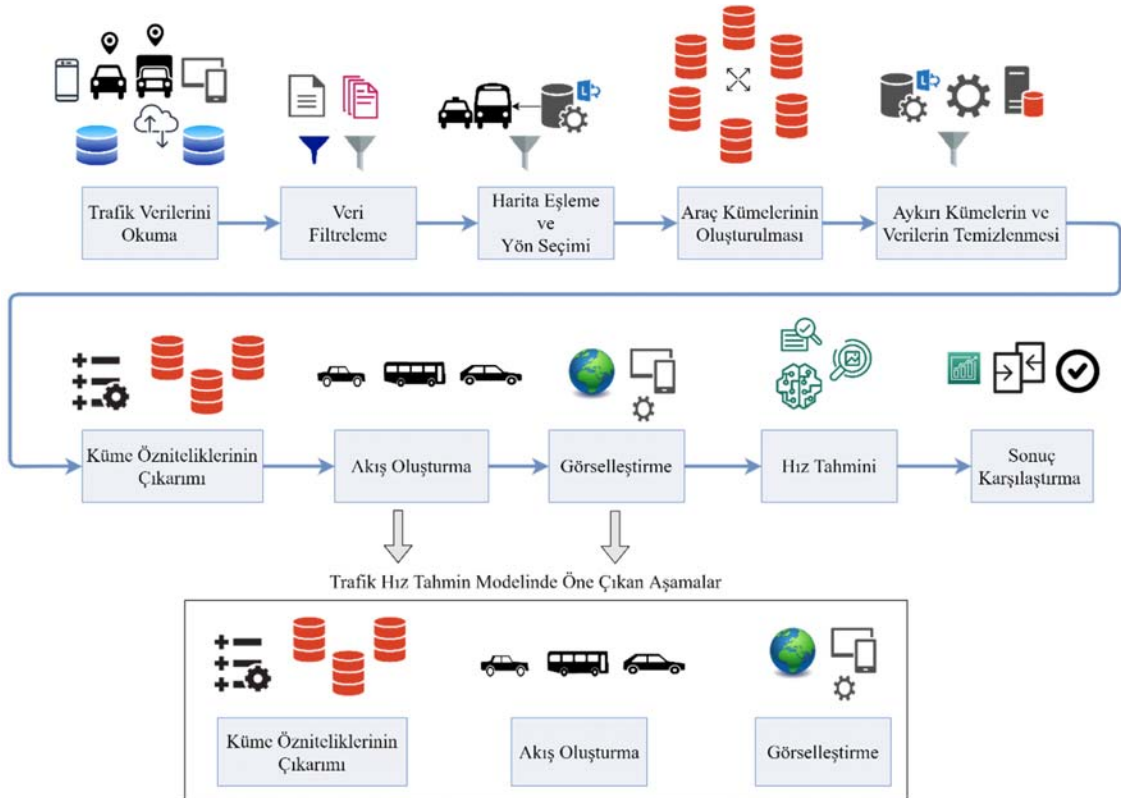
3.5. Haversine uzaklığı (Haversine distance)

Haversine uzaklığı, küre bir düzlem üzerindeki iki nokta arasındaki en büyük daire mesafesini hesaplamak için kullanılmaktadır [38]. Dünyanın şekli kutuplardan basık bir küre olduğundan, GPS noktaları arasındaki mesafe Haversine uzaklığı ile hesaplanmaktadır ve bu hesaplama Eş. 10'da verilmiştir [39]. Formülde, uzaklık hesabı "d", dünya yarıçapı "r", enlem değerleri θ , boylam değerleri φ ile gösterilmiştir.

$$d = 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\theta_2 - \theta_1}{2} \right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2 \left(\frac{\varphi_2 - \varphi_1}{2} \right)} \right) \quad (10)$$

4. GELİŞTİRİLEN MODEL (DEVELOPED MODEL)

Geliştirilen tahmin modeli dokuz aşamadan ve birbirini tamamlayan sıralı metotlardan oluşmaktadır. Başlangıç aşamaları filtreleme, harita eşleme, küme oluşturma ve aykırı veri temizleme gibi verilerin tahmin için hazırlıklarını içerirken sonraki aşamalar öznitelik üretimi, akış oluşturma, görselleştirme, hız tahmini ve doğrulama adımlarını kapsamaktadır. Geliştirilen trafik tahmin modelinin genel yapısı ve akışı Şekil 1'de gösterilmiş ve takip eden başlıklarda her bir aşama açıklanmıştır.



Şekil 1. Kısa zamanlı trafik hız tahmin model süreci ve öne çıkan aşamalar (Short term traffic speed prediction and highlights)

4.1. Trafik Verilerini Okuma (Reading Traffic Data)

Çalışmada kullanılan trafik verileri 3 milyon 800 bin satırdan oluşmakta ve 231 Megabayt boyutundadır. Veriler, araçların ID bilgisi, zaman damgası, enlem-boylam bilgisi, anlık hız değeri ve açılal yön bilgisinden oluşmaktadır ve ham veri Python-Pandas kütüphanesi kullanılarak okunmuş ve veri çerçevesine aktarılmıştır.

4.2. Veri Filtreleme (Data Filtering)

Elde edilen ve okunan veriyi istenilen yola yerleştirmek ve tahmin yapmak için trafik verisi belirli filtreleme ve harita eşleme işlemlerinden geçirilmiştir. Çalışmamızda Ankara iline ait verilere öncelikle hedef bölge olan İstanbul yolu ile Eskişehir yolu güzergâhları için filtreleme işlemleri uygulanmıştır. Filtreleme işlemi uygulanırken iki farklı Enlem ve iki farklı Boylam noktası seçilmiş ve araç verileri bir dörtgen arasında kalmıştır.

