



Akıllı telefon algılayıcıları ve makine öğrenmesi kullanılarak ulaşım türü tespiti

Transportation mode detection by using smartphone sensors and machine learning

Ensar Arif SAĞBAŞI^{1*}, Serkan BALLI¹

¹Bilişim Sistemleri Mühendisliği Bölümü, Teknoloji Fakültesi, Muğla Sıtkı Koçman Üniversitesi, Muğla, Türkiye.
arifsagbas@mu.edu.tr, serkan@mu.edu.tr

Geliş Tarihi/Received: 21.07.2015, Kabul Tarihi/Accepted: 26.08.2015

doi: 10.5505/pajes.2015.63308

* Yazışılan yazar/Corresponding author

Araştırma Makalesi/Research Article

Öz

Bu çalışmada akıllı telefon algılayıcıları kullanılarak kullanıcıların ulaşım türü tespitinin yapılması amaçlanmaktadır. Bunun için kullanıcıdan yürürken, koşarken, bisiklet sürerken, araba veya otobüs ile seyahat ederken GPS (Global Positioning System), ivmeölçer ve jiroskop algılayıcılarından elde edilen veriler toplanmıştır. Veriler 12'şer saniyelik aralıklarla etiketlenmiş ve toplamda 2500 örüntü elde edilmiştir. Bu verilerden 14 öznitelik elde edilmiştir. Oluşturulan veri seti ile makine öğrenmesi yöntemleri kullanılarak testler gerçekleştirilmiştir. En iyi sonuç GPS, ivmeölçer ve jiroskop algılayıcılarının kombinasyonundan, %99.4 doğruluk oranı ile Random Forest yönteminden elde edilmiştir.

Anahtar kelimeler: Ulaşım türü, Sınıflandırma, Akıllı telefon, Algılayıcı verisi

Abstract

The aim of this study is to detect transportation modes of the users by using smartphone sensors. Therefore, GPS (Global Positioning System), accelerometer and gyroscope sensor data have been collected while walking, running, cycling and travelling by bus or by car from the smartphone of the user. Sensor data were tagged with 12 second interval and 2500 pattern were obtained. 14 features were acquired from the dataset. Machine learning methods were tested on the dataset. Best result was obtained from GPS, accelerometer and gyroscope sensor combination and Random Forest method with 99.4% accuracy rate.

Keywords: Transportation mode, Classification, Smartphone, Sensor data

1 Giriş

Teknolojideki hızlı gelişme ile akıllı telefonlar insanların sürekli yanlarında taşıdıkları, kullanıcılarına telefon görüşmesinden daha fazlasını sunan cihazlar haline gelmiştir. Günümüz akıllı telefonları haberleşmenin yanında ivmeölçer, jiroskop, dijital pusula, GPS (Global Positioning System), mikrofon ve kamera gibi zengin algılayıcılar sağlamaktadır. Bu algılayıcı verilerinin toplanması; sağlık, sosyal medya, taşıma, çevresel izleme ve güvenlik gibi çeşitli uygulamaların geliştirilebilmesini mümkün kılmaktadır [1]. Bu algılayıcılar aynı zamanda konum tespitine ve ivme hesaplanmasına da olanak sağlamaktadır. Böylece akıllı telefonların sahip olduğu algılayıcı verilerini kullanarak kişinin yapmış olduğu yürüme, koşma, bisiklet sürme, motorlu taşıt ile seyahat etme gibi gündelik aktivitelerin tespiti de mümkün olmaktadır [2]. Bu çalışmanın konusu akıllı telefon algılayıcı verileri kullanılarak ulaşım türü tespitinin gerçekleştirilmesidir.

Literatür incelendiğinde ulaşım türü tespiti çeşitli makine öğrenmesi sınıflandırma algoritmaları kullanılarak gerçekleştirildiği görülmektedir. Reddy ve diğ. [2] tarafından yapılan çalışmada akıllı telefonların GPS ve ivmeölçer algılayıcılarından elde edilen veriler ile ulaşım türü tespiti gerçekleştirilmiştir. Sınıflandırma algoritması olarak Decision Tree (Karar Ağacı) kullanılmış ve mobil cihazı taşımada en doğru sonuçlar cepte taşımada elde edilmiştir. Zheng ve diğ. [3] Decision Tree algoritması kullanarak ham GPS verilerinden ulaşım türü öğrenilmesi konusunu çalışmıştır. Gyorbíró ve diğ. [4], akıllı telefonlar için eylem tanıma sistemi önermiştir. Sınıflandırmada ileri beslemeli sinir ağları kullanılmıştır. Wang ve diğ. [5], akıllı telefonlarda bulunan algılayıcılardan sadece dâhili ivmeölçeri kullanarak ulaşım türü tespiti

gerçekleştirmiştir. Stenneth ve diğ. [6], telefon ve coğrafi bilgi sistemleri bilgileri ile ulaşım türü tespiti üzerinde çalışmıştır. Lara ve diğ. [7], kullanıcı göğüs kafesi üzerinde bulunan algılayıcı verileri ile insan eylemlerini tanıyabilen ve Additive Logistic Regression (Eklemeli Lojistik Regresyon) algoritmasını kullanan bir sistem geliştirmiştir. Widhalm ve diğ. [8] Hidden Markov Model kullanarak, akıllı telefonu taşıyan kişinin ulaşım türünü tespit etmede yeni bir model önermişlerdir. Köse ve diğ. [9] akıllı telefonların dâhili ivmeölçer algılayıcısını kullanarak çevrimiçi eylem tanıma konusunda çalışmışlardır. Ayrıca çeşitli sınıflandırma metodlarının performanslarını çevrimiçi eylem tanıma analiz etmişlerdir. Bolbol ve diğ. [10] GPS verileri ile ulaşım türü tespitinde meydana gelen sınıflandırma problemini çözümünde Support Vector Machines (Destek Vektör Makineleri) tabanlı bir yapı sunmuştur. Hemminki ve diğ. [11] tarafından yapılan çalışmada akıllı telefondan sağlanan veriler ile ivmeölçer tabanlı ulaşım türü tespiti yapılmış ve önceki çalışmalarla karşılaştırılmıştır. Feng ve Timmermans [12] GPS ve ivmeölçer verileri kullanarak ulaşım türü tespiti gerçekleştirmiştir. Çalışmada sınıflandırma yöntemi olarak Bayesian Belief Network (Bayes Ağları) kullanılmıştır. Ellis ve diğ. [13] GPS, ivmeölçer algılayıcıları ve makine öğrenmesi algoritmaları kullanarak ulaşım türündeki değişimleri tespit etmiştir. Shin ve diğ. [14] ulaşım türü tespiti gerçekleştirmek için akıllı telefonun dâhili ivmeölçer ve ağ konumu algılayıcısını kullanan bir mobil uygulama geliştirmiştir. Sökün ve diğ. [15], tarafından yapılan çalışmada üç eksenli ivmeölçerden alınan sinyaller kullanılarak temel fiziksel hareket sınıflandırılması yapılmıştır. Geliştirilen yöntem ile ivmeölçeri taşıyan insanın araç ile mi, yoksa yaya olarak mı seyahat ettiğinin tespit edilmesi amaçlanmıştır. Alınan veriler belirli zaman aralıklarında incelenmiş ve gerçek zamanlı

olarak kNN (k En Yakın Komşu) algoritması kullanılarak sınıflandırılmıştır. Xia ve diğ. [16] tarafından yapılan çalışmada akıllı telefon algılayıcıları ve Support Vector Machines algoritması kullanılarak ulaşım türü sınıflandırması yapılmıştır.

