

# SİSMİK DARBELERİN SINIFLANDIRILARAK DEPREM TEHLİKESİNİN TAHMİN EDİLMESİ

Mehmet BİLEN<sup>1</sup>, Ali Hakan IŞIK<sup>2</sup>, Tuncay YİĞİT<sup>3</sup>

<sup>1</sup>Mehmet Akif Ersoy Üniversitesi, Çavdır Meslek Yüksek Okulu, Burdur

<sup>2</sup>Mehmet Akif Ersoy Üniversitesi, Mühendislik Mimarlık Fakültesi, Bilgisayar Mühendisliği, Burdur

<sup>3</sup>Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği, Isparta

## ÖZET

Deprem tahmin edilmesi en zor olan doğal afetlerden biridir. Bu yüzden deprem tahmini alanında yapılan çalışmalar olası can ve mal kayıplarının önüne geçilebilmesi için büyük bir öneme sahiptir. Bu çalışmada Polonya maden ocaklarından elde edilen sismik darbe verileri KNN (K-nearest Neighbors: K En Yakın Komşular), SVM (Support Vector Machine: Destek Vektör Makinesi), YSA (Yapay Sinir Ağları) algoritmaları kullanılarak sınıflandırılmıştır. Sınıflandırma hızının ve doğruluğunun artırılması için PCA (Principal Component Analysis: Temel Bileşen Analizi) gibi öznelik seçme algoritmalarından yararlanılarak verinin boyutu azaltılmıştır. Sınıflandırma sonucunda, deprem tehlikesi %94,12'ye kadar çıkan bir başarı oranı ile doğru tahmin edilmiştir ve bulgular paylaşılmıştır.

**Anahtar Kelimeler:** Deprem tahmini, KNN, YSA, SVM, PCA

## SEISMIC HAZARD PREDICTION WITH CLASSIFICATION OF SEISMIC PULSES

### ABSTRACT

Earthquake is one of the most difficult disasters to predict. Therefore, studies in this area are very important to prevent possibility of loss of life and property. In this study, seismic pulse data taken from coal mines of Poland are classified with using KNN (K-nearest Neighbors), SVM (Support Vector Machine) and ANN (Artificial Neural Networks) algorithms. Dimension of data is reduced to increase the speed and accuracy of classification with help of feature selection algorithms such as PCA (Principal Component Analysis). According to the classification results, seismic hazard is predicted correctly with up to 94.12% successful rates and results are shared.

**Keywords:** Earthquake prediction, KNN, ANN, SVM, PCA

## 1. GİRİŞ

Deprem yer derinliklerinde biriken enerjinin bir anda boşalması sonucu oluşan elastik dalgaların, zemin yüzeyinde meydana getirdiği titreşim hareketleridir (Ketin, 2005). Bir başka tanımla depremler karmaşık olaylar sonucu gelişen doğal afetlerdir (Karaman, 2006). Bu karmaşıklık depremin zamanı, yeri ve şiddeti gibi birçok parametrenin tahmin edilmesini zorlaştırmaktadır. Birinci derecede deprem bölgesinde olan ülkemizde meydana gelmesi olası bir depremde can ve mal kayıplarını en aza indirmek için bu tahminler hayati önem taşımaktadır. Olası bir depremin önceden tahmin edilebilmesi için birçok öğeden yararlanılmaktadır. Bunlar kabuktaki yamulmalar, sismik darbeler kayaçlardaki jeokimyasal değişimler, yer altı sularındaki değişimler, su kaynaklarının azalması veya önce artıp daha sonra azalması, manyetizma, elektrik, bozulma değişimleri, bölgedeki fiziksel, kimyasal ve biyolojik olaylarda meydana gelen değişimler, canlılarda meydana gelen davranış değişiklikleridir (Göker, 2010). Sismik darbeler kullanarak deprem tehlikesinin tahmin edilmesi üzerine literatürde birçok çalışma yapılmıştır (Kabiesz et al., 2013; Zhou ve Zhu, 2014; Çelik vd., 2014) Bu çalışmada

International Burdur Earthquake & Environment Symposium (IBEES2015)

Uluslararası Burdur Deprem ve Çevre Sempozyumu

7-9 May 2015, Mehmet Akif Ersoy University, Burdur-Türkiye

<http://ees2015.mehmetakif.edu.tr> – <http://ees2015.maku.edu.tr>

ise sismik darbeler kullanılarak KNN, SVM ve YSA gibi öğrenilebilir algoritmaların sayesinde literatürdeki çalışmalardan daha iyi bir deprem tehlikesi tahmini elde edilmeye çalışılmıştır.