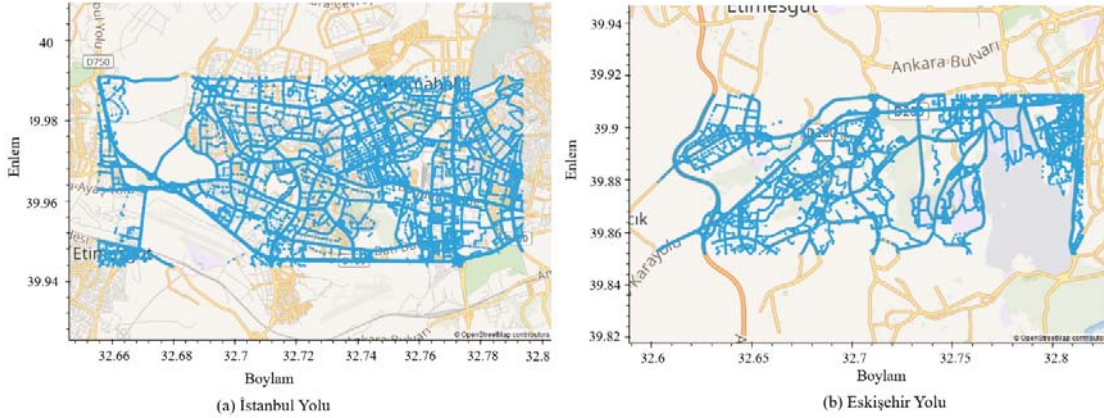
Eskişehir yolu için;

- 32,815056° K, 32,603345° K Enlem noktaları ve 39,851896° D, 39,912720° D Boylam noktaları,

İstanbul yolu için;

- 32,793608° K, 32,655502° K Enlem noktaları ve 39,991057° D, 39,943762° D Boylam noktaları seçilmiştir.

Kısa zamanlı bir tahmin yürütüleceği için yoğun trafiğin olduğu zaman dilimi 15 Ocak 2020 15:10 ve 15 Ocak 2020 15:20 arası olarak belirlenmiş ve konum filtresinden geçirilen veriler bu zaman dilimine çekilmiştir. Şekil 2’de filtrelenen araç verileri görülmektedir.



Şekil 2. Filtrelenen araç verilerine ait konum noktaları (Location points of filtered vehicle data)

ID	ADI	GEOMETRY
60086652	Eskişehir Yolu	LINestring (32.78312 39.90847, 32.78326 39.90848)
60349877	Eskişehir Yolu	LINestring (32.71574 39.90578, 32.72120 39.90599)
60038219	Eskişehir Yolu	LINestring (32.62835 39.86292, 32.62147 39.86035)
60178416	Eskişehir Yolu	LINestring (32.75085 39.90737, 32.74946 39.90732)
60210126	Eskişehir Yolu	LINestring (32.74969 39.90740, 32.74657 39.90726)
60160378	Eskişehir Yolu	LINestring (32.78561 39.90860, 32.78637 39.90864)
60036053	Eskişehir Yolu	LINestring (32.64494 39.87201, 32.64431 39.87157)
60053019	Eskişehir Yolu	LINestring (32.62147 39.86035, 32.61999 39.85979)

(a) SHP dosya bilgileri

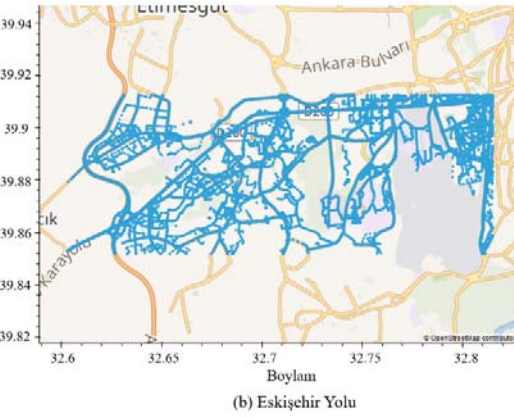
4.3. Harita Eşleme ve Yön Seçimi (Map Matching and Direction Selection)

Hedef bölge ve tahmin aralığı oluşturulduktan sonra ilgili araç verileri için harita eşleme işlemi gerçekleştirilmiştir. Harita eşleme işlemi için geometrik şekillerden oluşan bir Shape (SHP) dosyası kullanılmış, her bir GPS noktası tampon oluşturularak seçilen SHP dosyası ile eşleştirme işlemi yapılmıştır. Eşleme işleminde SHP dosyalarının mekân indeksini kullanılmıştır. Eskişehir yolu için örnek bir SHP dosyası Şekil 3’te verilmiştir.

Bu makale çalışmasında geliştirilen trafik hız tahmin modelinde, tahmin yapılacak yol için harita eşleme işlemi bir ön işlemdir. Bu işlem Ankara ili Eskişehir ve İstanbul Yolları için yapılmıştır ve bu yollar iki yönlüdür. Yola eşlenen araç verileri iki yönde hareket etmesinden dolayı aynı yönde giden araçlar için açılal yön dikkate alınarak ekstra bir seçim işlemi gerçekleştirilmiştir. Kuzey yönü açılal yönü 0° olacak şekilde, Eskişehir yolu doğu ve kuzey-doğu yönünde hareket eden araçların açılal yönü 0-90° arası, İstanbul yolu için batı ve kuzey-batı yönünde hareket eden araçların açılal yönü 240-359° arasında olacaktır. Bu değerler dikkate alınarak açılal yön filtreleme uygulandığında Eskişehir Yolu batı-doğu yönünde, İstanbul Yolu doğu-batı yönünde giden araçlar seçilmiştir.

4.4. Araç Kümelerinin Oluşturulması (Obtaining Vehicle Clusters)

Harita eşleme işlemi uygulandıktan ve hedef yol için yönlü araçlar seçildikten sonra enlem ve boylam bilgileri dikkate alınarak yoğunluk tabanlı kümeleme yöntemi OPTICS ile



(b) SHP dosyasının çizimi

Şekil 3. Eskişehir Yolu için SHP dosya bilgileri ve çizimi (SHP file information and drawing for Eskişehir Road)

araç verileri kümelendirilmiştir. Kümeleme işlemi için minimum araç sayısı 25 olarak belirlenmiş ve bu sayı isteğe bağlı olarak değişmektedir. DBSCAN yerine OPTICS yönteminin kullanılması ile küme için sadece eleman sayısının verilmesi ve herhangi bir minimum mesafe belirlenmemiştir. Minimum mesafenin belirlenmemesi ile araç verileri kendi olasılığı içinde kümelenecek ve yol karakteristiğini oluşturacak araçlar belirlenecektir.