Bu çalışmanın konusu, akıllı telefonların dâhili GPS, ivmeölçer ve jiroskop algılayıcılarını kullanarak ulaşım türü tahmini gerçekleştirmektir. Amaç, akıllı telefonu cebinde taşıyan kullanıcının yürüme, koşma, bisiklet sürme ve araba veya otobüs ile seyahat etme eylemlerinden hangisini gerçekleştirdiğini tespit edebilmektir. Bu işlem için 8 saatin üzerinde akıllı telefon algılayıcısı verisi toplanmıştır. Toplanan veriler algılayıcılara göre ayrı ayrı Naive Bayes, kNN, C4.5, Bayesian Network (Bayes Ağları), Random Forest (Rastgele Orman) ve Logistic Regression yöntemleri ile incelenip, en iyi sonucu veren algılayıcı veya algılayıcı grubu sınıflandırma işlemi için seçilmiştir.

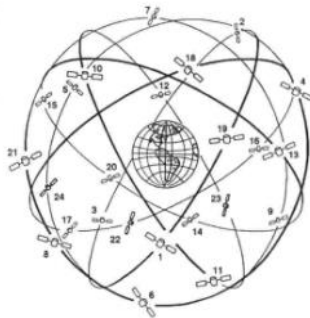
Çalışmanın ikinci bölümde, kullanılan akıllı telefon algılayıcıları hakkında, üçüncü bölümde ise sınıflandırmada kullanılan yöntemler hakkında bilgiler verilecektir. Dördüncü bölümde gerçekleştirilen uygulama, algılayıcılar ve yöntemlere göre ulaşılan deney sonuçları incelenecektir. Beşinci bölümde çıkarılan sonuçlar değerlendirilecektir. Altıncı bölümde ise gelecek çalışmalar için yapılan planlardan bahsedilecektir.

2 Akıllı telefon algılayıcıları

Çalışmada akıllı telefonu cebinde taşıyan kullanıcının gerçekleştirdiği eylemi ya da hangi araç ile seyahat ettiğini tespit edebilmek için akıllı telefonun dâhili GPS, ivmeölçer ve jiroskop algılayıcıları kullanılmıştır. Aşağıdaki alt bölümlerde bu algılayıcılar hakkında bilgiler verilecektir.

2.1 GPS

GPS, 1970'lerin başlarında U.S. Department of Defens (DoD) tarafından geliştirilen uydu tabanlı yönlendirme sistemidir. GPS, askeri ihtiyaçları karşılamak amacı ile geliştirilmiştir. Sonradan kişisel kullanım mümkün olmuştur. Şu anda bu sistem hem askeri hem sivil kullanıma uygundur. GPS, kullanıcıya dünya üzerinde herhangi bir yerde herhangi bir hava şartı altında konum ve zaman bilgisi sağlamaktadır. GPS, limitsiz kullanıcıya servis sağlayabilen pasif bir sistemdir. Yani, kullanıcılar sadece uydu sinyallerini alırlar. GPS dünya etrafında yörüngede bulunan Şekil 1'de gösterilen 24 operasyonel uyduya sahiptir [17].

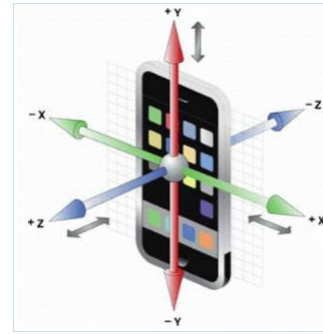


Şekil 1: Dünya yörüngesindeki 24 GPS uydusu.

2.2 İvmeölçer

İvmeölçer algılayıcısı Şekil 2'de gösterilen eksenler doğrultusunda akıllı telefona etki eden ivmeyi ölçer. Ham algılayıcı bilgisi ivmeölçerden üç eksende g cinsinden elde edilir. Ham ivmeölçer algılayıcısı verisi $Acc_i = \langle x_i, y_i, z_i \rangle, i = (1,2,3, \dots)$ vektörlerini içerir. İvmelenme değerlerinin yanında ayrıca zaman bilgisi de elde edilir. Mevcut çoğu ivmeölçer kullanıcı arayüzünde örnekleme hızını ayarlamaya imkân sunmaktadır. Böylece kullanıcı en uygun örnekleme hızını seçebilmektedir.

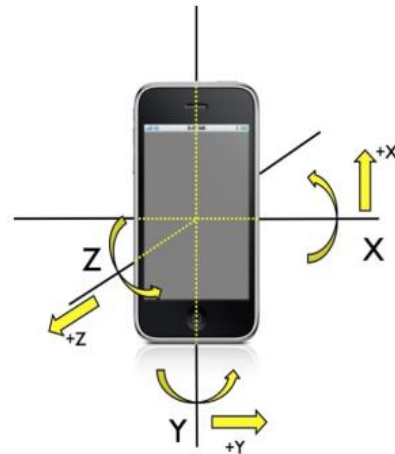
İvmeölçer, akıllı telefon tabanlı eylem tanıma uygulamalarında sıkça kullanılmaktadır. Bu algılayıcının popülerliği konumlandırıldığı cihazın veya taşıyan kullanıcının fiziksel hareketini direkt olarak hesaplayabilmesinden gelmektedir. Örneğin, kullanıcı yürür durumdan zıplar duruma geçerse ivmeölçer sinyallerinin şekli dikey ekseninde değişecektir [18].



Şekil 2: İvmeölçer algılayıcısı eksenleri.

2.3 Jiroskop

Jiroskop algılayıcısı akıllı telefonun x, y ve z ekseninde yapmış olduğu açısal hızı vermektedir. Eksen yörüngeleri Şekil 3'te gösterilmektedir. Jiroskop algılayıcısından elde edilen ham veriler akıllı telefonun üç fiziksel eksen etrafında dönüşünü rad/sn (radyan / saniye) cinsinden bildirmektedir: $Rotation_i = \langle x_i, y_i, z_i \rangle, i = (1,2,3, \dots)$. Karakter yönlendirme yapılan akıllı telefon oyunlarında jiroskop algılayıcısından yararlanılmaktadır. Eylem tanıma araştırmalarında bu algılayıcı, yön tespitini gerçekleştirmede yardımcı olarak kullanılmaktadır [18,19].