Bu çalışma 4 başlık şeklinde hazırlanmıştır. İlk başlıkta deprem tahmininin önemi ve zorluğundan bahsedilerek sismik darbeler ile yapılan çalışmalar verilmiştir. İkinci kısımda bu çalışmada kullanılan veriler ve algoritmalarından bahsedilmiştir. Üçüncü kısımda yapılan çalışmalar sonucunda elde edilen bulgular değerlendirilerek tartışılmıştır. Son olarak sonuçlar verilerek gelecekte yapılması düşünülen çalışmalardan bahsedilmiştir.

## 2. MATERYAL VE YÖNTEM

Hazırlanan çalışmada seismic-bumps adı verilen  $10^4$  J'den daha büyük sismik darbelerin tahmin edilmesi için kullanılan veri kümesi tercih edilmiştir (Sikora ve Wrobel, 2010). Polonya madenlerinden 8 saatlik vardiya aralıkları ile paralel alınan bu verilerde 2584 adet örnek 1'i sınıf parametresi olmak üzere 19 tanede öznitelik bulunmaktadır. Veri kümesi içerisinde 0 ve 1 olmak üzere iki farklı sınıf bulunmaktadır. 0 bir sonraki zaman aralığında deprem tehlikesinin olmadığı, 1 ise deprem tehlikesinin varlığı anlamına gelmektedir. Veri kümesindeki öznitelikler vardiya sonundaki sismik durum, sismo akustik durum, vardiya türü, jeofonlar üzerinden kaydedilen sismik enerji, önceki vardiyadan jeofonlar üzerinden kaydedilen darbe sayısı, sekiz önceki vardiyada kaydedilen ortalama enerji, sekiz önceki vardiyada kaydedilen ortalama darbe, sismo akustik yöntem ile vardiya sonucu, önceki vardiya için kaydedilen darbe sayısı, önceki vardiya içinde kayıtlı sismik darbe sayısı enerji aralıkları, önceki vardiyadaki toplam sismik enerji, önceki vardiyadaki sismik darbe maksimum enerjisidir (Çelik vd., 2014).

Veri kümesi içerisindeki sınıf parametrelerini tahmin etmek deprem tahmini ile aynı anlama gelmektedir. Bu sınıflandırma işlemi için KNN, SVM, YSA algoritmaları tercih edilmiştir ancak veri boyutunun çok yüksek olması karmaşıklığı ve işlem için harcanan süreyi olumsuz yönde artırmaktadır. Veri boyutunu azaltmak için örnekler arasında seçim yapılmış her sınıftan 170'er adet örnek alınmıştır. Her bir örnekte fazla sayıda bulunan öznitelik sayısını azaltmak için ise PCA kullanılmıştır.

### 2.1 Temel Bileşen Analizi

Temel bileşen analizi (PCA) örnekler içerisindeki bir biri ile bağımlı özniteliklerin değişimlerini korumaya çalışarak, öznitelikleri farklı bir boyuta taşıyarak öznitelik sayısını azaltmaya yarayan bir dönüşüm tekniğidir. (Çilli, 2007). Yapılan analiz sonucunda yeni bir değişken uzayı elde edilmektedir. Ve bu uzayı ilk değişkenlerin temel bileşenleri adı verilir. Bu bileşenler varyans değeri yüksek olandan düşük olana kadar sıralanarak en yüksek değerler seçilir.

Seismic-bump veri kümesi içerisindeki örneklerin 18 adet öznitelik barındırması hazırlanan çalışmada kullanılan algoritmaların hesaplamalarını çok uzun sürelerde tamamlamasına ve karmaşıklığa yol açmaktadır. Veri kümesinin öznitelikleri bulunarak 2 boyuta kadar indirilmiştir. Sonuç olarak daha basit bir veri kümesi, daha hızlı ve doğruluğu yüksek sınıflandırma sonucu elde edilmiştir. Çünkü önemsiz olan öznitelikler temel bileşen analizi sonrası etkisini yitirerek sınıflandırma sonucuna olan etkileri azaltılmıştır.

