DÜMF Mühendislik Dergisi

web: http://dergipark.gov.tr/dumf

Yazma davranışlarının mobil cihaz sensörleri kullanılarak sınıflandırılması

Asım Sinan YÜKSEL*
Süleyman Demirel Üniversitesi, Bilgisayar Mühendisliği Bölümü, Isparta
asimyuksel@sdu.edu.tr, Tel: (246) 241 13 80

Fatih Ahmet ŞENEL
Süleyman Demirel Üniversitesi, Bilgisayar Mühendisliği Bölümü, Isparta

İbrahim Arda ÇANKAYA
Süleyman Demirel Üniversitesi, Bilgisayar Mühendisliği Bölümü, Isparta

Geliş: 03.08.2017, Kabul Tarihi: 09.11.2017

Öz

Sensör teknolojilerinin gelişimi ve mobil cihazlara uyumu ile birlikte mobil kullanıcı davranışlarının ve aktivitelerinin modellenmesi için yeni fırsatlar ve çalışma alanları ortaya çıkmıştır. Mobil cihaz sensörleri özellikle ayakta durma, koşma, yürüme, basamak çıkma gibi fiziksel aktivitelerin otomatik olarak sınıflandırılması, sağlık hizmetlerinde hastaların izlenmesi ve gelişmiş kullanıcı arayüzlerinin geliştirilmesi gibi uygulamalar için oldukça çekici hale gelmiştir.

Bu çalışmada diğer çalışmalardan farklı olarak yazma eylemi üzerinde durulmuştur. Tasarlanan mobil bir arayüz aracılığı ile aynı cihaz üzerinde 30 farklı kullanıcıdan 20 farklı cümleyi yazmaları istenmiş; yazım esnasında mobil cihazın üç eksenli ivmeölçer ve jiroskop sensörlerinden alınan veriler kaydedilmiştir. Elde edilen veriler önişlemden geçirildikten sonra en uygun öznitelik vektörleri elde edilmiş, bu öznitelikler kullanılarak da kullanıcıların yazma davranışları sınıflandırılmıştır. Sınıflandırmada en çok kullanılan algoritmalardan Yapay Sinir Ağları (YSA), k En Yakın Komşuluk (k-EK), Destek Vektör Makineleri (DVM) algoritmaları uygulanmıştır. Yaptığımız kullanıcı çalışmalarında YSA ile %97.91, k-EK ile %100, DVM'de ise %97.91 başarı oranı yakalanmıştır. En başarılı sonuçlar k-EK ile 1 komşuluk sayısı kullanılarak elde edilmiştir.

Çalışmamız sadece mobil cihazdan alınan sensör verileri üzerinde çalışması ve oluşturulan öznitelikler ile de yüksek ikili sınıflandırma başarısına ulaşması nedeniyle literatüre büyük katkı sağlamaktadır. Buna ek olarak her insanın kendine has yazma davranışının olduğunu ortaya çıkarılmıştır. Mobil cihaz kullanıcılarının yazma davranışları biyometrik özellikler göstermekte olup kullanıcıları birbirinden ayırt edecek, cihazların istenmeyen kişilerin eline geçmesi durumunda gerekli güvenlik önlemlerini alacak bir güvenlik mekanizması olarak kullanılabilir.

Anahtar Kelimeler: *Fiziksel aktivite tanıma; sınıflandırma; mobil cihaz sensörleri; yazma davranışı;*

* Yazışmaların yapılacağı yazar

DOI:

Giriş

Tablet, akıllı telefon gibi mobil cihazların popülerleşmesi ve kullanımının yaygınlaşmasıyla birlikte insanlar neredeyse tüm işlerini sahip oldukları mobil cihazlarla yapar hale gelmişlerdir. Buna bağlı olarak mobil cihaz satışlarında son yıllarda büyük bir artış gözlemlenmiştir. Teknoloji sürekli gelişmekte ve daha akıllı, daha fazla özellik içeren cihazlar ortaya çıkmaktadır. Bu cihazlar arama, mesaj gönderme dışında depolama, mobil bankacılık, oyun oynama, yön bulma, alışveriş gibi işlemler için de kullanılabilir hale gelmiştir. Mobil cihazlara eklenen sensörler sayesinde de insan aktivitelerini (koşma, yürüme, bisiklet sürme) tanıma, davranışlarını izleme ve raporlama, nabız kontrolü gibi işlemler gerçekleştirilebilmektedir (Mannini ve Sabatini, 2010). Bu da araştırmacılar için yeni çalışma alanlarının ortaya çıkmasını sağlamıştır.

Günümüzde hemen hemen her akıllı telefonda üç eksenli ivmeölçer, jiroskop ve manyetometre gibi sensörler standart olarak bulunmaktadır. Bu sensörlerin devasa boyutlarda veri üretebilmeleri özellikle yapay zekâ, makine öğrenmesi, yaygın bilişim ve çevre ayırdında bilişim ile uğraşan araştırmacıların ilgisini çekmiş ve bunun neticesinde son yıllarda bu sensör verilerini yorumlayan, sınıflandıran çalışmaların sayısı artmıştır. Yapılan çalışmaların çoğu da yürüme, koşma, uyuma, bisiklet sürme gibi insan aktivitelerini tanıma ve sınıflandırma yönünde yoğunlaşmıştır.

