

Sosyal Medya Verileri için Etkili Bir Sınıflandırma Yaklaşımı

Fatih Ertam

Fırat Üniversitesi Enformatik Bölümü, 23100 Elazığ/Türkiye
fatih.ertam@firat.edu.tr

(Geliş/Received: 08.02.2017; Kabul/Accepted: 10.05.2017)

Özet

Sosyal medyanın kullanım oranlarının hızlı bir şekilde artması, sosyal medya madenciliği kavramının gelişmesine sebep olmuştur. Günümüzde sosyal medya araçlarından birisi olan Facebook, kullanıcı sayısı ve günlük aktivite açısından diğer sosyal medya araçlarına göre lider konumdadır. Bu çalışmada Facebook verilerinden oluşturulmuş bir veri seti üzerinde sınıflandırıcı modelleri uygulanarak etkili bir sınıflandırıcı modeli seçilmeye çalışılmıştır. Sınıflandırıcı modelleri olarak bayes ağ, çok katmanlı algılayıcı ağlar, ardışık minimal optimizasyon ve uç öğrenme makinesi kullanılmıştır. Uç öğrenme makinesi içerisinde kullanılan gizli katman hücre sayısının seçimi için bir ile yüz arasında ki tüm değerler alınarak birden fazla aktivasyon fonksiyonu ile çalışma yapılmıştır. Sosyal medya madenciliği yapılabilmesi için etkili ve hızlı bir sınıflandırıcı yöntemi önerilmeye çalışılmıştır.

Anahtar kelimeler: Sosyal Medya Madenciliği, Makine Öğrenmesi, Uç Öğrenme Makineleri, Sınıflandırma Modelleri

An Effective Classification Approach for Social Media Data

Abstract

The rapid increase in social media usage rates has led to the development of the social media-mining concept. Today, one of the social media tools, Facebook is a leader in terms of number of users and daily activity compared to other social media tools. In this study, an effective classifier model was tried to be chosen by applying classifier models on a data set created from Facebook data. BayesNET, multilayer perceptron, sequential minimal optimization and extreme learning machines are used as classifier models. For the selection of the number of hidden layer cells used in the extreme learning machine, all values between one and one hundred were taken and worked with more than one activation function. Efforts have been made to provide an effective and rapid classification method for social media mining.

Keywords: Social Media Mining, Machine Learning, Extreme Learning Machines, Classification Models

1. Giriş

Sosyal Medya kavramı günümüzde birçok işletme yöneticisinin gündeminin en başında gelmektedir [1]. İnternet üzerindeki kullanıcı sayısı her geçen gün hızlı bir şekilde artış göstermektedir. Araştırma sitelerinden Statista'ya göre 2020 yılında internet kullanıcı sayısı 2.95 milyar seviyesine ulaşacaktır [2]. Sosyal medyanın kullanım oranlarının artması internet kullanımını da doğrudan artırmaktadır. Özellikle popüler sosyal medya araçlarından Facebook, kullanım oranı olarak uzun zamandır liderliğini korumaktadır. Aynı istatistik sitesinden alınan veriye göre 2016 yılının son çeyreğinde Facebook 1.86 milyar aylık aktif kullanıcı sayısına

ulaşmıştır [3]. Sosyal medyanın yüksek kullanım oranları özellikle reklam alanında çalışan firmalar için oldukça cezbedici bir hal almıştır. Hızlı gelişimi göz önüne alındığında, sosyal medya markaların yakın gelecekte müşterilerine ulaşması için en önemli medya kanalı haline gelebileceği görülmektedir [4], [5]. Veri madenciliği teknikleri kullanılarak tahmini bilgilerin işlenmemiş verilerden çıkarılması sağlanabilir [6]. Sosyal medya madenciliği yeni bir araştırma başlığının ortaya çıkmasını sağlamıştır. Sosyal medya madenciliği, sosyal medya madenciliğinin temellerini ve potansiyellerini anlamak için tutarlı bir platform

sağlayabilmek adına sosyal medya, sosyal ağ analizi ve veri madenciliğini bütünleştirir. Sosyal medya verilerinden kaynaklanan benzersiz sorunları tanıtır ve temel kavramları ortaya çıkan sorunları ve ağ analizi ile veri madenciliği için etkili algoritmalar sunar [7].