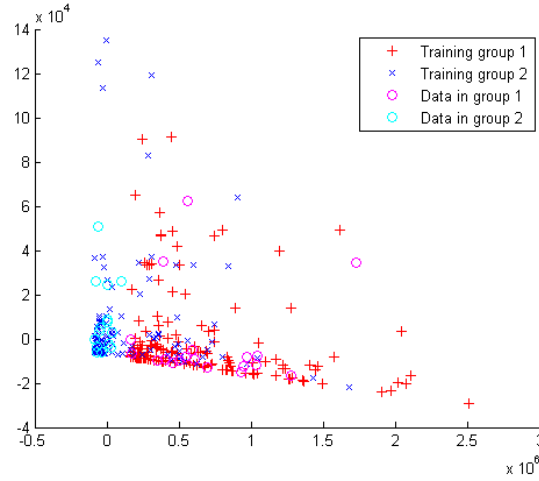
### 2.2 KNN Algoritması

K en yakın komşular algoritması sorgulanan örneğin daha önceki örnekler içerisindeki en yakın K adet vektörün hangi sınıfa ait olduğuna bakarak sınıflandırma yapmaktadır (Kutlu, 2001). KNN kullanımı basit olmasına rağmen bir o kadar da etkili bir algoritmadır. KNN her ne kadar bir eğitici öğrenme kullanan sınıflandırma algoritması olsa da karmaşık ve zaman alan bir öğrenme algoritması kullanmaya gerek kalmadan anlık verilere bakarak sınıflandırma yapabilmektedir. Sınıflandırılacak örneğin sahip olduğu özniteliklerin daha önceki sınıf parametresi bilinen örneklerdeki öznitelikler arasındaki uzaklık farkı en az olan örneklere bakılır. Bu uzaklık farkı hesaplanırken Manhattan, Minkowski, Öklid, Cityblock, Korelasyon, Cosine, Hamming, Jaccard, Mahalanobis, Spearman, Seucclidean uzaklık fonksiyonlarından biri tercih edilebilmektedir. Çalışmada bu uzaklık hesaplamalar içerisinde veri kümesine uygun ve en sık kullanılan Öklid (Euclidian Distance Metric:

Öklid Uzaklık Ölçümü) fonksiyonu seçilmiştir. Öklid uzaklık hesaplama fonksiyonu Eşitlik 1.'de gösterilmiştir (Karimifard et al., 2006).

$$d_{st} = \sqrt{\sum_{j=1}^n (x_{sj} - y_{tj})^2} \quad (1)$$

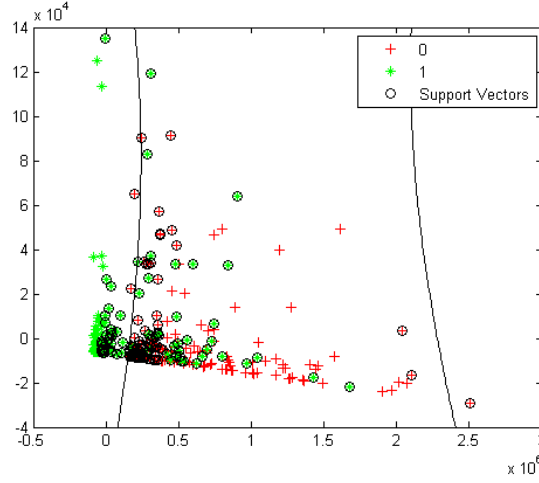
Eşitlik içerisinde kullanılan  $d_{st}$  iki örnek arasındaki uzaklığı ifade etmektedir. n toplam öznitelik sayısını j ise her bir öznitelik değerinin indis numarasını ifade etmek için kullanılmaktadır.  $x_s$  sınıflandırılmaya çalışılan örneği  $y_t$  ise sınıfı bilinen bir örneği ifade etmektedir. Sınıflandırılmak istenen örneğin bütün veri kümesine olan uzaklıkları bulunduktan sonra uzaklıklar yakından uzağa doğru sıralamaya tabi tutulur. Bu noktadan sonra K parametresi büyük bir önem taşımaktadır. Çünkü K adet en yakın örneğe bakılacak ve en çok hangi sınıf var ise sınıflandırılmaya çalışılan örneğin o sınıfa ait olduğuna karar verilecektir. K parametresini belirlemek için bir çok çalışma olsa da her veri kümesi için farklı bir optimum K parametresi olabilmektedir. Bu yüzden yapılan çalışmada K parametresi deneme yanılma yolu ile en yüksek başarı oranını veren 50 olarak kabul edilmiştir. KNN algoritması ile yapılan sınıflandırma işlemi sonucunda eğitim için kullanılan veriler ve sınıflandırma işlemine tabi tutulan verilerin grafiksel gösterimi Şekil 2.1'de verilmiştir.