(Machad vd., 2015) tarafından yapılan çalışmada ivmeölçer tabanlı aktivite tanıma üzerine yoğunlaşılmış ve yeni bir interaktif bilgi keşfi çatısı geliştirilmiştir. Sensör bilgilerinden elde edilen özellikler sayesinde insan aktivitelerinin tanınmasında ve sınıflandırılmasında kişiye bağımlı olarak %99.29, kişiden bağımsız olarak ise %88.57 oranında başarı elde edilmiştir. (Alberts vd., 2015) tarafından yapılan başka bir çalışmada Apple iPad marka tablet türündeki mobil cihazın ivmeölçer ve jiroskop gibi sensörlerinden veri alınarak duruş bozukluğu tespiti yapılıp yapılamayacağı araştırılmıştır. iPad 2 cihazının yeterli duyarlılıkta ve doğrulukta veriler sağladığı, klinik çevrelerde duruş bozukluğunun tespitinde böyle bir yaklaşımın kullanılabileceği ispatlanmıştır. Başka bir çalışmada ise mobil cihazlarda yer alan üç eksenli ivmeölçer sensörü aracılığı ile el hareketlerinin tanınması için bir metot geliştirilmiştir (Choi vd., 2005). Sistem havada yazılan 1 ile 9 arasındaki sayıları ve 5 ayrı sembolü tanıyabilmekte, interaktif oyun oynama ve müzik aleti çalabilme gibi işlemler için üretilen el hareketlerini tanıyabilmektedir. İstem dışı gelişen insan hareketlerinden biri de düşme eylemidir. 65 yaşın üstündeki insanların %33'ünden fazlası senede en az bir defa düşmektedir. Dolayısıyla bu tarz bir eylemin tespiti ve gerçekleşmesi sonucunda gerekli işlemlerin ivedilikle yapılması önem arz etmektedir. Mobil cihazlarda yer alan ivmeölçer ve jiroskop verileriyle çalışan algoritmalar ile bu tarz aktiviteler %91 oranında başarıyla tespit edilebilmektedir (Li vd., 2009).

Bir mobil cihaz kendi başına bir sensör platformu olarak kullanılarak kişinin emniyetini ya da güvenliğini sağlamak için bir mekanizma haline getirilebilir. Özellikle araç kullanırken sürücünün kendi emniyetini ve diğer sürücülerin emniyetini sağlamak, kazaların önüne geçebilmek için faydalı olabilir. Sürücülerin davranışları izlendiğinde; sürüş davranışlarıyla ilgili geribildirim yapıldığında, sürüş esnasındaki agresif sürüş davranışları raporlandığında, sürücüler için daha güvenli bir sürüş deneyimi sağlanabilir (Johnson ve Trivedi, 2011). Mobil cihazın güvenliğinin sağlanması, yetkisiz kişilerin kullanımın önlenmesi ve cihaz kullanımının kısıtlanması için de yine mobil cihazlarda yer alan ivmeölçer, jiroskop, manyetometre verileri kullanılabilir. Kullanıcıya ait mobil cihaz kullanım davranışları ortaya çıkarılabilir. Bu sayede mobil cihazın sadece sahibi tarafından kullanabileceği, farklı bir kullanım

algılandığında da güvenlik önlemlerinin oluşturulacağı sistemler geliştirilebilir (Zhu vd., 2013).

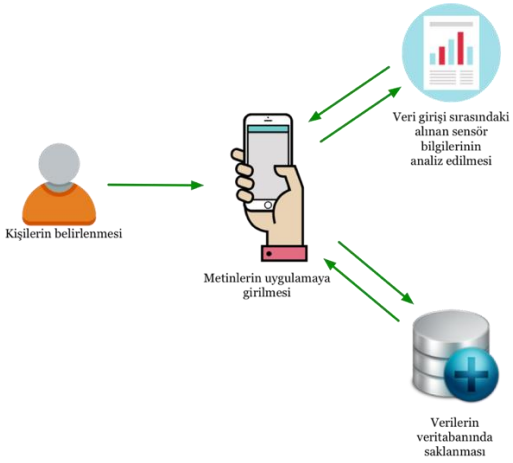
Bu çalışmada, ivmeölçer ve jiroskop gibi sensörlerden kaynaklanan faktörler üzerinde odaklanılmıştır. Bu faktörler bir kişinin fiziksel özelliklerini tanımlamaktadır. Kullanıcının telefonu nasıl tuttuğu, telefonun klavyesindeki tuşlara nasıl bastığı gibi gözlemlediğimiz çoğu karakterler biyometrik ve kişiyi tanımlayıcı özelliktedir. Çalışmamızda diğer çalışmalardan farklı olarak yazma eylemi üzerinde durulmuştur. Veri toplama işlemi Apple iPhone 7 Plus marka telefon üzerinde bulunan ivmeölçer, jiroskop sensörleri aracılığı ile gerçekleştirilmiştir. Önceki çalışmalara bakıldığında insan aktivitelerinin tanınması için genellikle vücuda bağlı sensörlerin kullanıldığı ve genelde yürüme, koşma, bisiklet sürme gibi eylemlerin tanınmasına odaklanıldığı görülmektedir. Bu çalışmada ise mobil bir cihazdan doğrudan ayrıştırılmış veriler kullanılmış ve çeşitli popüler algoritmalar kullanılarak ikili sınıflandırma yapılmış ve elde edilen başarı oranları sunulmuştur.

