

EEG TABANLI BEYİN-BİLGİSAYAR ARAYÜZÜ SİSTEMLERİNDE ÖZNETELİK ÇIKARMA YÖNTEMLERİ

Mete YAĞANOĞLU*¹, Ferhat BOZKURT¹, F. Baturalp GÜNAY¹

¹ Atatürk Üniversitesi, Bilgisayar Mühendisliği Bölümü, 25240, Erzurum, Türkiye

Anahtar Kelimeler	Özet
Öznitelik Çıkarma Beyin Bilgisayar Arayüzü Sınıflandırma EEG Ön İşleme	Bu çalışmada Elektroensefalogram (EEG) tabanlı Beyin-Bilgisayar Arayüzü (BBA) sistemlerinde çeşitli öznitelik çıkarma yöntemleri karşılaştırılmıştır. BBA sistemleri genelde 3 temel aşamadan oluşur. Ön işleme aşamasında sınıflandırmada daha iyi sonuç elde etmek için normalizasyon yöntemi kullanılmıştır. Öznitelik çıkarmada çeşitli yöntemler kullanılmış ve karşılaştırılmıştır. Sınıflandırmada ise k-en yakın komşuluk (kNN) algoritması kullanılmıştır. Bu çalışmada, BCI Competition 2003 yarışmasında sunulan sağlıklı bir insandan alınmış veri seti kullanılmıştır. Bu veri seti, kişinin bilgisayar ekranında imleci yukarı ve aşağı yönde hareket ettirmesi esnasında altı adet elektrot ile kaydedilerek elde edilmiştir. Bu veri setinden alınan EEG tabanlı BBA verileri ön işleme yapıldıktan sonra, öznitelik çıkarma yöntemleri anlatılmış ve bu veri seti için uygun yöntem önerilmiştir.

FEATURE EXTRACTION METHODS FOR EEG BASED BRAIN-COMPUTER INTERFACE SYSTEMS

Keywords	Abstract
Feature Extraction Brain Computer Interface Classification EEG Preprocessing	In this study, we compare various feature extraction methods for Brain-Computer Interface (BCI) systems based on Electroencefalogram (EEG). BCI systems in general are executed in three major steps. In the first step, called preprocessing, we use certain normalization techniques to improve the accuracy of the results. As for the second step, we use various feature extraction techniques and compare their implications. In the third step we use K-nearest neighbor (kNN) algorithm for classification purposes. In our analysis, we use a data set presented in BCI Competition in 2003, which was obtained by performing a data collection technique performed on a healthy person. The data set was generated using six electrodes attached to that person while moving a cursor up and down on a computer screen. In this study, we describe how to perform feature extraction methods and suggest which one is the most suitable for this particular data set.

1. Giriş

Başlangıcından şimdiye kadar bilim toplulukları tarafından büyük bir ilgi duyulan elektroensefalogram (EEG), Hans Verger tarafından 1924 yılında geliştirilmiştir (Collura, 1993). EEG, beyni etkileyen birçok hastalığı erken teşhis etmek ve daha iyi anlamak için kullanılan önemli bir araçtır. Beyin Bilgisayar Arayüzü (BBA), beynin normal çıkış kanallarındaki sinirlerin ve kaslara bağlı olmadan kontrol kanallarının birbirleriyle iletişimini vermektedir (Xu ve Song, 2008).

EEG, beyindeki elektriksel aktiviteyi ölçer. Burada insan kafasına belli sayıda sensör bağlanır ve bu sensörler kablo yoluyla bilgisayara irtibatlanır. Bilgisayar beynin elektriksel aktivitelerini belirlenmiş

bir zaman aralığı için kaydeder. Çeşitli hastalıklar, beynin elektriksel aktivitelerindeki değişiklikler izlenerek gözlemlenir. Örneğin EEG'nin yoğun olarak kullanıldığı bir teşhis uygulaması epilepsi hastalığı olup, standart bir EEG çalışmasına göre daha fazla anormal durum gözlenebilmektedir. Ayrıca EEG koma tanısı, ensefalopati (beyin hastalığı), beyin ölümü, uyku bozuklukları gibi gece boyunca kayıt işleminin gerektiği durumlarda faydalıdır (Chen,2014).