Şekil 2.1. KNN algoritması ile yapılan sınıflandırma işlemi sonucu elde edilen çıktılar

### 2.3. Destek Vektör Makinesi

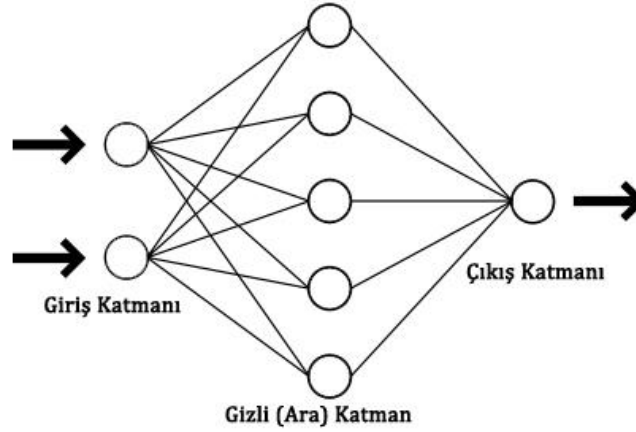
Destek vektör makineleri sınıflandırma problemlerinde oldukça faydalı sonuçlar veren basit algoritmalardan biridir. Sınıflandırma probleminde SVM iki farklı sınıf arasında bir sınır çizgisi çekerek iki farklı gruba ayırmaya çalışır (Şeker, 2008). Şekil 2.2.'de seismic-bump veri seti örneklerinin özniteliklerine uygulanmış SVM algoritması sonrası çizilen sınırlar gösterilmektedir.



Şekil 2.2. Özniteliklerin SVM algoritması ile sınıflandırılması

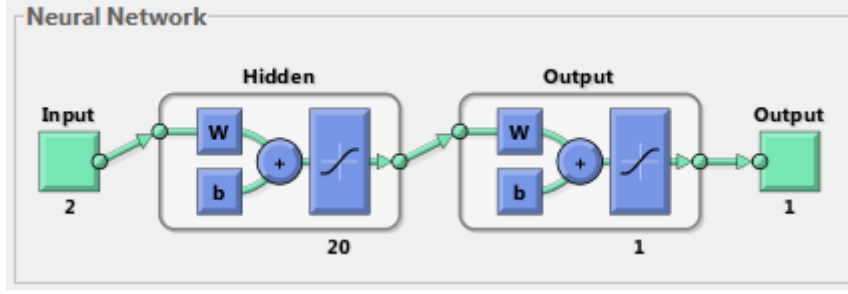
#### 2.4. Yapay Sinir Ağları

Yapay sinir ağları, biyolojik sinir hücrelerini basit bir şekilde taklit eden bilgisayar yazılımlarıdır. Bu özellikleri sayesinde sınıflandırma, tahmin, optimizasyon, örüntü tanıma, bellek yönetimi ve kontrol gibi birçok problemi etkili ve kolay bir şekilde çözebilmektedirler (Blanton, 1997). Yapay sinir ağları, örneklerden yola çıkarak bilgi toplamakta ve daha sonra hiç görmediği örnekler ile karşılaşınca öğrendiği bilgileri kullanarak örnekleri sınıflandırabilmektedir (Keleşoğlu ve Fırat, 2006). Basit bir yapay sinir ağının modeli Şekil 2.3'de gösterilmektedir.



Şekil 2.3. Basit bir YSA modeli

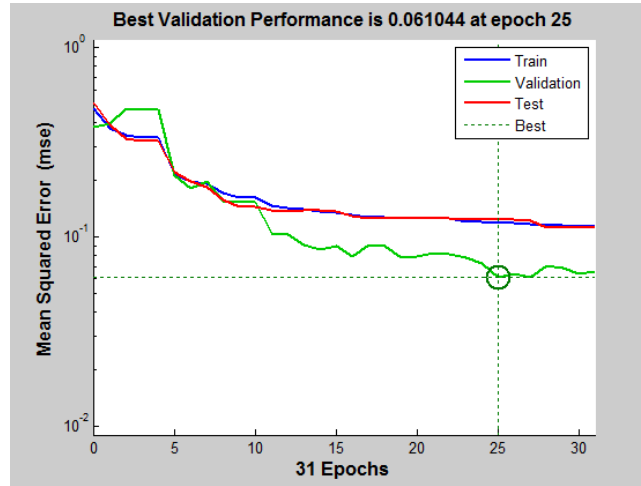
Giriş katmanı verilerin alınarak bir sonraki katmana teslim edildiği yerdir. Gizli katmandaki her bir sinir hücresi kendisinden bir önceki katmandaki sinir hücrelerinden gelen her bir veriyi belirli bir ağırlığa göre işleme tabi tutmaktadır. Bir YSA'da öğrenme bu ağırlıkların değişmesi ile meydana gelmektedir. Çünkü bu ağırlıklar her bir giriş değerinin ne oranda çıkışa etki edeceğine karar vermektedir. Hazırlanan çalışmada ağın eğitimi için kullanılan denklem Eşitlik 2'de, kullanılan ağın modeli ise Şekil 2.4'de verilmiştir.