Materyal ve yöntem

Mobil uygulama

Bu çalışmada ivmeölçer ve jiroskop verilerini elde etmek için iPhone 7 Plus telefon kullanılmıştır. Verilerin alınması için basit bir uygulama geliştirilmiştir. 20 adet metin belirlenmiş ve kişilerden bu metinleri uygulama ekranındaki metin kutusuna girmeleri istenmiştir. Metin kutusuna istenilen metinler girilirken de ivmeölçer ve jiroskop verileri, metnin girildiği süre bilgileri kaydedilmiştir. Çalışmanın sistem mimarisi Şekil 1'de gösterilmektedir.

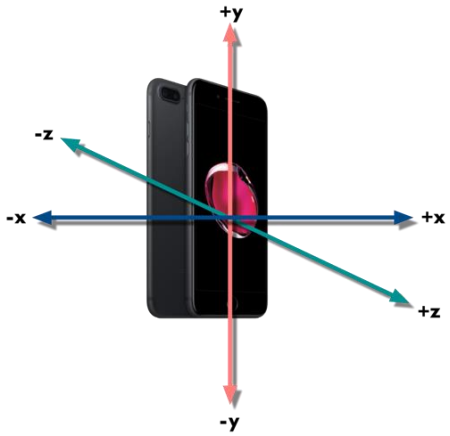
İvmeölçer sensörleri X, Y ve Z eksenlerinde telefona etki eden ivmelenmeyi ölçmektedir. İvmeölçer verileri, Şekil 2'de olduğu gibi X, Y ve Z eksenlerinden alınmıştır ve g cinsinden yani 9,81m/s^2 değerlerini vermektedir. İvmeölçer sensöründen alınan bilgilerde X, telefonun yatay eksenindeki; Y, telefonun dikey eksenindeki ve Z ise telefonun sırt üstü durma eksenindeki pozisyonunu göstermektedir.



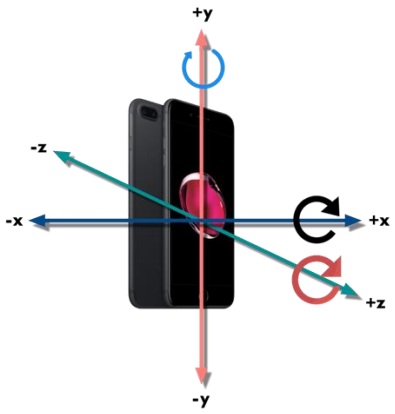
Şekil 1.Sistem mimarisi

İvmeölçer verileri toplama sıklığı kullanıcının ihtiyacına yönelik ayarlanabilmektedir. Bu ayarlamalar sayesinde zaman içerisindeki veri miktarı belirlenebilmektedir. Günümüzde birçok hareket uygulamasında da ivmeölçer verileri kullanılarak telefonun pozisyon bilgisine göre kullanıcı ile ilgili yorumlamalar yapılabilmektedir (LeMoyne vd., 2010; Győrbíró vd., 2009).

Jiroskop sensörleri akıllı telefonlarda cihazın X, Y ve Z eksenlerinde yapmış olduğu açısal hız bilgisini vermektedir. Jiroskop verileri hareket tanıma işlemlerinde ivmeölçer ile birlikte netlik miktarını artırmaktadır (Su ve Ji, 2014). Şekil 3'te jiroskop sensörü eksenleri gösterilmektedir.

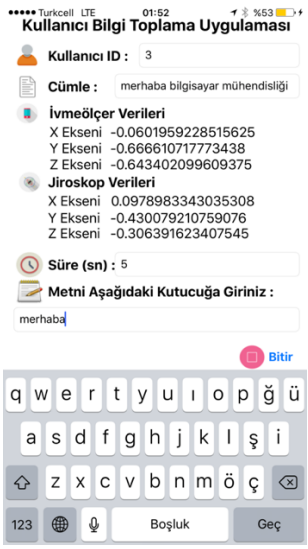


Şekil 2.İvmeölçer sensörü



Şekil 3.Jiroskop sensörü

Mobil uygulama ekranında kişi kendine atanan kimlik numarasını girdikten sonra ekranda belirtilen hazır metnin aynısını belirtilen yerdeki metin kutusuna girer.