Şekil 3: Jiroskop algılayıcısı eksenleri.

3 Hareket sınıflandırması için alternatif yöntemler

Ulaşım türü tespiti ve eylem tanıma uygulamalarında makine öğrenmesi sınıflandırma algoritmalarından yararlanılmaktadır. Makine öğrenmesinde sık kullanılan algoritmalar aşağıdaki alt bölümlerde açıklanacaktır.

3.1 Naive bayes (NB)

NB Sınıflandırıcısı güçlü nitelik bağımsızlığı varsayımına sahip, olasılığa dayalı basit bir algoritmadır. NB Sınıflandırıcısı test verisi üzerinden öğrenme işlemini gerçekleştirir ve en yüksek orandaki örneğini sınıfa dâhil eder. C bir sınıfı gösterdiği düşünölsün. $x = \langle x_1, x_2, x_3, \dots, x_m \rangle$ değerleri ise gözlemlenen niteliklerin değerleri olsun. c bilinen bir sınıf etiketini, $x = \langle x_1, x_2, x_3, \dots, x_m \rangle$ ise bilinen ve gözlemlenen niteliklerin değerlerini temsil etsin. x test verisine göre sınıfı tahmin etmek için Bayes Teoremi olasılık hesaplar:

$$p(C = c_j | X = x) = \frac{p(C = c_j)p(X = x | C = c_j)}{p(X = x)} \quad (1)$$

Sonrasında en yüksek olasılık ile sınıfı tahmin eder. Bu örnekte $X = x$ durumu $X_1 = x_1 \wedge X_2 = x_2 \wedge X_3 = x_3 \wedge \dots \wedge X_m = x_m$ ifade eder. $p(X = x)$ sınıflar arasında değişme göstermediği durumda ihmal edilir ve (1) numaralı denklem aşağıdaki şekle gelir.

$$p(C = c_j | X = x) = p(C = c_j)p(X = x | C = c_j) \quad (2)$$

($C = c_j$) ve $p(X = x | C = c_j)$ öğrenme verilerinden tahmin edilir. $x_1, x_2, x_3, \dots, x_m$ nitelikleri birbirlerine koşullu olarak bağımsızdırlar. Bu durumda (2) numaralı denklem aşağıdaki şekli alır:

$$p(C = c_j | X = x) = p(C = c_j) \prod_{i=1}^m p(X_i = x_i | C = c_j) \quad (3)$$

NB denilen (3) numaralı denklemi kullanarak test örneklerini hesaplamak ve öğrenme verilerinden tahminde bulunmak daha kolaydır [20].

Yapılan ulaşım türü tespiti çalışmasında 14 koşullu bağımsız nitelik üzerinden sınıflandırma işlemi gerçekleştirilmiştir.

3.2 Bayes ağları (BA)

BA düğümlerinin değişkenleri, okların bu düğümler arasında bağlantıyı temsil ettiği belirsizlik altında karar vermek için kullanılan grafiksel modellerdir. BA kesin olmayan çevrede durumu ifade etmeyi ve sonuç çıkarmayı mümkün kılan grafiksel yapılardır. Bir bayes ağında düğümler çevreden elde edilen rasgele değişkenleri temsil eder ve yönlü oklar ile birbirine bağlanır. Bu oklar düğümler arasındaki bağımlılığı temsil eder. Ayrık iki düğüm arasında bağlantının gücü, o iki düğüm arasındaki koşullu olasılık ile ölçülür. İlk olarak değişkenler arasındaki ilgi belirlenir. Bunun sonucunda düğümlerin neleri temsil ettiği ve hangi değerleri aldığı saptanır. Ayrık değerlere sahip düğümler ele alındığında, değerler hem birbirini dışlayan hem de kapsamlı olmalıdır. Bu da, bu değerlerin tam değerler almasıdır. Ağın yapısı veya topolojisi değişkenler arasındaki nitel ilişkiyi ele alır. Özellikle, iki düğüm birbirini etkiliyor veya birinin gerçekleşmesine neden oluyorsa bu iki düğüm direk bağlantılı olarak gösterilir. Ağ yapısından bahsederken aile metaforu kullanmak yararlıdır. İki düğüm arasında yönlü oklar ile bağlantı varsa

ebeveyn ve çocuk düğüm terimleri kullanılır. Ağ yapısında bir düğüm başka bir düğümün soyundan geliyorsa, ağ yapısında önce gelen düğüm sonra gelen düğümün atasıdır. Başka yararlı bir kavram ise Markov Blanket'tir. Bu kavram düğümün ebeveynlerinden, çocuklarından ve çocukların ebeveynlerinden oluşur. Başka bir yaygın kullanılan terminoloji ise ağaçtır. Ebeveyni olmayan düğüm kök olarak adlandırılır. Çocuk düğümüne sahip olmayan düğümler ise yaprak olarak adlandırılır. Bunun dışında kalan düğümlere ise (yaprak ve kök olmayan) ara düğüm adı verilir [21].

Gerçekleştirilen çalışmada kullanılan 14 parametre ile sınıfları (Araba, Otobüs, Bisiklet, Koşma, Yürüme) temsil eden düğümlere bağlanarak bir Bayes Ağı yapısı oluşturulmuştur.

3.3 K en yakın komşu (kNN)

kNN en teorik ve basit sınıflandırma yöntemlerinden biridir. KNN sınıflandırıcı $f(X) = k_i/k$ şeklinde tanımlanmış olup burada $k_i(c_i)$, X 'in k komşuluğunda c_i sınıfında yer alan örneklerin sayısını göstermektedir. kNN algoritmasında sınıflandırılacak olan yeni örneğe eğitim setinden en yakın mesafedeki k tane örneğe bakılır ve bu k örnek çoğunluk olarak hangi sınıfa dâhil edilmiş ise yeni örnek de o sınıfa dâhil edilir [15]. Mesafe hesabı olarak ise yaygın bilinen eşitlik 4'te verilen Manhattan uzaklık ölçütü, eşitlik 5'te verilen Öklid uzaklık ölçütü ve eşitlik 6'da verilen Minkowski uzaklık ölçütü formülleri aşağıdaki gibidir:

$$d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}| \quad (4)$$

$$= \sqrt{|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2} \quad (5)$$

$$= \sqrt[p]{|x_{i1} - x_{j1}|^p + |x_{i2} - x_{j2}|^p + \dots + |x_{ip} - x_{jp}|^p} \quad (6)$$

Yapılan çalışmada k değeri 3 olarak alınmıştır. Mesafe hesabı için ise Öklid uzaklık ölçütü kullanılmıştır.