4.5. Aykırı Kümelerin ve Verilerin Temizlenmesi (Cleaning Outlier Clusters and Data)

Yoğunluk tabanlı kümeleme işlemi sonucunda kümelenemeyen GPS kayıtları -1 değeri olarak işaretlenmektedir ve kümelenemeyen bu kayıtlar veriden temizlenmiştir. Ayrıca kümeleme işlemi yapıldıktan sonra kavşaklarda GPS sinyal bilgisi üreten araçların farklı yönde ve yolda gittiği tespit edilmiş ve Açılmal Yönl (AY) için standart sapma değeri ile aykırı araçlar kümelere çıkarılmıştır. Küme içindeki araç sayısı "n" değerini göstermek üzere Standart Sapma ve AY değeri ortalama hesabı Eş. 11, Eş. 12'de gösterilmiştir.

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (AY_i - \overline{AY})^2} \quad (11)$$

$$\mu = \frac{1}{n} \sum_{i=1}^n AY_i \quad (12)$$

4.6. Küme Özniteliklerinin Çıkarımı (Feature Extraction of Clusters)

Öznitelik üretimi veri boyutunu azaltma ve uygulama tahmin performansını iyileştirmek gibi iki ayrı amaç için gerçekleştirilmektedir [40]. Geliştirilen tahmin modelinde yol karakteristiği oluşturmak için her bir küme için öznitelikleri üretilmiştir. Üretilen küme öznitelikleri Tablo 3 ile listelenmiş ve hesaplama eşitlikleri gösterilmiştir.

4.7 Akış Oluşturma (Creating Flow)

Trafik hız tahmin modelimiz için akış oluşturma, belirlenen bir başlangıç noktası ve küme merkezinin bu noktaya

uzaklığı ile belirlenmektedir. Bu işlem ile başlangıç noktasından itibaren oluşan kümeler yakınlığına göre sıralanmaktadır. Eskişehir yolunda trafikte doğu yönünde hareket eden araçlar için 32,603345 K - 39,851896 D (Yaşamkent bölgesi), İstanbul yolu için trafikte batı yönünde hareket eden araçlar 32,655502 K - 39,991057 D (Çiftlik Kavşağı) başlangıç noktaları seçilmiştir. Küme merkezlerinin başlangıç noktalarına uzaklığı Haversine uzaklığı ile hesaplanmıştır.

4.8. Görselleştirme (Visualization)

Akış oluşturulan ve başlangıç noktasına yakınlığına göre sıralanan kümelerin hızları oluşan küme bölgelerinde trafiğin anlık hızını oluşturmuş, trafiğin anlık hız gösterimi için renklendirme yapılmış ve harita üzerinde görselleştirilmiştir. Anlık trafik hız gösterimi için farklı renkler kullanılmıştır. 60 km/s hızı ve üzerindeki değerler için trafiğin akıcı olduğunu belirtmek için mavi renk seçilmiştir. Daha yoğun trafiği göstermek için sırasıyla; 0-25 km/s kırmızı, 25-45 km/s kahverengi, 45-60 km/s turuncu renk ile gösterilmiştir.

4.9. Hız Tahmini ve Sonuçları Karşılaştırma (Speed Prediction and Comparing Results)

Geliştirilen tahmin modelinin trafik hız tahmini aşamasında RNN varyantı olan LSTM ve GRU modelleri kullanılmış, veriler eğitim ve test olarak ikiye ayrılmış, test verileri kullanılarak tahmin yapılmıştır. Tahmin sonucu üretilen hız değerleri ve orijinal hız değerlerinin karşılaştırılması, RMSE, MAPE ve MAE hata fonksiyonları kullanılarak yapılmıştır. Tahmin modeli için küme özniteliklerinden küme hızı girdi olarak kullanılmış ve çıktı olarak yine küme hızı elde edilmiştir. Tahmin modelinde uygulanan sinir ağının modeli giriş ve çıkış katmanları hariç Şekil 4'de gösterilmiştir.

5. DENEYSEL SONUÇLAR (EXPERIMENTAL RESULTS)

Bu makale kapsamında, önerilen model ve bu modelin testlerine ait çalışmalar bu bölümde açıklanmıştır. Modellerin test edilmesi için öncelikle, trafik hız tahmin modelimizin başlangıç aşamalarında ham veri okunmuş,

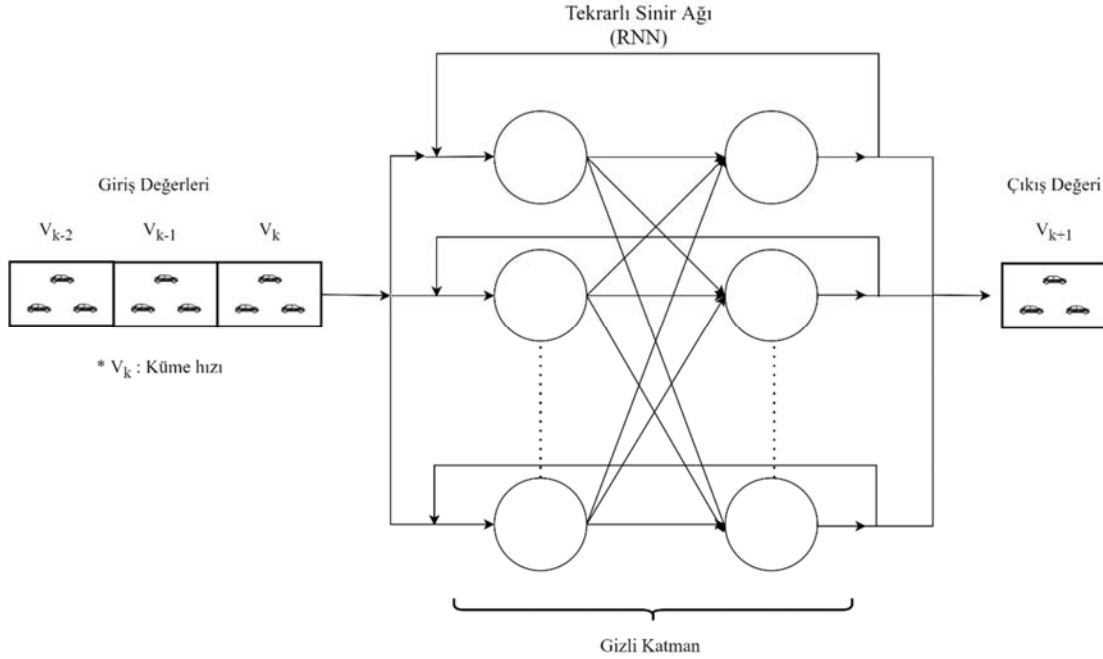
Tablo 3. Küme Öznitelikleri (Cluster Features)