Şekil 4.Uygulama ekranı

Kişi bu bilgiyi girerken uygulama tarafından o andaki ivmeölçer, jiroskop ve süre bilgileri bir metin dosyasına kaydedilir. Kişi hazır metni 20 kere girdikten sonra uygulama sonlanır. Uygulamanın ekran görüntüsü Şekil 4'te gösterilmektedir. Çalışmada geliştirilen mobil uygulama ile birlikte kişiye atanan kimlik numarası, belirlenen hazır metin, kişinin saniye cinsinden metni yazma süresi, ivmeölçer verileri, jiroskop verileri gibi bilgiler toplanmıştır.

Veri seti oluşturma

Çalışmada her bir kullanıcıdan farklı cümleleri telefonun klavyesini kullanarak 20'şer defa girmeleri istenmiştir. Kullanıcının cümleyi girme süresi boyunca her saniyede beş defa Jiroskop (Gyro) ve İvmeölçer (Acc) sensörlerinden veriler alınmıştır. 30 farklı kullanıcı ve her kullanıcı için 20 farklı cümle örneği alınmıştır. Kullanıcıların bir cümleyi girmeleri süresince alınan veriler kullanılarak toplamda 26 adet öznitelik ve 600x26 boyutunda veri seti oluşturulmuştur. Bu öznitelikler ve açıklamaları Tablo 1'de detaylı olarak verilmiştir. Her bir satırdaki öznitelik, kullanıcının bir cümleyi girdiği süre boyunca oluşan verilerden elde edilmiştir. Tablo 1'de kullanılan büyüklük değerinin hesaplanması Denklem 1'de verilmiştir. Burada, Mag değişkeni büyüklük değerini, X, Y ve Z değerleri ise jiroskop ve ivmeölçer sensörlerinden alınan verilerin ortalamalarını ifade etmektedir.

$$Mag = \sqrt{X^2 + Y^2 + Z^2} \quad (1)$$

Korelasyon tabanlı öznitelik seçimi

Bu çalışmada Korelasyon Tabanlı Öznitelik Seçimi (KTÖS) algoritması kullanılmıştır. KTÖS, sınıflandırma işlemini başarılı bir şekilde yapabilen, öznitelik alt kümesi oluşturan bir öznitelik filtreleme yöntemidir (Mursalin vd., 2017; Hall ve Smith, 1999). KTÖS algoritmasına göre, iyi bir öznitelik alt kümesi, sınıf etiketiyle yüksek korelasyona, kendi aralarında ise düşük korelasyona sahip olmalıdır. KTÖS, bu şartlar altındaki en iyi alt kümeyi oluşturmaktadır. KTÖS algoritması kısmen sezgisel olarak çalışmaktadır. Sezgiselliğin hesaplanması Denklem 2'de gösterilmektedir.

Tablo 1. Kullanılan öznitelikler ve açıklamaları

Öznitelik Adı	Açıklaması
minGyroX	Gyro sensöründen alınan X değerlerinin minimum olanı
minGyroY	Gyro sensöründen alınan Y değerlerinin minimum olanı
minGyroZ	Gyro sensöründen alınan Z değerlerinin minimum olanı
minAccX	Acc sensöründen alınan X değerlerinin minimum olanı
minAccY	Acc sensöründen alınan Y değerlerinin minimum olanı
minAccZ	Acc sensöründen alınan Z değerlerinin minimum olanı
maxGyroX	Gyro sensöründen alınan X değerlerinin maksimum olanı
maxGyroY	Gyro sensöründen alınan Y değerlerinin maksimum olanı
maxGyroZ	Gyro sensöründen alınan Z değerlerinin maksimum olanı
maxAccX	Acc sensöründen alınan X değerlerinin maksimum olanı
maxAccY	Acc sensöründen alınan Y değerlerinin maksimum olanı
maxAccZ	Acc sensöründen alınan Z değerlerinin maksimum olanı
meanGyroX	Gyro sensöründen alınan X değerlerinin ortalaması
meanGyroY	Gyro sensöründen alınan Y değerlerinin ortalaması
meanGyroZ	Gyro sensöründen alınan Z değerlerinin ortalaması
meanAccX	Acc sensöründen alınan X değerlerinin ortalaması
meanAccY	Acc sensöründen alınan Y değerlerinin ortalaması
meanAccZ	Acc sensöründen alınan Z değerlerinin ortalaması
stdGyroX	Gyro sensöründen alınan X değerlerinin standart sapması
stdGyroY	Gyro sensöründen alınan Y değerlerinin standart sapması
stdGyroZ	Gyro sensöründen alınan Z değerlerinin standart sapması
stdAccX	Acc sensöründen alınan X değerlerinin standart sapması
stdAccY	Acc sensöründen alınan Y değerlerinin standart sapması
stdAccZ	Acc sensöründen alınan Z değerlerinin standart sapması
magGyro	Gyro sensöründen alınan X, Y, Z değerleri ortalamasının büyüklük değeri
magAcc	Acc sensöründen alınan X, Y, Z değerleri ortalamasının büyüklük değeri