3.4 C4.5

C4.5 algoritması bilgi entropi kavramı kullanarak bir eğitim setinden karar ağacı inşa eder. Eğitim seti $S = S_1, S_2, \dots, S_n$ verilerinin sınıflandırılmış örnekleridir. Her bir S_i örneği p boyutlu vektörden $X_{1,i}, X_{2,i}, \dots, X_{p,i}$ oluşmaktadır. Burada X_j örneklerin niteliklerini veya özelliklerini göstermektedir. Ağacın her bir düğümünde C4.5, alt kümelerdeki zenginleştirilmiş verinin niteliğini seçer. Bölünme ölçütü normalize bilgi kazancıdır. En yüksek normalize bilgi kazancına sahip nitelik karar için seçilir. Sonrasında C4.5 algoritması daha küçük alt listelerden çekilir [22].

C4.5 algoritması bu çalışmada 16 yaprağa sahip ağaç boyutu 31 olan bir yapı oluşturmuştur.

3.5 Random forest (RF)

RF (Rastgele Orman) yöntemi çok sayıda karar ağacından oluşan orman sınıflayıcısıdır ve bu yöntemle, sınıflama veya regresyon ağaçları kurulabilmekte ve kümeleme yapılabilmektedir. Veri setindeki "sınıf değişkeni" kategorik ise sınıflama, sürekli ise regresyon ağaçları kurulmaktadır. Ormanda yer alan her bir karar ağacı, bootstrap tekniği ile orijinal veri setinden örneklem seçilmesi ve her karar düğümünde tüm değişkenler içinden belirlenen sayıda

rastgele değişkenin seçilmesi ile oluşturulmaktadır. Bu yöntemde CART algoritması ile ağaçlar oluşturulur ve ağaçlar budanmaz. CART algoritması veri setinin hangi değişkenden başlayarak dallara ayrılacağına "bilgi kazancı" kullanarak karar verir. Ayrıca dallara ayrılmak için seçilen değişkenin uygun test ölçütü "Gini Katsayısı" ile belirlenir [23]. Çalışmada 4 rastgele öznelik üzerinden 100 rastgele ağaç inşa edilmiştir.

3.6 Lojistik regresyon (LR)

LR bir veya çok ön göstergeli değişkene bağlı kategorik bağımlı değişkenlerin sonuçlarını tahmin etmede kullanılan bir tür regresyon analizidir. Geçmişte, temel LR modelinin genişletilmiş hali olarak farklı türde modeller geliştirilmiştir. Multinomial (Çok terimli) LR modeli ikiden fazla ayrı sonuca izin veren lojistik regresyonu genelleştiren bir modeldir. Yani bağımsız değişkenleri verilen, kategorik dağıtılan bağımlı değişkenlerin farklı olası sonuçlarını tahmin etmek için kullanılan modeldir. m nitelikli n örnekleri için k sınıflarının olduğu varsayılınsın. Bu durumda son sınıf hariç j sınıfının olasılığı:

$$P_j(X_i) = \frac{\exp(x_i b_j)}{\sum_{j=1}^{k-1} \exp(x_i b_j) + 1} \quad (7)$$

Son sınıf olasılığı

$$1 - \sum_{j=1}^{k-1} P_j(X_i) = \frac{1}{\sum_{j=1}^{k-1} \exp(x_i b_j) + 1} \quad (8)$$

Böylece multinomial log-likelihood:

$$L = - \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} (Y_{ij} \ln(P_j(x_i))) + \left(1 - \sum_{j=1}^{k-1} Y_{ij}\right) \cdot \ln(1 - \sum_{j=1}^{k-1} P_j(x_i)) + ridge \cdot B^2 \quad (9)$$

$ridge$, log-likelihood fonksiyonunda önceden verilmesi gereken bir parametredir. Minimize edildiği L B , bir Yarı-Newton Yöntemi $m * (k - 1)$ değişkenlerinin uygun değerlerini aramak için kullanılır [22]. Çalışmada $ridge$ değeri $1.0E-8$ olarak ayarlanmıştır.

4 Uygulama

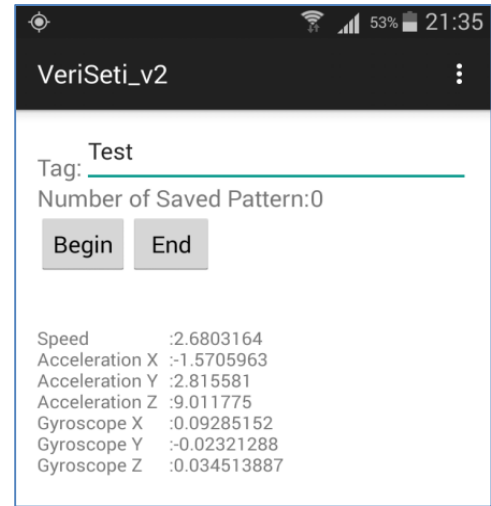
Çalışmada akıllı telefon kullanıcısının hangi ulaşım türünü kullandığını tespit etmek hedeflenmiştir. Bu amaçla veri setinin oluşturulacağı ve sınıflandırma işleminin gerçekleştirileceği bir Android tabanlı mobil uygulama geliştirilmiştir. Sınıflandırma işlemi için hangi değişkenlerin kullanılacağı gerçekleştirilen testler ile karar verilmiştir.

4.1 Verilerin toplanması

Çalışmada kullanılacak veri setinin hazırlanabilmesi için Android işletim sistemine sahip akıllı telefonlar üzerinde çalışabilen, ekran görüntüsü Şekil 4'te gösterilmiş olan bir mobil uygulama geliştirilmiştir. Geliştirilen uygulama, gerçekleştirilen eylemlerin ham algılayıcı verilerini etiketlenerek depolanmasına imkân sunmaktadır. Çalışmanın veri setini oluşturma aşamasında Samsung Galaxy Note 2 (GT-N7100) [24] marka akıllı telefon kullanılmıştır. Bu cihaz Quad Core 1.6 GHz işlemci, 2 GB RAM ve dahili ivmeölçer, GPS,

jiroskop, barometre, ortam ışığı, yakınlık algılayıcılarına sahiptir. Saniyede en fazla 100 örnekleme algılayıcı bilgisi sağlayabilmektedir. Algılayıcı verilerini örnekleme hızı olarak SENSOR_DELAY_UI ayarlanmıştır. Bu ayar 50 Hz örnekleme hızına izin vermektedir.