Küme Öznitelikleri	Küme Öznitelik Çıkarımında Kullanılan Hesaplamalar
Küme Araç Sayısı (k_s)	Küme içerisinde her bir aracının ürettiği sinyal toplamı ($\sum_{i=1}^n a_i$)
Küme Hızı (k_v)	Her aracın v hızı olmak üzere toplam hızın küme araç sayısına oranı ($\frac{\sum_{i=1}^n v}{k_s}$)
Ön Araç X	$\text{Min}(X_{\text{Küme}})$ Batıdan Doğuya hareket eden araçlardan minimum X (Boylam) değeri
Ön Araç Y	$\text{Min}(X_{\text{Küme}})$ değerine sahip aracın Y (Enlem) değeri
Arka Araç X	$\text{Max}(X_{\text{Küme}})$ Doğudan Batıya hareket eden araçlardan minimum X (Boylam) değeri
Arka Araç Y	$\text{Max}(X_{\text{Küme}})$ değerine sahip aracın Y (Enlem) değeri
Küme Uzunluğu (k_u)	Ön araç ve Arka Araç koordinatları arası Haversine uzaklığı (metre)
Küme Yoğunluğu (k_y)	Küme uzunluğunun küme araç sayısına oranı ($\frac{k_s}{k_u}$)

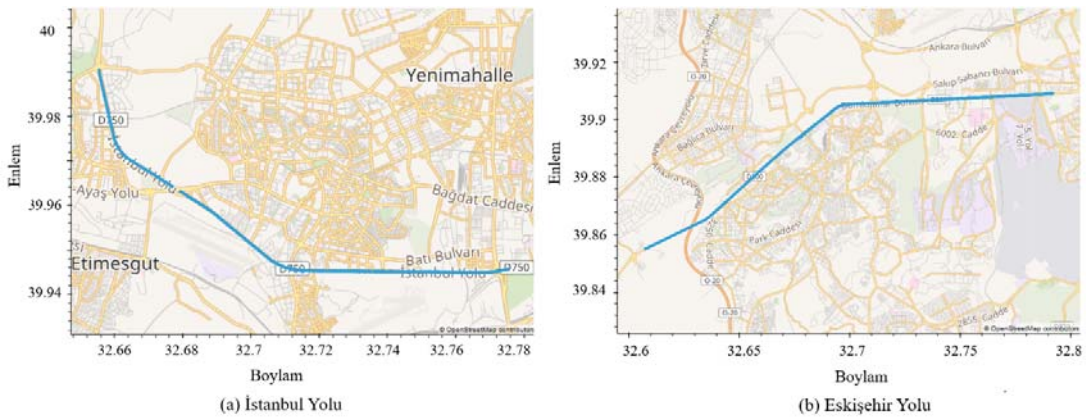
belirlenen bölgeye filtrelenmiş ve yol ağına eşleştirilmiş, yoğunluk tabanlı kümeler elde edilerek veri temizleme işlemi yapılmıştır. Veri temizleme işlemi sonrasında elde edilen kümelerin özniteliklerinin üretilmiş ve bir başlangıç noktasına göre kümeler sıralanarak bir akış oluşturulmuştur. Bu işlemlerle yola ait karakteristik oluşturmuş ve sonrasında küme öznitelikleri ile önce görselleştirme yapılmış ve sonrasında küme hız değerleri LSTM ve GRU ağları için girdi sağlamış ve trafik hız tahmini yapılmıştır. Modelimizde kullanılmak üzere Ankara iline ait bir saatlik veri 10 dakikalık zaman dilimine (15:10-15:20) indirgenmiştir ve kısa zamanlı bir tahmin gerçekleştirilmiştir. Tahmin yapılacak yol için Eskişehir yolu için batı-doğu yönünde ve İstanbul yolu için doğu-batı yönünde giden araçlar seçilmiştir. Filtreleme ve ön işlem yapılmayan Ankara verisi yaklaşık 3,8 milyon GPS kaydı içerirken, zaman indirgeme, yol eşleme işlemi ve aynı yönde geçen araçlar seçimi sonrası GPS noktası her iki yol için 6 bin civarına düşmüştür. Detayları 4.2 ve 4.3 bölümünde anlatılan veri filtreleme ve harita eşleme işleminin sonuçları Şekil 5’de harita üzerinde

görselleştirilmiştir. Her iki yola yerleştirilen araç verileri yoğunluk tabanlı kümeleme yöntemi ile kümelendi ve kümeleme işlemi minimum araç sayısı olarak 25 araç seçilerek, İstanbul yol güzergâhı için 73 ve Eskişehir yol güzergâhı için 80 küme oluşmuştur. Oluşan kümelerin harita üzerinde gösterimi Şekil 6’da görülmektedir.