Yapılan çalışma ile Facebook verilerinin olduğu bir veri seti üzerinde veri madenciliği yaklaşımları kullanılarak denetimli öğrenme yaklaşımı gerçekleştirilmiştir. Çalışma için UCI veri tabanında bulunan veri seti kullanılmıştır [8]. Moro ve ark. yapmış olduğu çalışma için elde ettikleri veri seti 19 öznelikten oluşmaktadır [9]. Moro ve ark. bir model geliştirilmeye çalışılarak modelin tahmini değerinin ölçülmesine, model tarafından öngörülen değer ile gerçek değerinin arasındaki farkın değerlendirilmesine çalışılmıştır. Aynı veri seti kullanılarak yapılan çalışmada ise tüm öznelik değerleri kullanılarak daha önce sıklıkla kullanılan bayes ağ, çok katmanlı algılayıcı ağlar, ardışık minimal optimizasyon gibi sınıflandırıcılar ile doğruluk oranı test edilmiş hem de daha yeni bir sınıflandırıcı olarak kabul edebileceğimiz uç öğrenme makinelerinin farklı aktivasyon fonksiyonları ile modelleme yapılmıştır.

Çalışmanın ikinci bölümünde elde edilen veri setinin özelliklerinden bahsedilmiştir. Aynı zamanda kullanılan veri madenciliği modellerinden söz edilerek karşılaştırma için kullanılan metriklerin neler olduğu belirtilmiş ve matematiksel işlevlerine yer verilmiştir. Üçüncü bölümde yapılan deneysel çalışmaların sonuçları birbirleriyle kıyaslanarak hem grafiksel hem de tablo halinde sunulmuştur. Son bölümde elde edilen sonuçlar yorumlanmış ve ileriki çalışmalarda nasıl ilerlenebileceğine dair bir yol haritası çizilmeye çalışılmıştır.

2. Materyal ve Metot

2.1 Veri seti

Bu çalışmada kullanılan veri seti UCI veri tabanından alınmıştır [8]. Bu veri setinde toplam 500 kayıt bulunmaktadır. Veri seti 19 Adet özneliğe sahiptir. Veri setinde bulunan tip özneliği sınıf olarak seçilmiştir. Facebook' da

kullanılan durum, fotoğraf, bağlantı ve video olmak üzere 4 sınıf bulunmaktadır.

Kullanılan özneliklere ait açıklamalar Tablo-1 de verilmiştir.

Tablo 1. Kullanılan öznelikler ve açıklamaları

No	Açıklaması
1	Toplam beğenilen sayfa sayısı
2	Kategori
3	Gönderinin yayınlandığı ay (1-12)
4	Gönderinin yayınlandığı gün (1-7)
5	Gönderinin yayınlandığı saat (1-24)
6	Reklam için ödeme yapılıp yapılmadığı (1,0)
7	Gönderiyi gören tekil kişi sayısı
8	Gönderiye tıklanıp görüntülenme sayısı
9	Gönderide her hangi bir yeri tıklayan tekil kişi sayısı
10	Gönderide her hangi bir yeri tıklayan kişi sayısı
11	Gönderideki her hangi bir yere tıklanma sayısı
12	Sayfayı beğenen kişilerin toplam sayısı
13	Sayfayı beğenen kişilerin sonradan erişim sayısı
14	Sayfayı beğenen ve gönderiye katılan kişi sayısı
15	Gönderideki yorum sayısı
16	Gönderideki beğeni sayısı
17	Gönderinin paylaşım sayısı
18	Beğeni, yorum ve paylaşımların toplamı
19	Tip (Durum, Fotoğraf, Link, Video)

2.2 Kullanılan sınıflandırıcılar

Çalışma için sınıflandırıcı model olarak kullanılan bayes ağ (BA), ardışık minimal optimizasyon (AMO), çok katmanlı algılayıcı ağ (ÇKAĞ) ve uç öğrenme makineleri (UÖM) kullanılarak performansları test verisi için doğruluk oranlarına göre karşılaştırılmıştır.