Şekil 2.4 Kullanılan YSA'nın yapısı

$$\Delta A_{jm}^{AC}(t) = \gamma \delta_m \zeta_j^A + \alpha A_{jm}^{AC}(t-1) \quad (2)$$

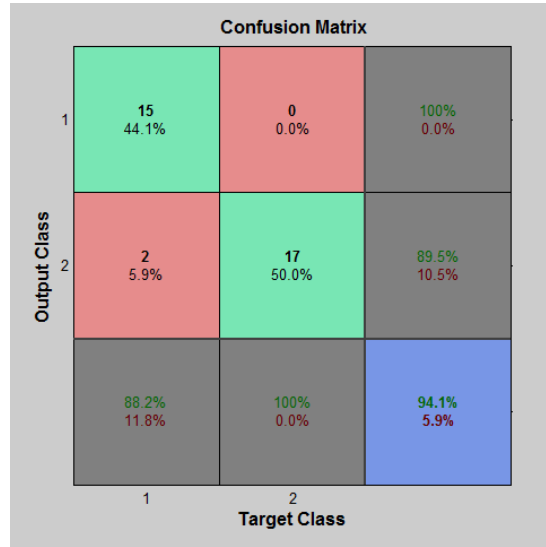
$\Delta A_{jm}^{AC}$  her bir j sinirinden m sinirine giden ağırlıkta meydana gelen değişim miktarını ifade ederken  $A_{jm}^{AC}(t-1)$  bir önceki adımdaki ağırlığı ifade etmek için kullanılmaktadır.  $\gamma$  ağırlık eğitimi için kullanılan parametre  $\alpha$  ise devinirlik değeridir.  $\delta_m$  m ağırlığının yapmış olduğu hata oranını göstermektedir.  $\zeta_j^A$  ise ara katmandaki j sinirinin çıkış değeridir. Eğitim işlemi gerçekleştirildikten sonra elde edilen ağırlıklar ağırlık hafızasını oluşturmaktadır. Bu ağırlıklar kullanılarak daha önce verilmemiş örnekler ağırlık üzerinden işleme sokularak sınıflandırma sonucu elde edilmiştir. Şekil 2.5'de bu ağırlık eğitimi, doğrulama ve test aşamalarının her bir adımı (epok) gösterilmiştir.



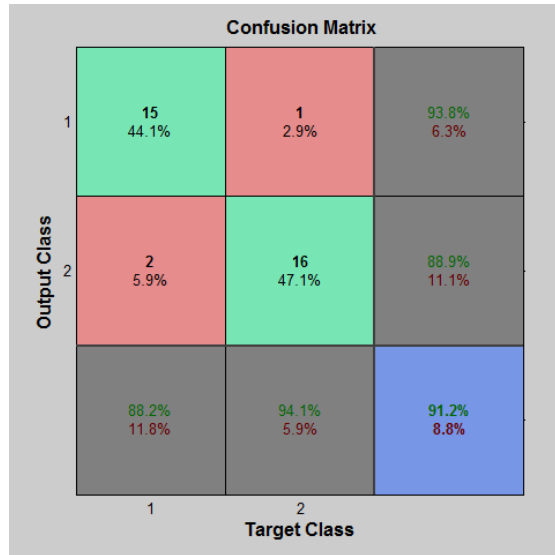
Şekil 2.5 YSA ağırlık eğitimi, doğrulama ve test işlemlerinin grafiksel çıktıları

### 3. ARAŞTIRMA BULGULARI VE TARTIŞMA

Deprem tahmini yapabilmek için kullanılan seismic-dump veri kümesi sınıflandırma algoritmalarının eğitimi için 10 parçaya bölünmüştür. 9 kısım eğitim için 1 kısım ise test için kullanılmıştır. KNN ile yapılan sınıflandırma işlemleri sonucunda %94.1176 oranında doğru sınıflandırma başarısı elde edilmiştir. KNN algoritması sonucunda elde edilen performans değerlerini göstermek için yararlanılan karşılıklı değerler matrisi (confusion matrix) Şekil 3.1'de verilmiştir. Yapılan sınıflandırma işlemleri sonucunda SVM algoritması %91.1765 başarılı sınıflandırma oranına ulaşabilirken, YSA %82.6 başarılı sınıflandırma oranında kalmıştır. SVM ve YSA algoritmaları ile yapılan sınıflandırma işlemleri sonucunda elde edilen performans değerleri için karşılıklı değerler matrisi ise Şekil 3.2 ve Şekil 3.3.'de gösterilmiştir.