Kümeleme işlemi sonrasında tahmin modelimiz için öznitelik üretim işlemi gerçekleştirilmiştir. Her küme için öznitelikler; küme içindeki araç sayısı, küme araçlarının ortalama hızı, küme içerisindeki en ön ve arkadaki aracın enlem ve boylam bilgileri, metre cinsinden küme uzunluğu ve küme içerisindeki küme araç yoğunluğu olarak üretilmiştir. Eskişehir yol güzergâhı kümeleri içerisinde en fazla araç sayısına 99 araç ile 23. kümedir. Bu sayı İstanbul Yol güzergâhı için 81 araç ile 62. kümedir. Küme yoğunluğunun en fazla olduğu küme Eskişehir yolu için 0. küme, İstanbul yol güzergâhı için 72. küme ve küme yoğunluğunun en fazla olduğu bu iki küme için hız ortalamaları sırasıyla 17,7 ve 14,91 km/h’dir. Araç sayısının



Şekil 4. Trafik hız tahmini için uygulanan sinir ağı modeli (Implemented neural network model for traffic speed prediction)



Şekil 5. Yol ağ bilgisi ile eşlenen noktaların görünümü (View of points matching with road network information)

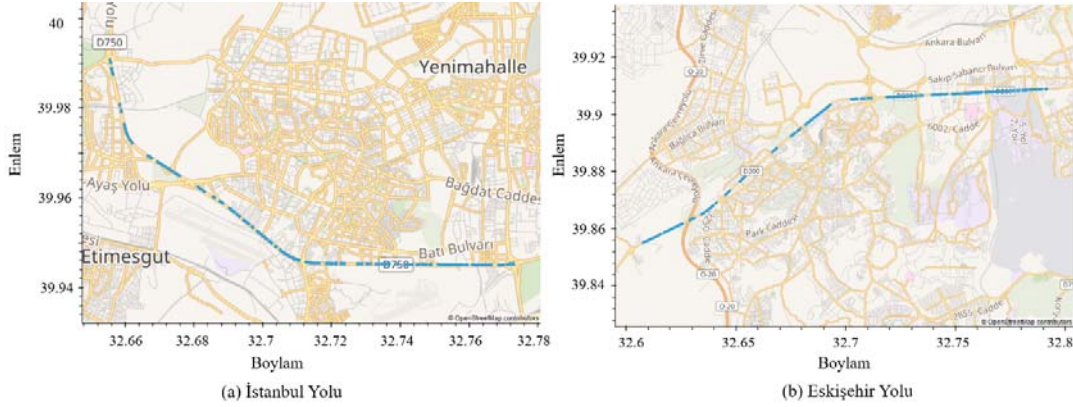
ve yoğunluğunun artması bize hız değerlerinin düştüğünü göstermektedir. Eskişehir yoluna ait kümelerle ilgili özet bilgiler Tablo 4 ile görülmektedir.

Küme özet bilgileri üretimi sonrasında belirlenen başlangıç noktasına göre Haversine uzaklığı ile kümeler sıralanmış ve akış oluşturulmuştur. Her iki yol için başlangıç noktasına en yakın 5 küme ve metre uzaklık değerleri Tablo 5 ile gösterilmiştir. Akış oluşturulan kümeler ve bu kümelerle ilgili özet bilgiler kullanılarak harita üzerinde anlık trafik hız gösterimi yapılmıştır. Yoğunluk haritalarında mavi renk 60 km/s üzerindeki hızı gösterirken, koyu renkler daha düşük hızda yoğun trafiği göstermektedir. Harita üzerinde anlık trafik hızı kümelerle ilgili ön ve arka araç değerleri kullanılarak çizdirilmiş, İstanbul Yolu ve Eskişehir Yolu güzergâhları için anlık hız trafiğinin gösterimi Şekil 7'de gösterilmiştir.

Tablo 5. Başlangıç noktalarına göre en yakın 5 küme ve uzaklık değerleri (The 5 Closest Clusters and Distance Values According to the Starting Points)

Eskişehir Yol Güzergâhı		İstanbul Yol Güzergâhı	
Küme No	Başlangıç Noktasına Uzaklık (km)	Küme No	Başlangıç Noktasına Uzaklık (km)
79	0,685643	72	2,320122
78	1,007149	71	2,390518
77	1,376619	70	2,470207
76	1,801005	69	2,513389
75	2,123413	68	2,659491

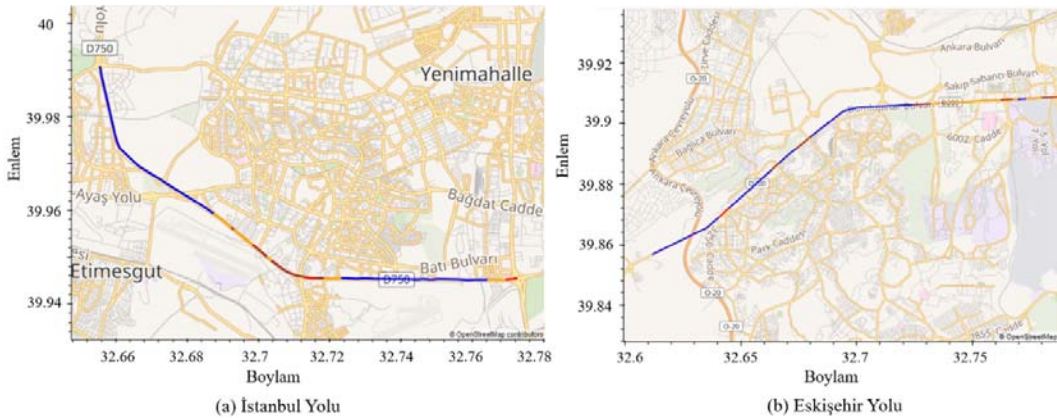
Şekil 4'de verilen örnek model temel alınarak testler yürütülmüştür. Kümelerle ilgili küme hızı değerleri LSTM ve GRU modelleri için giriş değerleridir. Her iki yolda oluşan kümeler için %80 eğitim verisi seçilmiş, %20 küme sayısı ile



Şekil 6. Elde Edilen Yoğunluk Tabanlı Kümeler (Obtained density based clusters)

Tablo 4. Eskişehir yol güzergâhı için örnek küme özet bilgileri (Samples of Cluster Features for Eskişehir Road)