2.2.1 Bayes ağ (BayesNET)

BayesNET veya BayesAĞ olarak adlandırabileceğimiz sınıflandırıcı Bayes teoremine göre çalışmaktadır. Bayes teoremi iki farklı olaya bağlı olarak olasılık hesabına dayanan bir teorem olarak karşımıza çıkmaktadır. Birbirinden bağımsız X ve Y olayları için iki olayın aynı anda oluşma olasılığı iki olayın ayrı ayrı oluşma olasılıkları çarpımına eşittir. Bir birine bağlı olaylar için bu olasılık Denklem-1 de verildiği şekilde elde edilmektedir:

$$P(X \wedge Y) = P(X) * P(X|Y) \quad (1)$$

Denklem-1 X ve Y durumlarının birlikte olma ihtimalini vermektedir. Denkleme göre $P(X|Y)$, Y olayına ait durumun bilinmesi halinde X olayının meydana gelme ihtimalini verir. Bu durumda Bayes teoremi Denklem-2 de verildiği gibi gösterilebilir.

$$P(X \wedge Y) = P(X) * P(X|Y) = P(Y) * P(X|Y) \quad (2)$$

$$P(X|Y) = \frac{P(X) * P(Y|X)}{P(Y)}$$

BA algoritması ile zor problemler verimli ve güvenilir bir şekilde çözülebilmesi amacıyla var olan yöntemlerin geliştirilmesi ile ortaya çıkmıştır. BA algoritmasının denklemi aşağıda verildiği şekilde kullanılmaktadır [10].

$$p(X) = \prod_{i=0}^{n-1} p(X_i | \pi_i) \quad (3)$$

2.2.2 Ardışık minimal optimizasyon

AMO, destek vektör makinesi sınıflandırıcısının eğitilmesi esnasında meydana gelen ikinci dereceden programlama probleminin çözümü için kullanılan bir yöntemdir [11]. Destek vektör makinelerinde oluşan ikinci dereceden programlama problemi denklem-4 de gösterildiği şekilde ifade edilmektedir.

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j K(x_i, x_j) \alpha_i \alpha_j \quad (4)$$

Denklemden i değerinin 1 ile n ' e kadar tüm değerleri için α değerinin $0 \leq \alpha_i \leq C$ aralığında olması gerekir. Aynı zamanda α_i ve y_i değerlerinin çarpımlarının toplamının sıfır olması gerekir. Bu algoritma ile destek vektör makinelerinin eğitilmesinde meydana gelen ikinci dereceden programlama problemini iteratif olarak daha küçük alt problemlere bölünmeye çalışılır. Elde edilecek olan küçük alt problemler lagrange çarpanı barındırmaktadır. Denklemden α_1 değerinin sıfırdan büyük eşit olması durumunda ve α_2 değerinin C değerinden küçük eşit olması

durumunda denklem-5 de verildiği gibi olmaktadır.

$$y_1 \alpha_1 + y_2 \alpha_2 = k \quad (5)$$

Bu durumdan sonra tek boyutlu ikinci dereceden fonksiyonun minimum değeri bu aşamadan sonra bulunmalıdır. Bu amaçla önce problem için uygun bir α_1 ve α_2 çarpanı seçilmekte ve (α_1, α_2) çiftinin optimize edilmektedir. Bu durum istenilen yakınsama değeri bulunulana kadar devam ettirilir.