Beyin-Bilgisayar arayüzü (BBA) araştırmasının amacı, insan niyetlerinin oluşturduğu uygun beyin işaretlerinin çeşitli çıkış cihazları için kontrol işareti olarak yorumlayabilen ve önemi artan bir haberleşme sisteminin kurulmasıdır. Bu çıkış cihazlarına örnek olarak bilgisayar uygulaması veya nöroprotrezi olarak verilebilir (Wolpaw vd., 2002).

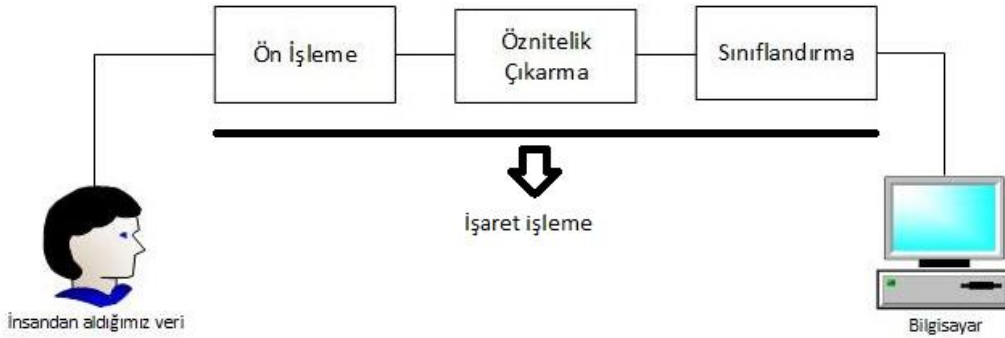
*İlgili yazar: yaganoglu@atauni.edu.tr

Günümüzde EEG-tabanlı BBA arayüzü araştırmalarının üç önemli noktası vardır. Veri kalitesi, bilinen motor çıkışlarından bağımsız olarak bütün performansı etkiler. Genelleştirme, çevrimdışı sonuçları çevrimiçi deneylerle genelleştirir. Geribeslemede ise, geribesleme sağlandığı zaman, geribeslemesiz çalışma durumlarında veri toplanmasında uygulanacak yöntemler geliştirilir (Blankertz vd., 2004). BBA 2003 yarışmasının tasarımı ilk iki konuyu kapsamaktadır. Üçüncü husus olan geribesleme konusu çevrimiçi deneylerle ifade edilebilen deneysel bir konudur. Rakiplerin katkılarında emin olmak için yalnızca merkezi sinir sistemiyle ilgili olmalıdır. Motor hareketlerinden kaynaklanan yapay durumlar (veya işlem hatası) dikkate alınmamalıdır. Yapılan çalışmalarda EEG tabanlı BBA arayüzünde kullanılan dört veri seti kullanılıp ağırlıklı olarak işlem hataları ve bu hataların veriden çıkarılmasını konu alınmıştır (Kotchoubey vd., 1997; Vaughan vd., 1998; Blankertz vd., 2003).

BBA sistemleri, beyin işaretleriyle bir bilgisayarı, elektronik kolu veya herhangi bir beyinle alakalı mekanik yapıyı kontrol eder. Bu tip arayüzler şahsın beyin işaretlerini ölçerek, niyetlerine ve özelliklerine bağlı olarak yorumlayıp harici cihazlarda kullanılacak kontrol işaretleri haline getirilmektedir. BBA teknolojisi özellikle nöromüsküler rahatsızlıklar için yardımcıdır. Bu tip rahatsızlıklara örnek olarak ALS, beyin kökü felci, omurilik yaralanması, beyin felci vb. rahatsızlıklar