Küme No	Küme Araç Sayısı	Küme Hızı	Ön Araç X	Ön Araç Y	Arka Araç X	Arka Araç Y	Küme uzunluğu(m)	Küme Yoğunluğu
23	99	30,9091	32,7791	39,9083	32,7785	39,9083	74,4719	1.32936
16	76	38,3026	32,7735	39,9081	32,7728	39,908	70,728	1,07454
14	41	37,2195	32,7719	39,9081	32,7717	39,9081	25,6613	1,59773
70	46	75,3478	32,6396	39,8681	32,6381	39,867	190,749	0,241155
0	50	17,7	32,7639	39,9077	32,7637	39,9076	24,5476	2,03686

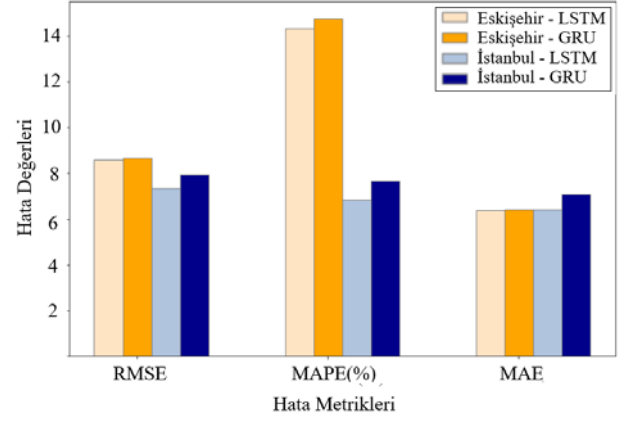


Şekil 7. Anlık Trafik Hız Gösterimleri (Instant display of traffic speed)

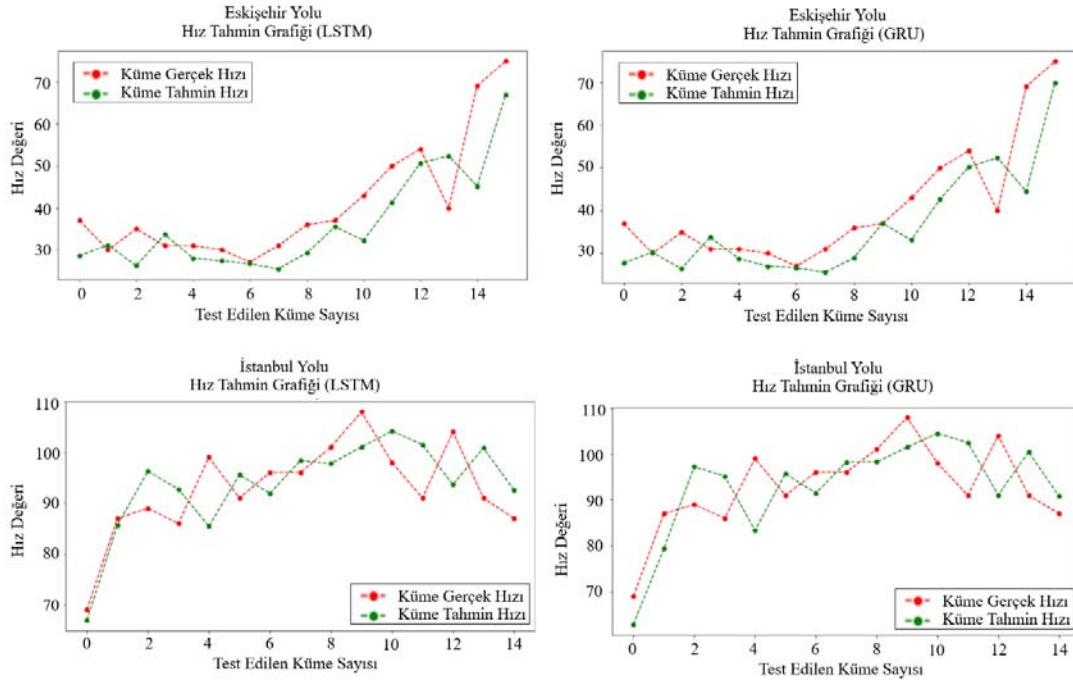
modeller test edilmiştir. LSTM-GRU sinir ağları modelleri için 2 katman oluşturulmuş, birinci katmanda 128 nöron, ikinci katmanda 64 nöron kullanılırken, optimizasyon için “Adam” algoritması seçilmiş ve yığın boyutu 32 olarak belirlenmiştir. Bu ağlar, 200 turda (epok) eğitilmiş olup elde edilen gerçek ve tahmini hız grafikleri Şekil 8’de verilmiştir.

Geliştirilen modellerin testleri, belirlenen test kümesi kullanılarak yapılmıştır. Yapılan tahmin sonucunun hata oranları RMSE, MAPE ve MAE hata metrikleriyle ölçülmüştür. Hata değerleri sırasıyla Eskişehir Yolunda hata oranları LSTM-GRU modelleri için 8,595-8,656, %14,339-%14,766, 6,387-6,421 ve İstanbul Yolunda LSTM-GRU için 7,331-7,955, %6,846-%7,656, 6,407-7,098 olarak elde edilmiştir. RMSE, MAPE ve MAE değerleri Tablo 6’da, grafik olarak Şekil 9’da gösterilmiştir. Sinir ağları modelleri için son üç küme hız değeri zaman penceresi olarak seçildiğinde tekrarlayan örüntülerle orijinal hız verilerine yakın değerler üretilmiştir ve birbirine yakın hata değerleri ölçülmüştür. Eskişehir yol güzergâhında bulunan 12, 13 ve 14. kümelerde ani düşüş ve çıkış gözlenmektedir. Bu noktada her iki sinir ağ modelinin hata oranı artmış fakat diğer kümelerde çok yakın tahmin üretmiştir. Aynı şekilde,

İstanbul yol güzergâhında bulunan 8, 9 ve 10. kümelerde mevcuttur. Bu durum her iki yol için kavşaklardaki yavaşlamalardan kaynaklanmaktadır. Kavşaklardaki yavaşlamalar ve hız değişimleri hata oranı yükseltse de genel toplam hata oranı literatürdeki çalışmalar değerlendirildiğinde tatmin edicidir.