2.2.3 Çok katmanlı algılayıcı ağlar

ÇKAĞ, girdi katmanı ara katman ve çıktı katmanından oluşan doğrusal olmayan aynı zamanda geriye yayılım modeli olarak da bilinen bir yapay sinir ağı tekniğidir [12]. Girdi katmanından verilen veri herhangi bir işleme tabi tutulmadan bir sonraki katmana iletilir. Bir katmandaki bütün iş akışı bir üst katmandaki işleyişe bağlı olduğu için veri akışı sürekli ileriyeye doğru yapılmaktadır. ÇKAĞ' de amaç ağın istenilen çıktısı ile üretilen çıktısı arasındaki hatayı olabildiğince düşürmektir. Ağın eğitimi sırasında girdi ve girdilere karşılık üretilen çıktı değerleri verilir. Ağın görevi her giriş için o girişe karşılık gelecek olan çıkışı üretmektir. Kullanılan eğitim yöntemine göre ağın çıkışı ile elde edilmek istenilen çıkış arasındaki hata tekrar tekrar geriye doğru yayılarak hata en düşük değerini alıncaya kadar ağın ağırlıkları güncellenir. Gizli katmanda bulunan her j nöronu w_{ji} bağlantı ağırlığının giriş değerlerinin çarpımının toplamıdır. Çıkış olan y_j bu toplamın bir fonksiyonudur. Denklem-6 da bu durum matematiksel olarak gösterilmiştir.

$$y_i = f\left(\sum w_{ji} x_i\right) \quad (6)$$

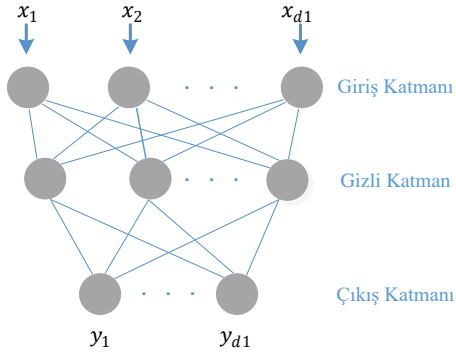
Denklem-6 da verilen f bir aktivasyon fonksiyonudur ve nörona etki eden işaretlerin ağırlıklı toplamını için bir çıkış değeri oluşturur. Çıkış için istenilen değer ile elde edilen değer arasındaki farkların kareleri toplamı olan e Denklem-7 de verildiği gibi ifade edilebilir.

$$e = \frac{1}{2} \sum_j (y_j^* - y_j)^2 \quad (7)$$

Denklem-7 de verilen y_j^* , j . çıkış nöronunun üretilen değeri; y_j ise j . çıkış nöronunun istenilen değerini ifade etmektedir.

2.2.4 Uç öğrenme makinesi (extreme learning machine)

Daha önce bahsedilen sınıflandırma algoritmalarının yanında daha yeni ve gelişmekte olan UÖM, Huang ve ark. tarafından 2004 yılında önerilmiştir [13]. Diğer ileri beslemeli yapay sinir ağlarından farklı olarak UÖM hata fonksiyonuna göre ağırlığı sürekli olarak ayarlamak zorunda değildir. Gizli katmanda bulunan nöronların öğrenme parametreleri rasgele bir şekilde oluşturulurken, çıktı ağırlıkları Moore-Penrose genelleştirilmiş ters matrisi hesaplanarak belirlenir. Bu nedenle eğitim hızı ve genelleme doğruluğu yüksektir [14], [15]. Bu nedenlerden dolayı UÖM, sınıflandırma çalışmalarında sıklıkla kullanılmaya başlamıştır. UÖM temel olarak bir giriş, bir gizli katman bir de çıkış katmanı içermektedir. Şekil-1 de UÖM'nin yapısı gösterilmektedir[16].



Şekil 1. Uç öğrenme makinesi yapısı

UÖM ağırlıkları hata fonksiyonuna göre sürekli bir şekilde ayarlama ihtiyacı duymaz. Giriş katmanı ile gizli katman arasındaki öğrenme parametreleri rasgele oluşturularak çıktı ağırlıkları analitik olarak en küçük kareler yöntemi kullanılarak hesaplanır [13]. UÖM'nin

matematiksel çıkarımı şöyle özetlenebilir: $\{(I_n, T_n) | n = 1, 2, \dots, N; I_n \in RW; T_n \in RV\}$ eğitimi verildiğinde I_n giriş vektörü ve T_n her örnek için hedef değer vektörüdür. K gizli düğümleri olduğu düşünülürse UÖM'nin çıktısı Denklem- 8 de verildiği gibi olur.