Şekil 3.1. KNN sınıflandırma sonucu karşılıklı değerler



Şekil 3.2. SVM sınıflandırma sonucu karşılıklı değerler

Output Class	1	18 39.1%	1 2.2%	94.7% 5.3%
	2	7 15.2%	20 43.5%	74.1% 25.9%
		72.0% 28.0%	95.2% 4.8%	82.6% 17.4%
		1	2	
		Target Class		

Şekil 3.3. YSA sınıflandırma sonucu karşılıklı değerler

#### 4. SONUÇ

Deprem oluşumu Polonya kömür madenlerinden alınan sismik veriler kullanılarak tahmin edilmeye çalışılmıştır ve %94.1176 oranıyla KNN algoritmasıyla yüksek bir sınıflandırma başarısı elde edilmiştir. Yapılan çalışma KNN algoritmasının deprem tahmininde kullanılmasının başarılı ve hızlı sonuçlar vereceğini göstermiştir.

Gelecek çalışmalarda farklı bölgeler ve madenlerden alınan veriler ile veri kümesi genişletilerek daha sağlıklı sonuçlar almak için deneysel çalışmalar yapılacaktır.

#### KAYNAKLAR

- Blanton, H. 1997. An introduction to neural networks for technicians, engineers and other non PhDs." Proceedings of the 1997 Artificial Neural Networks in Engineering Conference. St. Louis: ANNIE'97.
- Çelik, E, M Atalay, and H Bayer 2014. Yapay Sinir Ağları ve Destek Vektör Makineleri ile Deprem Tahmininde Sismik Darbelerin Kullanılması. IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014). Trabzon. 730-733.
- Çilli, M. 2007. İnsan Hareketlerinin Modellenmesi ve Benzeşiminde Temel Bileşen Analizi Yönteminin Kullanılması. Ankara: Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü.
- Göker, D. 2010. Sürekli Radon Gazı Ölçümlerinin Deprem Tahmin Parametresi Olarak Kullanılması: İzmir Seferhisar Doğanbey Fayhattı Örneği. İzmir: Ege Üniversitesi Fen Bilimleri Enstitüsü.
- Kabiesz, J, B Sikora, M Sikora, and L Wrobel. 2013. Application of rule-based models for seismic hazard prediction in coal mines." Acta Montanistica Slovaca, 262-272.
- Karaman, M E. 2006. Yapısal Jeoloji ve Uygulamaları. Ankara: Yazarın kendi yayını.
- Karimifard, S, A Ahmadian, M Khosnevisan, and M S Nambakhs. 2006. Morphological heart arrhythmia detection using Hermitian basis functions and KNN classifier. 28th IEEE EMBS Annual International Conference. New York, 1367-1370.
- Keleşoğlu, Ö, and A Fırat. 2006. Tuğla Duvardaki ve Tesisattaki Isı Kaybının Yapay Sinir Ağı ile Belirlenmesi." Fırat Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 139-141.
- Ketin, İ. 2005. Genel Jeoloji: Yerbilimlerine Giriş. Ankara: İTÜ Vakfı.
- Kutlu, F. 2001. Categorization in Hierarchically Structured Text Database. İstanbul: Bilkent Üniversitesi,.

- Sikora, M, and L Wrobel. 2010. Application of rule induction algorithms for analysis of data collected by seismic hazard monitoring systems." Archives of Mining Sciences, 55(1) 91-114.
- Şeker, S E 2008. SVM - Bilgisayar Kavramları.  
<http://bilgisayarkavramlari.sadievrenseker.com/2008/12/01/svm-support-vector-machine-destekci-vektor-makinesi/> (accessed 12 10, 2014).
- Zhou, F, and X Zhu. 2014. Eartquake Prediction Based on LM-BP Neural Network." Proceedings of the 9th International Symposium on Linear Drives for Industry Applications. Berlin: Springer-Verlag, 13-20.