örnek gösterilebilir. Bir BBA uygulaması uygun fiziksel ve zihinsel görevleri belirleme, bu görevlerle ilgili elektrotları tayin etme, söz konusu elektrotlarla kaydedilen işaretlerden ilgili özellikleri elde etme, en yüksek sınıflandırma oranını veren bir algoritma geliştirme ve dönüştürücü algoritmaları sayesinde bu bilgileri iletişim ve kontrol ünitelerine sağlama görevlerini yerine getirir. EEG işaretleri durağan olmadığından dolayı, yüksek düzeyde nesneye özel olup, fiziksel ve zihinsel görevlere bağlı ve çeşitli işlem hatalarıyla bozulabilen işaretlerdir. Elektromiyogram (EMG), Elektrokulogram (EOG) özel olarak tasarlanabilen BBA sistemleri olup bu sistemlere göre BBA uygulamaları geliştirilebilmektedir. Bundan dolayı, mevcut çalışmalar beyine ve çıkış birimi arasındaki iletişime ait hız ve doğruluk oranını artırmaya yöneliktir. Kullanıcıya özgü hızlı ve kesin sınıflandırma yönteminin geliştirilmesi ve uygulamaya dair özel ihtiyaçların karşılanabilmesi, verimli olabilecek bir EEG tabanlı iletişim ve kontrol sistemi için öncelikli öneme sahiptir. Yapılan çalışma; sınırlı sayıda öğretim işaretleri, daha iyi sınıflandırma algoritmaları, daha az kanal ve çalışmayı daha belirleyici özelliklerin görüldüğü bir öğretim sürecidir (Kayıkcıoğlu ve Aydemir, 2010).

Şekil 1’de görüldüğü gibi, İnsan-Bilgisayar Arayüzü sistemlerinde insandan alınan veriler işlenerek bilgisayar ortamına aktarılır. Bu ortam elektronik cihaz, mekanik kol da olabilir. İşaret işleme kısmında ön işleme, öznitelik çıkarma ve sınıflandırma işlemleri yapılarak doğru bir tespit yapılabilir.



Şekil 1. İnsan-Bilgisayar Arayüzü için İşaret İşleme

Bu çalışmada, BCI Competition 2003 yarışmasında sunulan sağlıklı bir insandan alınmış veri seti kullanılmıştır. Bu veri seti, kişinin bilgisayar ekranında imleci yukarı ve aşağı yönde hareket ettirmesi esnasında altı adet elektrot ile kaydedilerek elde edilmiştir. Bu veri setinden alınan EEG tabanlı BBA verileri ön işleme yapıldıktan sonra, öznitelik çıkarma yöntemleri anlatılmış ve bu veri seti için uygun yöntem önerilmiştir.

2. Materyal ve Yöntem

2.1. Veri Seti Tanıtımı

2 sınıflı bir problem bulunmaktadır.

1. Sınıf (class 0): İmlecın yukarı hareketine ait EEG işareti
2. Sınıf (class 1): İmlecın aşağı hareketine ait EEG işareti

2.1.1. Train (Eğitim) Datası

Traindata_0 matris boyutu 135x5377 'dir. Birinci sınıfa ait veriler burada bulunmaktadır. Bu matriste 135 deneme sayısını göstermektedir. 135 tane ayrı denemelerde kişi bilgisayar imlecini yukarı doğru hareket ettiğini düşünmüştür. 5377'nin birinci elemanı hep 0'dır. Birinci sınıfa ait olduğunu teyit etmek için 0 alınmıştır. Yani Traindata_0 için birinci

elemanlar hep 0'dır. İşaret işleme aşamasında birinci eleman kullanılmamaktadır.