Şekil 9. Hata değerlerinin grafiği (Graph of error values)



Şekil 8. Farklı modeller ve güzergâhlar için gerçek ve tahmin hız grafikleri (Actual and predicted speed graphics for different models and routes)

Tablo 6. Trafik hız tahmini için hata değerleri (Error values of traffic speed prediction)

	Eskişehir Yolu Güzergâhı		İstanbul Yolu Güzergâhı	
	LSTM Modeli	GRU Modeli	LSTM Modeli	GRU Modeli
RMSE	8,595	8,656	7,331	7,955
MAPE (%)	14,339	14,766	6,846	7,656
MAE	6,387	6,421	6,407	7,098

6. SONUÇLAR (CONCLUSIONS)

Bu çalışmada, anlık trafik yoğunluğunu gösteren ve trafik parametrelerinden kısa zamanlı hız tahmini yapan yeni bir model geliştirilmiş, test edilmiş ve başarı ile sunulmuştur. 9 aşamadan oluşan bu modelde, başlangıç aşamaları, anlık trafik hız gösterimi ve hız tahmini yapmak için mevcut trafik verisini hazırlama aşamalarını içerirken, son aşamalar hız tahmini, test ve karşılaştırma aşamalarını kapsamaktadır. 14 küme ile yapılan hız tahmin testlerinde 2 farklı yol ve 2 farklı RNN model varyantı ile ortalama 8 km/h hız tahmin farkı ve ortalama yaklaşık %10 hata oranı ile tatmin edici sonuçlar elde edilmiştir. Trafik hız ve diğer parametrelerin tahmini üzerine mevcut çalışmalarda gerçekleştirilen segment uzunluğu kadar tahmin üretmek yerine, geliştirilen tahmin modelinde segmenti kendi olağan trafik akışında tahmin etmek için yoğunluk tabanlı kümeler elde edilmiştir. Oluşan her küme için öznitelikler üretilmiş ve küme öznitelikleri yol içinde küçük segmentler oluşturularak yol karakteristiği elde edilmiştir. Yol karakteristikleri elde edilen yollar için aynı zamanda harita üzerinde anlık trafik yoğunluk gösterimi ve renklendirilmesi yapılmıştır. Geliştirilen modellere küme hız değerleri girdi olarak verilmiş ve modeller test edilmiş ve tahmini hız değerleri bulunmuştur. Trafığın olasılıklı ve değişken yapısı için önerilen bu modelin, gerçek zaman uygulamaları için başarılı sonuçlar verdiği belirlenmiştir. Bu çalışmada önerilen model ve çözümlerin, gelecekte yapılacak olan gerçek zamanlı trafik parametrelerinin tahmininde farklı ve yeni çözümler sunacağı, farklı alanlarda uygulanabileceği, süreçleri hızlandıracağı ve en önemlisi ise kullanıcılara daha doğru ve hızlı hizmet verilmesine katkılar sağlayacağı değerlendirilmektedir.

TEŞEKKÜR (ACKNOWLEDGEMENT)

Bu çalışma, 3191873 numaralı TÜBİTAK proje kapsamında yapılmıştır. Yazarlar; başta TÜBİTAK ve Başarsoft Bilgi Teknolojileri AŞ'ye, verdiği teknolojik altyapı desteği için Gazi BİDİSEC'e teşekkür ederler.

KAYNAKLAR (REFERENCES)

1. A. Chesterton. How many cars are there in the world?. <https://web.archive.org/web/20210322032308/https://www.carsguide.com.au/car-advice/how-many-cars-are-there-in-the-world-70629>. Erişim tarihi Mart 22, 2020.
2. Djahel S., Doolan R., Muntean G.-M., Murphy J., A Communications-Oriented Perspective on Traffic Management Systems for Smart Cities: Challenges and Innovative Approaches, *IEEE Commun. Surv. Tutorials*, 17 (1), 125–151, 2015.
3. Boukerche A., Wang J., Machine Learning-based traffic prediction models for Intelligent Transportation Systems, *Comput. Networks*, 181 (2020), 107530, 2020.
4. Lingani G. M., Rawat D. B., Garuba M., Smart traffic management system using deep learning for smart city applications, 2019 IEEE 9th Annual Computing and Communication Workshop and Conference CCWC 2019, Las Vegas-USA, 101–106, 7-9 Ocak, 2019.
5. Lanke N., Koul S., Smart Traffic Management System, *Int. J. Comput. Appl.*, 75 (7), 19–22, 2013.
6. Dimitrakopoulos G., Demestichas P., Intelligent Transportation Systems, *IEEE Veh. Technol. Mag.*, 5 (1), 77–84, 2010.
7. Mahdavejad M. S., Rezvan M., Barekatin M., Adibi P., Barnaghi P., Sheth A. P., Machine learning for internet of things data analysis: a survey, *Digital Communications and Networks*, 4 (3), 161–175, 2018.
8. Rao K. V. K. Fundamental parameters of traffic flow. <https://web.archive.org/web/20210406205237/https://nptel.ac.in/content/storage2/courses/105101087/downloads/Lec-30.pdf>. Erişim tarihi Nisan 6, 2020.
9. El Sherief M. M., Ramadan I. M. I., Ibrahim A. M., Development of traffic stream characteristics models for intercity roads in Egypt, *Alexandria Eng. J.*, 55 (3), 2765–2770, 2016.
10. Kovačić K., Ivanjko E., Jelušić N., Measurement of Road Traffic Parameters based on Multi-Vehicle Tracking, *arXiv*, arXiv:1510.04860 (2015), 3–8, 2015
11. Barros J., Araujo M., Rossetti R. J. F., Short-term real-time traffic prediction methods: A survey, 2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Budapest-Macaristan, 132–139, 3-5 Haziran, 2015.
12. Akin M., Sağıroğlu S., Büyük Veri Perspektifinden Trafik Tahmini Traffic Prediction Based on Big Data Perspective, 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara-Türkiye, 1-6, 6-7, Kasım, 2019.
13. George S., Santra A. K., Traffic Prediction Using Multifaceted Techniques: A Survey, *Wirel. Pers. Commun.*, 115 (2), 1047–1106, 2020.
14. Silgu M. A., Çelikoğlu H. B., K-Means Clustering Method to Classify Freeway Traffic Flow Patterns, *Pamukkale Univ. J. Eng. Sci.*, 20 (6), 232–239, 2014.
15. Do L. N. N., Vu H. L., Vo B. Q., Liu Z., Phung D., An effective spatial-temporal attention based neural network for traffic flow prediction, *Transp. Res. Part C Emerg. Technol.*, 108, 12-28, 2019.
16. Li Z., Filev D. P., Kolmanovsky I., Atkins E., Lu J., A New Clustering Algorithm for Processing GPS-Based Road Anomaly Reports With a Mahalanobis Distance, *IEEE Trans. Intell. Transp. Syst.*, 18 (7), 1980–1988, 2017.
17. Diker A. C., Nasibov E., Estimation of traffic congestion level via FN-DBSCAN algorithm by using GPS data, 2012 IV International Conference "Problems of Cybernetics and Informatics" (PCI), Bakü-Azerbaycan, 1-4, 12-14 Eylül, 2012.
18. Cui Z., Ke R., Pu Z., Ma X., Wang Y., Learning traffic as a graph: A gated graph wavelet recurrent neural network for network-scale traffic prediction, *Transp. Res. Part C Emerg. Technol.*, 115, 102620, 2020.
19. Akin M., Sağıroğlu S., Degirmenci A., Traffic Flow Forecasting Model with Density Based Clustering Algorithm, 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara-Türkiye, 1-5, 6-7, Kasım, 2019.