Akıllı telefonu taşıyan kullanıcının hız bilgisi için GPS algılayıcısı, hareket bilgileri için ise ivmeölçer ve jiroskop algılayıcıları kullanılmıştır. Kullanıcı gerçekleştirmekte olduğu aktiviteyi yazıp "Begin" tuşuna basarak veri toplamaya başlayabilmektedir. Her bir örüntü için 12'şer saniye belirtilen algılayıcı verileri akıllı telefon kullanıcının cebinde konumlandırılmış bir şekilde toplanmıştır. 12 saniye içerisinde toplanan 600 satırlık veri etiketlenerek bir örüntüyü oluşturmuştur. Veriler akıllı telefonun dahili depolama biriminde, istenilen isimde, CSV formatında depolanmaktadır. Şekil 5'te araba sınıfına ait bir örüntünün ham verileri gösterilmektedir. Yürüme, koşma, bisiklet sürme, araba veya otobüs ile seyahat etme eylemlerinin her biri için 500 adet, toplamda 2500 adet örnek elde edilmiştir.



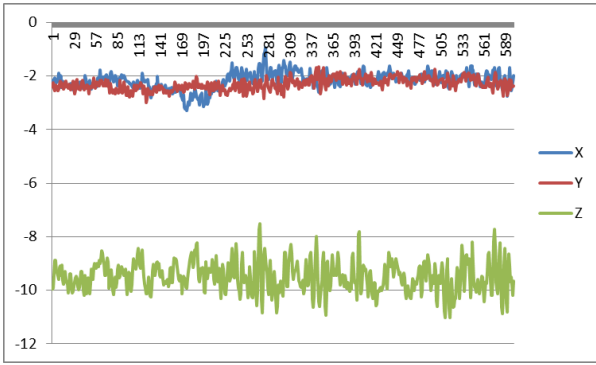
Şekil 4: VeriSeti Android uygulaması ekran görüntüsü.

	A	B	C	D	E	F	G
1	Speed	Acc X	Acc Y	Acc Z	Gyr X	Gyr Y	Gyr Z
2	35.47419	-1.96325	-5.65989	-8.3414	-0.01527	-6.11E-04	0.032376
3	35.47419	-1.87705	-5.40132	-8.29351	-0.00305	0.040012	-0.00794
4	35.47419	-1.77171	-5.62159	-7.79552	0.045509	0.110567	0.02474
5	35.47419	-1.77171	-5.61201	-8.61913	-0.01222	0.027184	0.045204
6	35.47419	-1.88663	-5.40132	-8.44674	0.022602	0.06353	0.033292
7	35.47419	-1.64721	-5.64074	-8.10198	0.036652	0.096822	0.022602
8	35.47419	-1.79086	-5.71735	-8.74363	-0.011	0.006109	0.014966
9	35.47419	-1.81959	-5.79397	-8.59997	0.02474	0.06017	0.056505
10	35.47419	-1.76213	-5.89931	-8.31267	0.018326	0.038485	0.065363
11	35.47419	-1.82917	-5.56412	-8.45632	-0.0281	-0.02199	0.014355
12	35.47419	-1.68552	-5.49709	-8.27436	-0.01283	-0.0058	0.025045
13	35.47419	-1.94409	-5.29597	-8.2169	-0.01313	0.031154	0.042455
14	35.47419	-1.75256	-5.15232	-7.54652	0.00733	0.04215	0.039401
15	35.47419	-1.79086	-5.42047	-7.54652	-3.05E-04	0.025045	0.0281
16	35.47419	-1.88663	-5.1619	-8.35098	-0.02352	0.009774	0.0281
17	35.97535	-1.90578	-5.36301	-8.43717	0.021991	0.081245	0.043677
18	35.97535	-1.99198	-5.81312	-8.14986	0.047342	0.095295	0.046731
19	35.97535	-2.20267	-5.91847	-8.53294	0.029932	0.039706	0.043066
20	35.97535	-2.02071	-5.70778	-8.72447	0.020769	0.025962	0.029932
21	35.97535	-2.07817	-5.64074	-8.53294	0.015882	0.049175	0.020464
22	35.97535	-2.07817	-5.37259	-8.42759	0.00672	0.053145	0.014355
23	35.97535	-2.03986	-5.09486	-7.69975	0.029322	0.058948	0.047037
24	35.97535	-2.08774	-5.17148	-7.44118	-0.00916	0.003054	0.042761
25	35.97535	-2.2314	-5.05655	-8.16902	-0.03879	-0.00641	0.025656

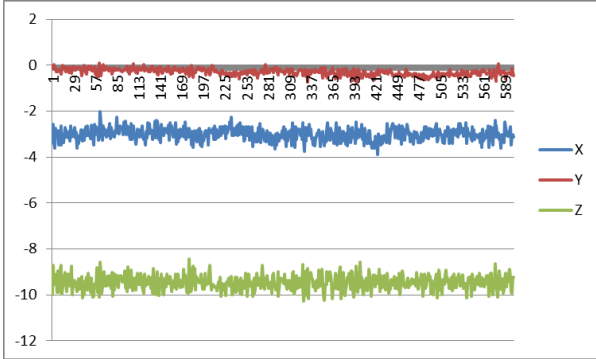
Şekil 5: Örnek veriler.

4.2 Kullanılacak parametrelerin seçilmesi

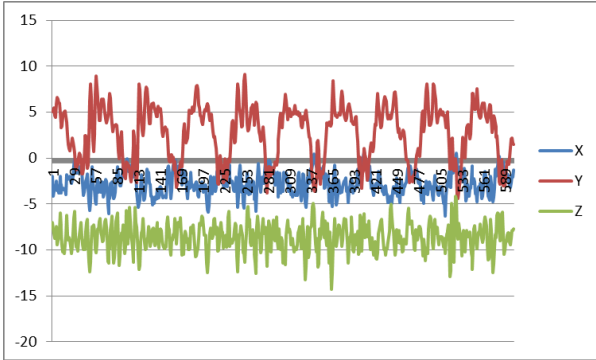
İvmeölçer algılayıcısından elde edilen verilerin ortalaması, cihazın hangi şekilde konumlandığı hakkında bilgi verirken, aynı verilerin standart sapması ise cihazın yapmış olduğu titreşim hakkında bilgi vermektedir. Jiroskop algılayıcısı, cihazın saniyede gerçekleştirdiği açılal hareketi ölçmektedir. Gerçekleştirilen çalışmada bu algılayıcıdan elde edilen verilerin standart sapma ve ortalama değerleri kullanılmıştır. GPS algılayıcısından elde edilen hız verisinin Feng ve Timmermans [12] tarafından yapılan çalışmadaki gibi ortalama değeri ve ulaştığı en yüksek değer hesaplamaya alınmıştır. Veriseti uygulamasından elde edilen örnekler adı geçen işlemler uygulanarak eyleme ait örüntü elde edilmiştir. Şekil 6-10'da sırasıyla araba, otobüs ile seyahat etme, bisiklet sürme, yürüme ve koşma eylemlerine ait örnek ivmeölçer verilerinin grafikleri gösterilmiştir.