$$M_S = \frac{k^* r_{cf}}{\sqrt{k + k^*(k\text{-}1)^* r_{ff}}} \quad (2)$$

Bu denklemde, M_S öznitelik alt kümesinin değerini; k, öznitelik sayısını; r_{cf}, ortalama özniteliklerin sınıf etiketiyle olan korelasyonunu; r_{ff} ise öznitelik alt kümesinde bulunan öznitelikler arası ortalama korelasyonu ifade etmektedir. Denklem 2 ile birbirleriyle yüksek korelasyona sahip özniteliklerin aynı alt kümede bulunması engellenmektedir. Diğer taraftan sınıf etiketiyle yüksek korelasyona sahip özniteliklerin ise alt kümeye dâhil olması sağlanmaktadır. Sınıflandırma problemlerinde genellikle öznitelikler kategori bazlı ifade edilmektedirler. Yani tüm öznitelikler aynı türden olmayabilir. Bu durumda öznitelikler arası korelasyon hesaplanmadan önce özniteliklerin bir ön işleme tabi tutularak ortak bir kategori türünde ifade edilmeleri gerekmektedir. Bu işlem için entropi değeri kullanılmaktadır. İki adet (X, Y) rasgele ayrık değişkende, X gözleminden önce ve sonraki Y değişkeninin entropi değeri Denklem 3 ve 4'te gösterilmiştir.

$$H(Y) = -\sum_{y \in Y} p(y) log_2 p(y) \quad (3)$$

$$H(Y|X) = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) log_2 p(y|x) \quad (4)$$

Burada, H(Y), Y'nin olasılığını; H(Y|X), X bilindiğinde Y'nin olasılığını ifade etmektedir. X değişkeni ile Y değişkeni hakkında bilgi sağlandıkça, Y değişkeninin entropi değeri azalma eğilimi göstermektedir. Denklem 5, Y değişkeninin X değişkenine olan bağlılığını (korelasyonu) ifade etmektedir.

$$C(Y|X) = \frac{H(Y)\text{-}H(Y|X)}{H(Y)} \quad (5)$$

Burada, C(Y|X), Y değişkeninin X değişkenine bağlılık oranını ifade etmektedir. KTÖS ile seçilen öznitelik alt kümesi Tablo 2'de verilmiştir.

Tablo 2. Seçilen öznitelik alt kümesi

Öznitelik Adı	Açıklaması
maxGyroZ	Gyro sensöründen alınan Z değerlerinin maksimum olanı
maxAccY	Acc sensöründen alınan Y değerlerinin maksimum olanı
meanAccX	Acc sensöründen alınan X değerlerinin ortalaması
meanAccY	Acc sensöründen alınan Y değerlerinin ortalaması
meanAccZ	Acc sensöründen alınan Z değerlerinin ortalaması
stdGyroZ	Gyro sensöründen alınan Z değerlerinin standart sapması

Çapraz doğrulama tekniği

Geliştirilen tüm modellerde Çapraz Doğrulama Tekniği uygulanmıştır. Bu teknik ile modellerin eğitimi aşamasında ezberlemenin önüne geçilmektedir. Çapraz Doğrulama (ÇD), veri setini k adet eşit parçaya böler. Her seferinde parçalardan birini test işlemi için, geriye kalan k-1 parça da eğitim işlemi için kullanır. Tüm parçalar hem eğitim hem de test işlemi için kullanıldıktan sonra tüm sonuçların ortalaması alınarak sonuç elde edilir.

Performans ölçümü

Modellerin çıktıları karşılaştırılırken, üç farklı performans ölçütü dikkate alınmıştır. Bunlar, Ortalama Mutlak Hata (OMH), Ortalama Karesel Hatanın Karekökü (OKHK) ve Doğru Sınıflandırma Yüzdesi (DSY)'dir. OMH, OKHK ve DSY hesaplama işlemleri sırasıyla Denklem 6, 7 ve 8'de verilmiştir. Burada N toplam örnek sayısını, A gözlemlenen değeri, P tahmin değeri, A_{ort} gerçek değerlerin ortalamasını ve son olarak DT ise doğru tahmin edilen örnek sayısını ifade etmektedir.

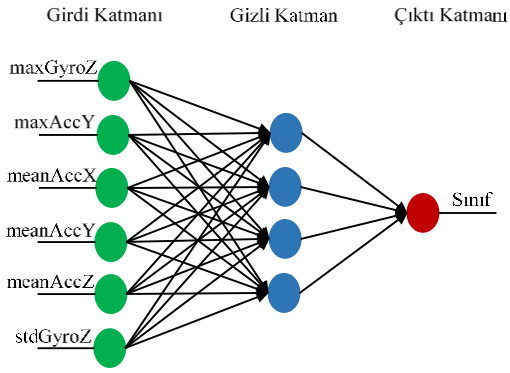
$$OMH = \frac{\sum_{i=1}^{N} |P_i\text{-}A_i|}{N} \quad (6)$$