$$T_n = \sum_{k=1}^K \beta_k G(\omega_k \cdot I_n + b_k)$$

Burada $\omega_k = [\omega_{k1}, \omega_{k2}, \dots, \omega_{kW}]$ $\omega_k = [\omega_{k1}, \omega_{k2}, \dots, \omega_{kW}]^T$ ve $\beta_k = T[\beta_{k1}, \beta_{k2}, \dots, \beta_{kV}]$ $k = 1, 2, \dots, K$ sırasıyla k . gizli katman nöronunun girdi ağırlıklarını ve çıktı ağırlıklarını göstermektedir. b_k ve $G(\omega_k \cdot I_n + b_k)$ k . gizli katman nöronunun eşik değerini ve çıktısını temsil etmektedir. $G(\cdot)$ gizli katmanın aktivasyonudur ve denklem-9 da verildiği gibi gösterilebilir:

$$G(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

Denklem-8 $H\beta = T$ gibi daha kısa ve etkili bir şekilde gösterilebilir. H gizli katmanının rasgele üretilmiş çıktı matrisini temsil eder ve denklem-10 da verildiği gibi gösterilebilir.

$$H = \begin{bmatrix} G(\omega_1 \cdot I_1 + b_1) & \dots & G(\omega_K \cdot I_1 + b_K) \\ \vdots & \ddots & \vdots \\ G(\omega_1 \cdot I_N + b_1) & \dots & G(\omega_K \cdot I_N + b_K) \end{bmatrix}_{N \times K}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_K^T \end{bmatrix}_{K \times V} \quad (10)$$

$$T = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times V}$$

$H\beta = T$ ifadesi en küçük kareler çözümüne göre hesaplanır ve gizli katman ile çıktı katmanı arasındaki ağırlıklar $\hat{\beta} = H^+T$ şeklinde elde edilir. H^+ , H çıkış matrisinin genelleştirilmiş Moore-Penrose matrisi olarak tanımlanmıştır.

UÖM Eğitim algoritmasını şu şekilde özetlemek mümkündür:

- Giriş ağırlıkları ω_k ve gizli katman nöronları b_k rasgele atanır
- Gizli kalman çıkış matrisi \mathbf{H} hesaplanır
- En küçük kareler çözümü kullanılarak çıkış ağırlıkları β hesaplanır.

2.3 Sınıflandırıcı modeli için performans karşılaştırması

Veri setine uygulanacak olan sınıflandırıcı modellerin performans karşılaştırılmasının yapılabilmesi için hata matrisinin oluşturulması gerekmektedir. Tablo-2 de veri setinden alınan gerçek değerler ile tahmin edilen değerlerin arasındaki ilişki gösterilmektedir.

Tablo 2. Performans değerlendirmesi için gerçek değerler ile tahmini değerlerin karşılaştırılması

Etiket	Gerçek Pozitif	Gerçek Negatif	Gerçek Toplam
Tahmini Pozitif	Doğru Pozitif (D_p)	Yanlış Pozitif (Y_p)	T_p
Tahmini Negatif	Yanlış Negatif (Y_n)	Doğru Negatif (D_n)	T_n
Tahmini Toplam	(p)	(n)	T

Tablo 2'deki değerlere göre doğru pozitif; gerçekte hem de tahminde pozitif olarak işaretlenen değerleri, yanlış pozitif; gerçekte negatif olmasına rağmen pozitif işaretlenen değerleri, yanlış negatif; gerçekte pozitif olmasına rağmen negatif olduğu tahmin edilen değerleri, doğru negatif ise hem gerçekte hem de tahminde negatif işaretlenen değerleri ifade etmektedir. Sınıflandırmanın doğruluğu Denklem 11'de gösterildiği şekilde hesaplanabilir.