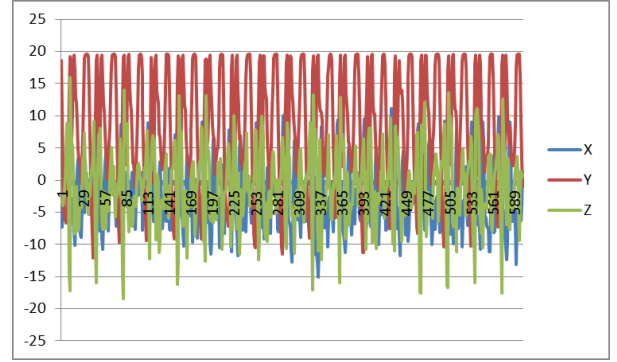
Şekil 6: Araba sınıfına ait ivmeölçer verileri.



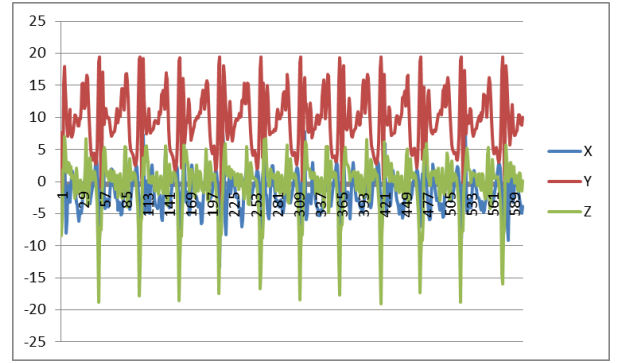
Şekil 7: Otobüs sınıfına ait ivmeölçer verileri.



Şekil 8: Bisiklet sınıfına ait ivmeölçer verileri.



Şekil 9: Koşma sınıfına ait ivmeölçer verileri.



Şekil 10: Yürüme sınıfına ait ivmeölçer verileri.

2500 verinin 1500 tanesi eğitim için, geri kalan 1000 tanesi ise test için ayrılmıştır. Veriler NB, kNN (k=3), C4.5, BA, LR ve RF yöntemleri ile Weka Toolkit [25] kullanılarak Tablo 1'de gösterilen algılayıcı kombinasyonları modelleri uygulanarak sınıflandırılmış ve performansları karşılaştırılmıştır.

Tablo 1: Değişkenlere göre model yapısı.

Model	İçerikler	Değişkenler
1	Sadece GPS	MaxHiz, OrtHiz
2	Sadece İvmeölçer	OrtXAcc, OrtYAcc, OrtZAcc, StdXAcc, StdYAcc, StdZAcc
3	Sadece Jiroskop	OrtXGyr, OrtYGyr, OrtZGyr, StdXGyr, StdYGyr, StdZGyr
4	GPS + İvmeölçer	MaxHiz, OrtHiz, OrtXAcc, OrtYAcc, OrtZAcc, StdXAcc, StdYAcc, StdZAcc
5	GPS + Jiroskop	MaxHiz, OrtHiz, OrtXGyr, OrtYGyr, OrtZGyr, StdXGyr, StdYGyr, StdZGyr
6	İvmeölçer + Jiroskop	OrtXAcc, OrtYAcc, OrtZAcc, StdXAcc, StdYAcc, StdZAcc, OrtXGyr, OrtYGyr, OrtZGyr, StdXGyr, StdYGyr, StdZGyr
7	GPS + İvmeölçer + Jiroskop	MaxHiz, OrtHiz, OrtXAcc, OrtYAcc, OrtZAcc, StdXAcc, StdYAcc, StdZAcc, OrtXGyr, OrtYGyr, OrtZGyr, StdXGyr, StdYGyr, StdZGyr

Tablo 1’de değişkenler sütununda gösterilen: MaxHiz, ulaşılan en yüksek hızı m/sn cinsinden, OrtHiz, ortalama hızı m/sn cinsinden göstermektedir. StdXAcc, ivmeölçerin X ekseninden ölçülen değerlerin standart sapma değerini, StdYAcc, ivmeölçerin Y ekseninden ölçülen değerlerin standart sapma değerini, StdZAcc, ivmeölçerin Z ekseninden ölçülen değerlerin standart sapma değerini vermektedir. OrtXAcc, ivmeölçerin X ekseninden elde edilen verilerin ortalama değeri, OrtYAcc, ivmeölçerin Y ekseninden elde edilen verilerin ortalama değeri, OrtZAcc, ivmeölçerin Z ekseninden elde edilen verilerin ortalama değeridir. StdXGyr, jiroskop algılayıcısının X ekseninden ölçülen değerlerin standart sapması, StdYGyr, jiroskop algılayıcısının Y ekseninden ölçülen değerlerin standart sapması, StdZGyr, jiroskop algılayıcısının Z ekseninden ölçülen değerlerin standart sapmasıdır. OrtXGyr, jiroskop algılayıcısının X ekseninde elde edilen verilerin ortalaması, OrtYGyr, jiroskop algılayıcısının Y ekseninde elde edilen verilerin ortalaması, OrtZGyr, jiroskop algılayıcısının Z ekseninde elde edilen verilerin ortalamasıdır.

Sınıflandırma işlemi için her bir algılayıcının ve algılayıcı kombinasyonlarının ne ölçüde başarılı olduğu, sonucu Tablo 2’de verilen testler sonucunda değerlendirilmiştir.

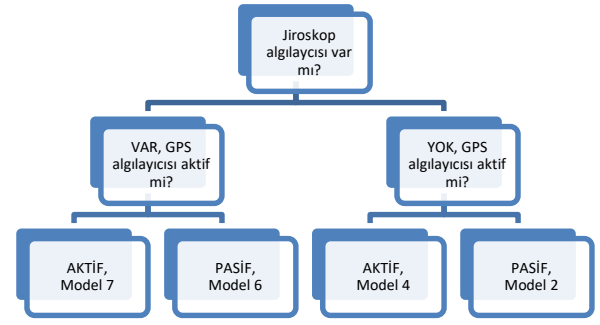
Tablo 2: Algılayıcı kombinasyonlarının, yöntemlere göre doğruluk oranları.

Model	1	2	3	4	5	6	7
Yöntem							
BA	80.6	97.1	80.4	97.7	91.7	98.4	97.9
NB	79.1	97.7	75.9	97.3	84.4	97	96.6
LR	76.5	97.6	78.1	97.6	88	97	97.2
kNN	78.4	99	84.5	98.8	89.2	98.9	98.8
C4.5	80.6	99.2	83.1	99	91.6	98.1	97.8
RF	78.7	99.1	90.5	98.8	93.3	99	99.4

Yapılan deneylerde her bir algılayıcı tek tek ve birbirleri ile kombinasyonlu bir şekilde teste tabi tutulmuştur. En iyi sonuç RF yöntemiyle GPS, ivmeölçer ve jiroskop algılayıcılarından elde edilen verilerin kombinasyonundan elde edilmiştir.