$$OKHK = \sqrt{\frac{\sum_{i=1}^{N} (P_i\text{-}A_i)^2}{N}} \quad (7)$$

$$DSY = 100 * \frac{DT}{N} \quad (8)$$

Yapay Sinir Ağları

YSA, sınıflandırma problemlerinde kullanılan popüler bir öğrenme algoritmasıdır. İlk defa 1943 yılında McCulloch ve Pitts (1943) tarafından öne sürülmüştür. Temel olarak YSA; girdi katmanı, gizli katman ve çıktı katmanlarından oluşmaktadır. Girdi katmanında, problemin girdi sayısı kadar nöron, çıktı katmanında ise problemin çıktı sayısı kadar nöron bulunmaktadır. Gizli katmanda ise problemin başarısını etkileyen ve kullanıcı tarafından belirlenen sayıda katman ve nöronlar bulunmaktadır. Bu çalışmada kullanılan YSA modeli farklı gizli katman sayıları ve nöronlarında test edilerek en uygun sonucu veren model tespit edilmiştir. Altı adet öznitelik seçildiğinden dolayı YSA modeli altı girişli ve tek çıkışlı olacak şekilde tasarlanmıştır. Gizli katman sayısı bir olarak seçilmiş ve nöron sayısı da dört olarak belirlenmiştir. Yapılan tüm eğitim işlemlerinde çapraz doğrulama testi uygulanmıştır. Çapraz doğrulama testi tüm modeller için daha iyi bir öğrenme işlemi gerçekleştirmektedir. Ancak nöron sayılarının belirlenmesinde en uygun nöron sayısı deneme yanılma yöntemiyle bulunmuştur. Şekil 5'te kullanılan YSA modeli gösterilmiştir.



Şekil 5.Kullanılan YSA modeli

k En Yakın Komşuluk Algoritması

k-EK, sınıflandırma işleminde kullanılan danışmanlı bir sınıflandırma algoritmasıdır. Kullanıcı tarafından belirlenen bir k sayısı kadar komşuya bakarak, yeni gelen bir örneğin hangi sınıfa ait olduğuna karar verme mantığıyla çalışmaktadır (Cover ve Hart, 1967). Yakınlık değeri ölçülürken minkowski, öklid, manhattan gibi ölçütler kullanılmaktadır. En yaygın olarak öklit uzaklık ölçütü kullanıldığından çalışmada öklit uzaklık ölçütü seçilmiştir. Ayrıca farklı k komşuluk değerleri ile problem test edilmiş ve en uygun k değerinin 1 olduğu tespit edilmiştir.

Destek Vektör Makineleri

DVM, sınıflandırma ve regresyon işlemlerinde kullanılan gelişmiş bir algoritmadır. Temel olarak iki farklı sınıfı birbirinden ayıran çok boyutlu düzlemin tahmin edilmesi amacına göre çalışmaktadır (Cristianini ve Shawe-Taylor, 2000; Haykin, 2009). Çok boyutlu düzlem belirlenirken iki sınıf arası mesafenin en fazla olacağı şekilde işlemler gerçekleştirilir. DVM doğrusal olarak ayrılabilen problemlerde kullanılabildiği gibi doğrusal olmayan problemlerde de başarıyla kullanılabilmektedir. Eğer problem uzayı doğrusal olarak sınıflandırmaya uygun değilse, problem daha yüksek dereceden bir problem uzayına dönüştürülerek doğrusal sınıflandırıcılarla sınıflandırılabilmektedir. DVM, doğrusal ayrılamayan problemleri farklı bir uzaya dönüştürerek başarıyla sınıflandırabilmektedir. Bu çalışmada radyal tabanlı çekirdek (Rbf), polinomsal çekirdek (Poly) ve Pearson VII fonksiyonu tabanlı çekirdek (Puk) kullanılarak DVM ile sonuçlar elde edilmiştir. Buna ek olarak, kullanılan çekirdek haricindeki tüm parametreler için WEKA programındaki varsayılan değerler kullanılmıştır. C değeri ise 1 olarak seçilmiştir.

Uygulama ve başarımlar

Üç farklı modelleme yöntemi, akıllı telefondan alınan sensör verileri ile çalıştırılmıştır. Modeller kullanılarak akıllı telefonun sahibi tarafından mı yoksa farklı bir kişi tarafından mı kullanıldığı tespit edilmiştir. Tablo 3'te elde edilen ikili sınıflandırma sonuçları gösterilmiştir. Sonuçlar incelendiğinde en başarılı algoritmanın k-EK algoritması olduğu görülmektedir. DVM ve YSA'da kabul edilebilir başarılı sonuçlar elde edilmiştir. Fakat k-EK algoritmasına göre daha az başarılı sonuçlar üretmişlerdir. YSA'da gizli katmandaki nöron sayısını artırmanın belli bir zamandan sonra sonucu etkilemediği, en ideal nöron sayısının 4 olduğu gözlemlenmiştir. DVM'de doğru çekirdek tipi seçiminin başarı oranını doğrudan etkilediği görülmüştür. En yüksek başarı oranı Pearson VII fonksiyonu tabanlı çekirdek kullanılarak elde edilmiştir. k-EK'de ise komşuluk sayısındaki artış belli noktadan sonra başarı oranını değiştirmemiştir. Komşu sayısının 1'e eşit olduğu durumda %100 başarı oranı yakalanmıştır. Sonuç olarak bir cihazın kullanıcısı dışında kullanımının tespiti yüksek başarı oranıyla yapılabilmektedir.