$$\text{Doğruluk Değeri} = \frac{D_p + D_n}{T} \quad (11)$$

3. Deneysel Çalışmalar

Deneysel çalışma için veri setinin %60'ı eğitim için ayrılmış geri kalanı ise test için kullanılmıştır. Sınıflandırıcılara ait performansların karşılaştırılabilmesi amacıyla Weka 3.8.1 ve Matlab 2015a yazılımlarından faydalanılmıştır. Çalışma Windows 2012 Server

işletim sistemine sahip, 16 GB RAM ve 2.93 GHz intel x5570 işlemcili bilgisayar üzerinde hazırlanmıştır. BA algoritması ile yapılan çalışmada test verilerinden 149 tanesi doğru 51 tanesi yanlış olarak sınıflandırılmıştır. Ortalama doğruluk oranı % 74.5 olarak bulunmuştur. ÇKAĞ algoritması ile yapılan çalışmada 172 tane test verisi doğru, 28 tane test verisi yanlış sınıflandırılarak % 86 doğruluk değerine ulaşılmıştır. AMO çalışmasında farklı çekirdek fonksiyonları seçilerek sınıflandırıcıya test verileri verilmiştir. Radyal tabanlı çekirdek fonksiyonu kullanılarak yapılan çalışmada doğruluk oranı %85.5, polinom çekirdek fonksiyonu kullanılması ile %89.5 değerine ulaşılmıştır.

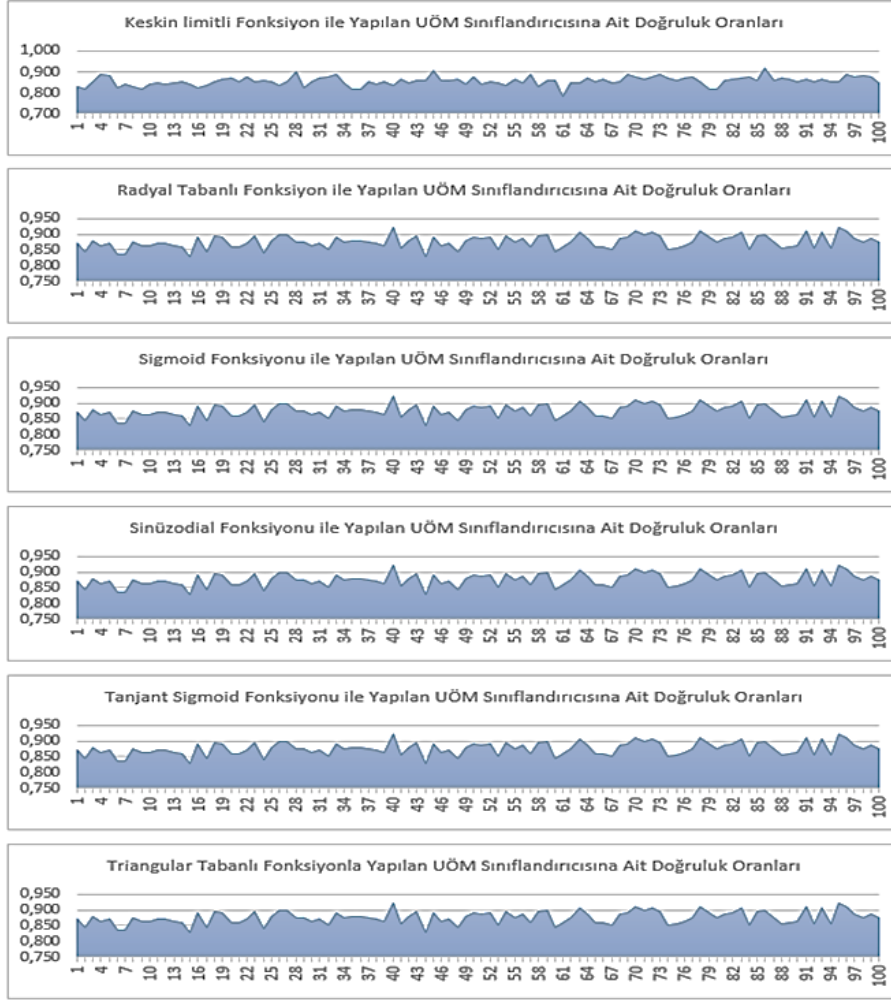
UÖM algoritması ile yapılan çalışmalarda farklı aktivasyon fonksiyonları ile gizli nöron sayıları kullanarak sınıflandırıcının doğruluk değeri belirlenmeye çalışılmıştır. Bu amaçla 1-100 arasında değişen gizli nöron sayıları; keskin limitli, radyal tabanlı, sigmoid, sinüzoidal, tanjant tabanlı ve triangular tabanlı aktivasyon fonksiyonu kullanılan UÖM sınıflandırıcılarına verilmiştir. Keskin limitli aktivasyon fonksiyonun kullanıldığı çalışma için en yüksek doğruluk oranı 86 gizli nöronun kullanıldığı çalışmada % 91.5, en kötü değer ise 61 gizli nöronun kullanıldığı çalışmada %78 olarak bulunmuştur. Radyal tabanlı aktivasyon fonksiyonun kullanıldığı çalışma için en yüksek doğruluk oranı 54 gizli nöronun kullanıldığı çalışmada % 93.5, en kötü değer ise 29 gizli nöronun kullanıldığı çalışmada %82 olarak bulunmuştur. Sigmoid aktivasyon fonksiyonun kullanıldığı çalışma için en yüksek doğruluk oranı 65 gizli nöronun kullanıldığı çalışmada % 94.5, en kötü değer ise 1 gizli nöronun kullanıldığı çalışmada %82 olarak bulunmuştur. Sinüzoidal aktivasyon fonksiyonun kullanıldığı çalışma için en yüksek doğruluk oranı 27 ve 33 gizli nöronun kullanıldığı çalışmada % 92.5, en kötü değer ise 1 gizli nöronun kullanıldığı çalışmada %77.5 olarak bulunmuştur.