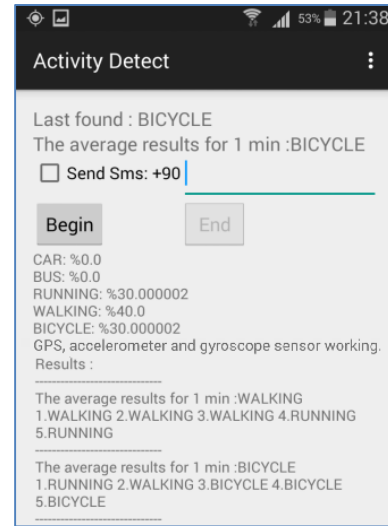
4.3 Çevrimiçi sınıflandırma

Gerçekleştirilen testler sonucunda en iyi sonucun GPS, ivmeölçer ve jiroskop algılayıcıları kombinasyonundan, RF yöntemi kullanılarak elde edildiği görülmüştür. Veri seti oluşturduğumuz akıllı telefon jiroskop algılayıcısına sahip olsa da bu algılayıcıya sahip olmayan akıllı telefonlar bulunmakta veya kullanıcı güç tasarrufu amacı ile GPS algılayıcısını kapalı durumda tutabilmektedir. Bu durumlarda geliştirilen Android uygulama Şekil 11’de gösterilen çalışma senaryosu uygulanarak sadece aktif olan algılayıcılardan veri toplanıp o algılayıcı veya algılayıcı kombinasyonunun en başarılı sonuç verdiği yöntemin kullanılacağı şekilde ayarlanmıştır. Geliştirilen uygulamada sınıflandırma hesaplamaları için Weka API [26] kullanılmıştır.



Şekil 11: Uygulama çalışma senaryosu.

Arka planda da çalışabilen uygulama, çalışmaya başladığı anda 12 saniyede bir örüntü oluşturmakta ve bunu aktif olan algılayıcı grubuna göre RF veya C4.5 yöntemi ile değerlendirmektedir. Yöntemlerinin eğitiminde, kullanılacak parametrelerin seçimi aşamasında 1500 tane eğitim için, 1000 tane test için ayırdığımız 2500 satırlık verinin tamamı kullanılmıştır. Geliştirilen uygulama, test etmek amacıyla bulunan sonuçları uygulama ara yüzünde raporlamakta ve eğer istenirse bir dakikalık aralıklara SMS gönderebilmektedir. Ulaşım türü sınıflandırma işlemini gerçekleştiren Android uygulamanın ara yüz ekran görüntüsü Şekil 12’de gösterilmektedir.



Şekil 12: Sınıflandırma uygulaması ekran görüntüsü

5 Tartışma

Yapılan çalışmada ulaşım türü tespiti için Android işletim sistemine sahip akıllı telefondan, Tablo 1’de gösterilen modellerdeki algılayıcı kombinasyonlarından elde edilen veriler kullanılarak bir örüntü oluşturulmuştur. Oluşturulan örüntüler makine öğrenmesi yöntemleri kullanılarak test edilmiştir. Gerçekleştirilen testler sonucunda en iyi sonuç RF yöntemi ile Model 7 kullanılarak elde edilmiştir. RF yöntemi ile Tablo 1’de verilen modellerden elde edilen ulaşım türlerine göre elde edilen karmaşıklık matrisleri Tablo 3-9’de gösterilmiştir.

Tablo 3: Model 1 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	114	77	3	0	0
Otobüs	54	147	1	1	1
Bisiklet	4	2	171	29	2
Koşma	0	0	20	166	7
Yürüme	0	0	3	9	189

Tablo 4: Model 2 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	189	5	0	0	0
Otobüs	4	200	0	0	0
Bisiklet	0	0	208	0	0
Koşma	0	0	0	193	0
Yürüme	0	0	0	0	201

Tablo 5: Model 3 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	173	20	1	0	0
Otobüs	40	164	0	0	0
Bisiklet	0	1	196	2	9
Koşma	0	0	0	192	1
Yürüme	0	0	21	0	180

Tablo 6: Model 4 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	189	4	1	0	0
Otobüs	7	197	0	0	0
Bisiklet	0	0	208	0	0
Koşma	0	0	0	193	0
Yürüme	0	0	0	0	201

Tablo 7: Model 5 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	157	36	1	0	0
Otobüs	26	178	0	0	0
Bisiklet	0	0	207	0	1
Koşma	0	0	0	191	2
Yürüme	0	0	1	0	200

Tablo 8: Model 6 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	189	4	1	0	0
Otobüs	5	199	0	0	0
Bisiklet	0	0	208	0	0
Koşma	0	0	0	193	0
Yürüme	0	0	0	0	201

Tablo 9: Model 7 karmaşıklık matrisi.

	Araba	Otobüs	Bisiklet	Koşma	Yürüme
Araba	191	2	1	0	0
Otobüs	3	201	0	0	0
Bisiklet	0	0	208	0	0
Koşma	0	0	0	193	0
Yürüme	0	0	0	0	201

Modellere göre karmaşıklık matrisleri tabloları incelendiğinde, sadece GPS verileri kullanıldığında özellikle otobüs ve araba gibi benzer hızlarda seyahat eden araçların sınıflandırılmasında sorun yaşandığı gözlemlenmiştir. Bu araçlar şehir merkezinde kalabalık trafik ortamında bir bisiklet veya yürüme kadar ortalama hızı sahip olabilmektedir. Bu da sadece GPS verileri kullanarak sınıflandırmada doğruluk oranını düşürmektedir. Sadece ivmeölçer algılayıcısı kullanıldığı durumda doğru sınıflandırma oranının büyük oranda arttığı hatta yürüme, koşma ve bisiklet sürme eylemlerinin tahmininde %100 oranında bir başarı sağlandığı gözlemlenmiştir. Sadece jiroskop algılayıcısı kullanılarak yapılan deneylerde, sadece GPS verileri ile yapılan deneyde olduğu gibi otobüs ve araba verilerin ayrılmasında sorunlar yaşandığı görülmüştür. Üç algılayıcının kombinasyonundan elde edilen veri seti ile işlem yapıldığında toplamda %99.4 oranında doğru sonuç elde edildiği görülmüştür.