Sonuçlar ve tartışma

Bu çalışmada, insanların yazma davranışlarının mobil cihaz sensörleri kullanılarak sınıflandırılması amaçlanmıştır. Çalışmada basit bir arayüz geliştirilerek mobil cihazdan sensör verileri alınmış ve özgün bir veri kümesi oluşturulmuştur. Sınıflandırıcı açısından incelendiğinde k-EK'nin YSA ve DVM sınıflandırıcılarına göre oldukça iyi bir başarı oranı yakaladığı görülmektedir.

Kullanılan veri seti boyutunun artırılmasının başarı oranını kısmen düşüreceği öngörülmektedir. K-EK algoritması ile hızlı bir

şekilde sonuca ulaşılabilmesi, ikili sınıflandırmanın gerçek zamanlı olarak işletilmesinin önünü açmaktadır. İkili sınıflandırma performansının artırılması için ve daha başarılı sonuçlar elde etmek için gerçek zamanlı bir uygulama geliştirebilir. Bu uygulama aracılığı ile elde edilecek yeni verilerle yeniden eğitim yapılarak model güncellenebilir. Çalışma sadece iOS platformunda iPhone 7 plus cihaz üzerinde denenmiştir. Farklı cihazlarda sensör duyarlılıkları farklı olacağından ölçülen değerlerde farklılıklar olacaktır. Ancak yazma davranışı sırasında eksenlerde oluşacak değişimler benzer olacaktır. Farklı cihazların ve farklı ekran boyutlarının performansa etkisinin detaylı analizi gelecek çalışma olarak geliştirilebilir.

Kullanılan sınıflandırma yöntemleri, eğitim aşamasını çok hızlı bir şekilde tamamlayabilmektedir. YSA, diğer iki sınıflandırma yönteminden zaman olarak geride kalsa da tüm yöntemler gerçek zamanda çalışabilmeye uygun hızlardadır. Tüm yöntemler ortalama olarak, kullanılan veri setimiz için 130 milisaniyede eğitim işlemini tamamlamaktadır. Kullanılan veri seti boyutuna ve donanıma göre zaman farklılıkları oluşabilir.

Sonuç olarak; kullanılan modellerin hepsi de telefonu kullanan kişinin telefonun gerçek sahibi olup olmadığını tespit edebilecek düzeydedir. Bu çalışmada elde edilen sonuçlar, sadece sensör bilgilerini kullanarak telefonun yabancı bir kişi tarafından kullanılıp kullanılmadığının tespit edilebilir olduğunu göstermiştir. Bunun yanında her kişinin kendine has yazma tarzı olduğu; telefonu tutma, tuşlara basma gibi işlemlerin kişiden kişiye değişiklik gösterdiği ve bunların biyometrik özellikler olarak kullanılabileceği ortaya koyulmuştur. Elde edilen sonuçlar ışığında, bu çalışmada önerilen yöntem geliştirilecek bir mobil uygulamaya ya da mobil işletim sistemine dahil edilerek yazma davranışı tabanlı bir kimlik doğrulama mekanizmasına dönüştürülebilir.

Tablo 3. Sınıflandırma sonuçları

Performans Ölçütleri	**YSA**			**k-EK**			**DVM**		
	Gizli Katmandaki Nöron Sayısı			**k Komşuluk Sayısı**			**Çekirdek Tipi**		
	2	4	6	1	3	5	POLY	RBF	PUT
OMH	0.0515	0.043	0.041	0.0113	0.0314	0.0376	0.2396	0.3021	0.0208
OKHK	0.168	0.150	0.1476	0.0113	0.1078	0.1064	0.4895	0.5496	0.1443
DSY (%)	95.83	97.91	97.91	100	98.95	98.95	76.04	69.79	97.91

Kaynaklar

Alberts, J. L., Hirsch, J. R., Koop, M. M., Schindler, D. D., Kana, D. E., Linder, S. M. ve Thota, A. K., (2015). Using accelerometer and gyroscopic measures to quantify postural stability, *Journal of athletic training*, **50**, 6, 578-588.

Choi, E. S., Bang, W. C., Cho, S. J., Yang, J., Kim, D. Y. ve Kim, S. R., (2005). Beatbox music phone: gesture-based interactive mobile phone using a tri-axis accelerometer, IEEE International Conference on In Industrial Technology, 97-102, China

Cover, T. ve Hart, P., (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, **13**(1), 21–27.

Cristianini, N. ve Shawe-Taylor, J., (2000). *An introduction to support vector machines: and other kernel-based learning methods*. Cambridge University Press.

Győrbíró, N., Fábián, Á. ve Hományi, G., (2009). An activity recognition system for mobile phones, *Mobile Networks and Applications*, **14**, 1, 82-91.

Hall, M. A. ve Smith, L. A., (1999). Feature Selection for Machine Learning: Comparing a Correlation-Based Filter Approach to the Wrapper, Twelfth International Florida Artificial Intelligence Research Society Conference, 235-239, Florida.