20. Vlahogianni E. I., Karlaftis M. G., Golias J. C., Short-term traffic forecasting: Where we are and where we're going, *Transp. Res. Part C Emerg. Technol.*, 43, 3–19, 2014.
21. Seo T., Bayen A. M., Kusakabe T., Asakura Y., Traffic state estimation on highway: A comprehensive survey, *Annu. Rev. Control*, 43, 128–151, 2017.
22. Lana I., Del Ser J., Velez M., Vlahogianni E. I., Road Traffic Forecasting: Recent Advances and New Challenges, *IEEE Intell. Transp. Syst. Mag.*, 10 (2), 93–109, 2018.
23. Jiang B., Fei Y., Vehicle Speed Prediction by Two-Level Data Driven Models in Vehicular Networks, *IEEE Trans. Intell. Transp. Syst.*, 18 (7), 1793–1801, 2017.
24. Hou Y., Chen J., Wen S., The effect of the dataset on evaluating urban traffic prediction, *Alexandria Eng. J.*, 60(1), 597-613, 2020.
25. Asif M. T. et al., Spatiotemporal patterns in large-scale traffic speed prediction, *IEEE Trans. Intell. Transp. Syst.*, 15 (2), 794–804, 2014.
26. Jeon S., Hong B., Monte Carlo simulation-based traffic speed forecasting using historical big data, *Futur. Gener. Comput. Syst.*, 65, 182–195, 2016.
27. Ma X., Tao Z., Wang Y., Yu H., Wang Y., Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, *Transp. Res. Part C Emerg. Technol.*, 54, 187–197, 2015.
28. Cheng Z., Chow M.-Y., Jung D., Jeon J., A big data based deep learning approach for vehicle speed prediction, 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh-UK, 389–394, Haziran 19-21, 2017.
29. Niu K., Zhang H., Zhou T., Cheng C., Wang C., A Novel Spatio-Temporal Model for City-Scale Traffic Speed Prediction, *IEEE Access*, 7, 30050–30057, 2019.
30. Zhao J. et al., Truck Traffic Speed Prediction Under Non-Recurrent Congestion: Based on Optimized Deep Learning Algorithms and GPS Data, *IEEE Access*, 7, 9116–9127, 2019.
31. Da Silva T. L. C et al., Efficient and distributed DBScan algorithm using mapreduce to detect density areas on traffic data, *ICEIS 2014 - Proc. 16th Int. Conf. Enterp. Inf. Syst.*, 1, 52–59, 2014.
32. Ankerst M., Breunig M. M., Kriegel H., Sander J., OPTICS: Ordering Points To Identify the Clustering Structure, *ACM SIGMOD Int. Conf. Manag. data*, Philadelphia-USA, 49–60, Mayıs 31-Haziran 3, 1999.
33. Martin Ester X. X., Kriegel H., Sander J., A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise, *KDD-96 Proceedings*, Portland-USA, 226–231, Ağustos 2-4, 1996.
34. Cui Z., Ke R., Pu Z., Wang Y., Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction, *arXiv*, 2018.
35. Wang Y., Zhang D., Liu Y., Dai B., Lee L. H., Enhancing transportation systems via deep learning: A survey, *Transp. Res. Part C Emerg. Technol.*, 99, 144–163, 2019.
36. Işık G., Artuner H., Turkish dialect recognition in terms of prosodic by long short-term memory neural networks, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35 (1), 213–224, 2020.
37. Fusco G., Colombaroni C., N. Isaenko, Short-term speed predictions exploiting big data on large urban road networks, *Transp. Res. Part C Emerg. Technol.*, 73, 183–201, 2016.
38. Winarno E., Hadikurniawati W., Rosso R. N., Location based service for presence system using haversine method, 2017 International Conference on Innovative and Creative Information Technology (ICITech), Salatiga-Indonesia, 1–4, Kasım 2-4, 2017.
39. Chopde N. Nichat M., Landmark Based Shortest Path Detection by Using A* and Haversine Formula, *Int. J. Innov. Res. Comput. Commun. Eng.*, 1 (2), 298–302, 2013.
40. Sondhi P., Feature construction methods: a survey, *Sifaka. Cs. Uiuc. Edu*, 69, 70–71, 2010.