6 Sonuç

Bu çalışmada Android işletim sistemine sahip akıllı telefon algılayıcıları kullanılarak ulaşım türü tespiti gerçekleştirilmiştir. Kullanılacak algılayıcı ya da algılayıcı kombinasyonu ve sınıflandırma yöntemi yapılan testler sonucunda belirlenmiştir. Literatür incelendiğinde eylem tanıma ve ulaşım türü sınıflandırmalarında ivmeölçer ve GPS algılayıcılarından elde edilen verilerden faydalandığı görülmektedir. Bu çalışmada tespit edilmek istenen beş farklı ulaşım türünün tanınmasında sadece ivmeölçer algılayıcısının başarılı sonuçlar verdiği görülmüştür. Fakat bu yüksek başarı oranları sadece jiroskop algılayıcısı ve sadece GPS algılayıcısı ile elde edilememiştir. Yalnızca ivmeölçer algılayıcısına göre bu algılayıcının GPS ve jiroskop algılayıcıları ile yapmış olduğu kombinasyonlardan elde edilen doğruluk oranları sadece BA yönteminde artış göstermiştir. GPS ve jiroskop algılayıcılarının ivmeölçer algılayıcısı ile oluşturduğu kombinasyonun ise veriyi zenginleştirerek doğruluk oranını artırdığı

gözlemlenmiştir. Sınıflandırma yöntemi olarak ise 6 farklı makine öğrenmesi algoritması denenmiştir. Bu çalışmada kullanılan veri setine en uygun yöntemin RF yöntemi olduğu belirlenmiştir. Gelecek çalışmalarda ulaşım türü tespiti uygulamasının çeşitli yollarla geliştirilmesi mümkündür. Bunlardan birincisi tren, metro, motosiklet gibi araçlar ekleyerek sınıf sayısını çoğaltmak, bir diğeri akıllı telefonu farklı konumlarda bulundurarak eğitim setini genişletmek, doğruluğu artırmak için daha etkili ve az sayıda öznetelik seçimi yapmaktır. Ayrıca bu çalışma, akıllı telefon algılayıcılarına sahip olan giyilebilir akıllı cihazlar ile genişletilecektir.

7 Kaynaklar

- [1] Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT. "A survey of mobile phone sensing". *Communications Magazine*, 48(9), 140-150, 2010.
- [2] Reddy S, Burke J, Estrin D, Hansen M, Srivastava M. "Determining transportation mode on mobile phones". *Wearable Computers, 12th IEEE International Symposium*, Pittsburgh, USA, 28 September-1 October 2008.
- [3] Zheng Y, Liu L, Wang L, Xie X. "Learning transportation mode from raw GPS data for geographic applications on the web". *17th World Wide Web Conference*, Beijing, China, 21-25 April 2008.
- [4] Györfbíró N, Fábíán Á, Hományi G. "An activity recognition system for mobile phones". *Mobile Networks and Applications*, 14(1), 82-91, 2009.
- [5] Wang S, Chen C, Ma J. "Accelerometer based transportation mode recognition on mobile phone". *2010 Asia-Pacific Conference*, Shenzhen, China, 17-18 April 2010.
- [6] Stenneth L, Wolfson O, Yu FS, Xu B. "Transportation mode detection using mobile phones and GIS information". *19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Chicago, IL, USA, 1-4 November 2011.
- [7] Lara OD, Pérez AJ, Labrador MA, Posada JD. "Centinela: A human activity recognition system based on acceleration and vital sign data". *Pervasive and Mobile Computing*, 8(5), 717-729, 2012.
- [8] Widhalm P, Nitsche P, Brandie N. "Transport mode detection with realistic Smartphone sensor data". *21st International Conference on Pattern Recognition (ICPR 2012)*, Tsukuba, Japan, 11-15 November 2012.
- [9] Kose M, Incel OD, Ersoy C. "Online human activity recognition on smart phones". *2nd International Workshop on Mobile Sensing*. Beijing, China, 16 April 2012.
- [10] Bolbol A, Cheng T, Tsapakis I, Haworth J. "Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification". *Computers, Environment and Urban Systems*, 36(6), 526-537, 2012.
- [11] Hemminki S, Nurmi P, Tarkoma S. "Accelerometer-based transportation mode detection on smartphones". *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, Rome, Italy, 11-14 November 2013.
- [12] Feng T, Timmermans HJP. "Transportation mode recognition using GPS and accelerometer data". *Transportation Research Part C: Emerging Technologies*, 37, 118-130, 2013.
- [13] Ellis K, Godbole S, Marshall S, Lanckriet G, Staudenmayer J, Kerr J. "Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms". *Frontiers in Public Health*, 2(36), 1-8, 2014.
- [14] Shin D, Aliaga D, Tunçer B, Arisona SM, Kim S, Zünd D, Schmitt G. "Urban sensing: Using smartphones for transportation mode classification". *Computers, Environment and Urban Systems*, 53, 76-86, 2015.
- [15] Sökün H, Kalkan H, Cetişli B. "Classification of physical activities using accelerometer signals". *Signal Processing and Communications Applications Conference (SIU)*, Muğla, Turkey, 18-20 April 2012.
- [16] Xia H, Qiao Y, Jian J, Chang Y. "Using smart phone sensors to detect transportation modes". *Sensors*, 14(11), 20843-20865, 2014.
- [17] El-Rabbany A. *Introduction to GPS: The Global Positioning System*. 2nd ed. Norwood, USA, Artech House, 2002.
- [18] Su X, Tong H, Ji P. "Activity recognition with smartphone sensors". *Tsinghua Science and Technology*, 19(3), 235-249, 2014.
- [19] Sağbaş E.A, Ballı S. "Akıllı Telefon Sensörlerinin Kullanımı ve Ham Sensör Verilerine Erişim". *Akademik Bilişim Konferansı*, Eskişehir, Türkiye, 4-6 Şubat 2015.
- [20] Chandra B, Gupta M, Gupt MP. "Robust approach for estimating probabilities in Naive-Bayes classifier". *Pattern Recognition and Machine Intelligence*. Kolkata, India, 18-22 December 2007.
- [21] Korb KB, Nicholson AE. *Bayesian Artificial Intelligence*. 2nd ed. Boca Raton, FL, USA, CRC Press, 2011.
- [22] Feng T, Timmermans HJP. "Comparative evaluation of algorithms for GPS data imputation". *13th WCTR*, Rio de Janeiro, Brazil, 15-18 July 2010.
- [23] Akman M, Genç Y, Ankaralı H. "Random forest yöntemi ve sağlık alanında bir uygulama". *Türkiye Klinikleri*, 3(1), 36-48, 2011.
- [24] Samsung. "Galaxy Note 2". <http://www.samsung.com/tr/consumer/mobile-devices/smartphones/galaxy-note/GT-N7100RWDTUR> (08.07.2015).
- [25] Garner SR. "Weka: The waikato environment for knowledge analysis". *2nd New Zealand Computer Science Research Students Conference*, Hamilton, New Zealand, 18-21 April 1995.
- [26] Weka. "Use WEKA in your Java code". <https://weka.wikispaces.com/Use+WEKA+in+your+Java+code> (08.07.2015).