Haykin, S. S., (2009). *Neural networks and learning machines*. Prentice Hall/Pearson.

Johnson, D. A. ve Trivedi, M. M., (2011). Driving style recognition using a smartphone as a sensor platform, 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), 1609-1615, Washington DC.

LeMoyne, R., Mastroianni, T., Cozza, M., Coroian, C. Ve Grundfest, W., (2010). Implementation of an iPhone for characterizing Parkinson's disease tremor through a wireless accelerometer application, Annual International Conference on Engineering in Medicine and Biology Society (EMBC), 4954-4958, Buenos Aires.

Li, Q., Stankovic, J. A., Hanson, M. A., Barth, A. T., Lach, J. ve Zhou, G., (2009). Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information, Sixth International Workshop on Wearable and Implantable Body Sensor Networks, 138-143, Washington DC.

Machado, I. P., Gomes, A. L., Gamboa, H., Paixão, V. ve Costa, R. M., (2015). Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization, *Information Processing & Management*, **51**, 2, 204-214.

Mannini, A. ve Sabatini, A. M., (2010). Machine learning methods for classifying human physical activity from on-body accelerometers, *Sensors*, **10**, 2 1154-1175.

McCulloch, W. S. ve Pitts, W., (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, **5**(4), 115–133.

Mursalin, M., Zhang, Y., Chen, Y. ve Chawla, N. V., (2017). Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier, *Neurocomputing*, **241**, 204-214.

Su, X., Tong, H. ve Ji, P., (2014). Activity recognition with smartphone sensors, *Tsinghua Science and Technology*, **19**, 3, 235-249.

Zhu, J., Wu, P., Wang, X. ve Zhang, J., (2013). Sensec: Mobile security through passive sensing, International Conference on Computing, Networking and Communications (ICNC), 1128-1133, San Diego CA.

Classification of writing behaviors using mobile device sensors

Extended abstract

New opportunities and research areas have emerged for modeling mobile user behaviors and activities with the development of sensor technologies and their integration with mobile devices. Mobile device sensors have become rather attractive particularly for applications such as the creation of systems that automatically classify physical activities performed by people such as standing, sitting, running, walking, climbing upstairs, monitoring of patients in health care, and the development of advanced user interfaces.

Different from other studies, in this study, we focused on writing activity. 20 different sentences were randomly presented to 30 different users and the users were asked to write the sentences in the same mobile device via designed mobile interface; the data received from three-axis accelerometer and gyroscope sensors were stored during the writing process. After preprocessing the gathered data, optimal feature vectors were obtained and classifications were done by using these features. The most employed algorithms Artificial Neural Networks (ANN), k-nearest neighbor (k-NN) and Support Vector Machine (SVM) were used for classification. In the user studies that we conducted, ANN, k-NN and SVM achieved the success rate of %97.91, %100 and %97.91 respectively. The best results were obtained when k-NN was employed with the number of neighbors was equal to 1.

Our work provides a great deal of contribution to the literature because it only works on the sensor data received from the mobile device and also achieves a very high classification performance with the created features. In addition to that, it has been revealed that every person has his or her own writing behavior. The writing behaviors of mobile users exhibit biometric properties and can be used as a security mechanism that will distinguish users from each other and take necessary security precautions when personal devices fall into the wrong hands.

ANN is a popular learning algorithm used in classification problems. Basically, ANN consists of input, hidden and output layers. In input layer, the number of neurons are equal to the number of problem inputs. In output layer, the number of neurons is equal to the number of problem outputs. k-NN is a supervised classification method that looks at the nearest neighbors in a set of classified instances and determine a new unclassified instance in order to identify the class which it belongs. k-NN classifies a sample (in our case sensor readings) by computing the distance from the sample to each member of the set by taking the K nearest neighbors and assigning the sample to the class most represented in the K nearest neighbors. Support Vector Machines relies on the idea of decision planes which define the boundaries of a decision. The decision plane separates a set of samples by assigning different class labels. In our case, it is the writing behavior that either belongs to the device owner or someone else.

In our study, we compared the performance of ANN, k-NN and SVM methods. We found that the k-NN method yielded the best performance in classifying the writing behaviors of users by using the data gathered from mobile device sensors. For the ANN classifier with three different hidden neuron structure, following performances were obtained: classification rate 95.83% for two hidden neurons, classification rate 97.91% for 4 hidden neurons, classification rate 97.91 % for six hidden neurons. It was seen that incrementing the number of hidden neurons did not affect the performance over time. For the k-NN classifier with three different neighbor value, the performances were obtained as follows; classification rate 100% for one neighbor, classification rate 98.95% for three and five neighbors. For the SVM classifier, the performance with three different kernels were obtained as follows; classification rate 76.04% for polynomial kernel, classification rate 69.79% radial basis kernel and classification rate 97.91% for Pearson VII function based kernel.

Our results showed that writing behaviors of users can be classified with high success rate. This suggests that this kind of approach can be used to identify the owner of a mobile device and help the user to take the necessary security precautions.

Keywords: *Physical activity recognition, classification, mobile device sensors, writing behavior*