Tanjant Sigmoid aktivasyon fonksiyonun kullanıldığı çalışma için en yüksek doğruluk oranı 26 ve 51 gizli nöronun kullanıldığı çalışmada % 94, en kötü değer ise 2 gizli nöronun kullanıldığı çalışmada %73.5 olarak bulunmuştur. Triangular tabanlı aktivasyon fonksiyonun kullanıldığı

çalışma için en yüksek doğruluk oranı 40 ve 95 gizli nöronun kullanıldığı çalışmada % 92, en kötü değer ise 15 ve 44 gizli nöronun kullanıldığı çalışmada %83 olarak bulunmuştur.

Şekil-2 de farklı aktivasyon fonksiyonlarının 1-100 arasında gizli nöron sayısının değiştirilmesi ile elde edilen doğruluk değerleri gösterilmiştir.

BA, ÇKAĞ, AMO ve UÖM ile yapılan çalışmalarda elde edilen doğruluk değerleri karşılaştırılmalı bir şekilde Tablo 3 de verilmiştir.



Şekil 2. Farklı aktivasyon fonksiyonları ve gizli katman hücre sayıları ile UÖM doğruluk değerleri

Tablo 3. Kullanılan sınıflandırıcıların karşılaştırılması

Metot	Aktivasyon Fonksiyonu	Nöron Sayısı	Doğruluk Yüzdesi
BA	-	-	74.5
ÇKAĞ	-	-	86
AMO	Radyal tabanlı	-	85.5
AMO	Polinom	-	89.5
UÖM	Keskin limitli	86	91.5
UÖM	Radyal tabanlı	54	93.5
UÖM	Sigmoid	65	94.5

4. Sonuçlar ve Tartışma

Sosyal medyanın kullanım oranlarının hızlı bir şekilde artması, sosyal medya madenciliği kavramının gelişmesine sebep olmuştur. Günümüzde sosyal medya araçlarından Facebook

kullanıcı sayısı ve günlük aktivite açısından diğer sosyal medya araçlarına göre lider konumdadır.

Bu çalışmada Facebook verilerinden oluşturulmuş ve herkesin ulaşabileceği bir veri seti üzerinde sınıflandırıcı modeller uygulanarak etkili bir sınıflandırıcı modeli seçilmeye çalışılmıştır. Bu amaçla BA, ÇKAĞ, AMO ve UÖM algoritmalarından faydalanılmıştır. Kullanılan algoritmalarından aktivasyon fonksiyonu olanlar için farklı aktivasyon fonksiyonları ile ayrı ayrı çalışma yapılmıştır.

Yapılan çalışmalarda BA ile elde edilen doğruluk oranının diğer algoritmalara göre daha düşük olduğu gözlenmiştir. ÇKAĞ ve AMO algoritmalarında elde edilen sonuçların birbirlerine yakın, UÖM ile elde edilen sonuçların ise daha iyi derecede sınıflandırma yaptığı gözlenmiştir. UÖM için farklı aktivasyon fonksiyonlarının 1 ile 100 arasında değişen değerlerde farklı sayıda gizli katman hücrelerine verdiği tepkiler için ayrı bir yazılım hazırlanarak sonuç alınmıştır. En iyi doğruluk oranının %94.5 ile sigmoid aktivasyon fonksiyonunun kullanıldığı ve gizli nöron sayısının 65 seçildiği uç öğrenme makinesi ile bulunduğu görülmüştür.

Sosyal medya verilerinin incelenerek analiz edilmesi ile özellikle bu alanda reklam verebilecek olan firmalar için bu verilerin önemli olduğu hangi sınıfın daha çok tercih edilerek reklam için bu alanda yoğunlaşabileceğinin önemli bir veri olacaktır. İlerleyen çalışmalarda gizli nöron hücre sayısının tespit edilmesinde daha hızlı modelleme tekniklerinden faydalanılarak daha fazla değerler daha hızlı bir şekilde seçilebilmesi hedeflenmektedir.

5. Kaynaklar

1. Kaplan, A. M. ve Haenlein, M. (2010). “Users of the world, unite! The challenges and opportunities of Social Media”, *Bus. Horiz.*, **53**, sayı 1, ss. 59–68.
2. Statista, (2016). “Number of Worldwide Social Network Users 2010-2019”, *Statista*, [Çevrimiçi].<http://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>.

3. Statista (2016). “Facebook users worldwide 2016”, *statista.com*, [Çevrimiçi].
<http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>.
4. Korschun D. ve Du, S. (2013). “How virtual corporate social responsibility dialogs generate value: A framework and propositions”, *J. Bus. Res.*, **66**, sayı 9, ss. 1494–1504.
5. Mangold W. G. ve Faulds D. J. (2009). “Social media: The new hybrid element of the promotion mix”, *Bus. Horiz.*, **52**, sayı 4, ss. 357–365.
6. Turban E., Sharda R. ve Delen D. (2011). *Decision Support and Business Intelligence Systems*, **8th**.
7. Zafarani, R., Abbasi, M. A. ve Liu H. (2014). *Social Media Mining. Cambridge: Cambridge University Press*.
8. U. M. L. Repository (2016). “Facebook metrics Data Set”, www.ics.uci.edu.
9. Moro, S., Rita, P. ve Vala, B. (2016). “Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach”, *J. Bus. Res.*, **69**, sayı 9, ss. 3341–3351.
10. Friedman, N., Geiger, D. ve Goldszmidt, M. (1997). “Bayesian Network Classifiers”, *Mach. Learn.*, **29**, ss. 131–163.
11. Platt, J. C. (1998). “Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines”, *Adv. kernel methods*, ss. 185–208.
12. Rumelhart, D. E., Hinton, G. E. ve Williams, R. J. (1986). “Learning representations by back-propagating errors”, *Nature*, **323**, sayı 6088, ss. 533–536.
13. Huang, G., Zhu, Q. ve Siew, C. (2004). “Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks”, *IEEE Int. Jt. Conf. Neural Networks*, **2**, ss. 985–990.
14. Huang, G.-B., Zhu, Q. ve Siew C. (2006). “Extreme learning machine: Theory and applications”, *Neurocomputing*, **70**, sayı 1–3, ss. 489–501.
15. Huang, G. Bin., Wang, D. H. ve Lan, Y. (2011). “Extreme learning machines: A survey”, *Int. J. Mach. Learn. Cybern.*, **2**, sayı 2, ss. 107–122.
16. Ertam F. ve Avcı, E. (2017). “A new approach for internet traffic classification: GA-WK-ELM”, *Measurement*, **95**, ss. 135–142.