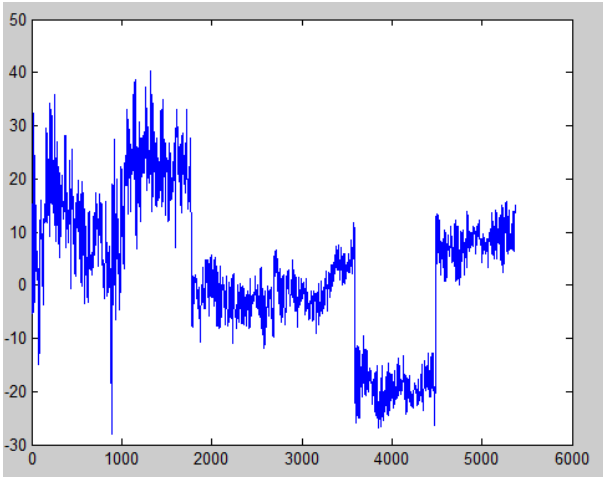
Birinci elemanı düşünmezsek 5376 işaret kalmıştır. Bu eğitim setinde 6 tane elektrottan veri alınmıştır. Her bir elektrottan 896 tane işaret elde edilmiştir. Bu işaretler peş peşe konularak oluşturulmuştur. Yani her bir elektrot boyutu 896 örnekten oluşmaktadır.

Traindata_1 matris boyutu 133x5377'dir. İkinci sınıfa ait veriler burada bulunmaktadır. 133 deneme yapılmış, birinci elemanı hep 1'dir. İkinci sınıfa ait olduğunu teyit etmek için hep 1 alınmıştır. Geri kalan veriler 6 adet elektrottan alınmıştır.

2.1.2. Test Datası

Test Data'nın boyutu 293x5376'dir. 293 deneme yapılmış, 6 tane elektrottan veriler alınmıştır.

Uygulama ortamı olarak MATLAB (Matrix Laboratory) kullanılmıştır. Matlab ortamında test datası için 1. Deneme Şekil 2'de görülmektedir. Bu şekilde, altı tane elektrottan veriler alınmıştır ve elektrotlar arası sert bir geçiş vardır.



Şekil 2. Test datası örnek

2.2. Ön işleme

Ön işleme aşamasında İstatiksel veya Z-Score normalizasyonu kullanılmıştır. Aynı veri seti üzerindeki bazı değerlerin 0'dan küçük, bazılarının ise daha büyük değerlere sahip olması, veriler arasındaki bu uzaklıkların özellikle verilerin baş veya son noktasındaki verilerin sonuçlar üzerinde daha etkin olacağını göstermektedir. Verilerin normalleştirilmesi sayesinde, eğitim giriş setindeki her bir parametrenin, modelin tahmin işlemine eşit ölçüde katkıda bulunması sağlanır. Her bir değişkene karşı gelen sütunların aritmetik ortalaması ve standart sapması bulunur. Daha sonra, aşağıdaki eşitlikte belirtilen formül ile veriler normalleştirilerek, veriler arasındaki uzaklıklar ortadan kaldırılır ve verilerdeki uç

noktaları azaltır (Yüksek vd., 2007; Yavuz ve Deveci, 2013).

$$x' = \frac{x_i - \mu_i}{\sigma_i} \quad (1)$$

Bu eşitlikte;

x' = Normalize edilmiş veriyi,

x_i = Girdi değerini,

μ_i = Girdi veri setinin ortalamasını,

σ_i = Girdi veri setinin standart sapmasını, ifade etmektedir.

2.3. Öznitelik çıkarma

2.3.1. Toplam

Ön işleme yöntemlerinden en basiti bu verilerin toplamında elde edilecek öznitelik olabilir.

$$toplama = \sum_{i=1}^n x_1 + x_2 + x_3 + \dots + x_n \quad (2)$$

Burada n, örnek sayısıdır. Veri setimizden alınan örneklerin toplamı alınarak öznitelik çıkartılmaya çalışılır.

2.3.2. Ortalama

Veri setimizde bulunan örneklerin toplamının, örnek sayısına bölünerek elde edilen öznitelik yöntemidir.

$$ortalama = toplama/n \quad (3)$$

2.3.3. Varyans (var)

Standart sapmanın karesidir. Dağılımın ölçüsüdür, elimizdeki veri setindeki değerlerin ortalamaya göre dağılımını gösterir. O an ki değerimiz ile ortalama arasındaki sapmaya (farka) göre değişim gösterir.

$$Var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \quad (4)$$

2.3.4. Standart Sapma (std)

Varyansın kareköküdür. Bir serideki değerlerin aritmetik ortalamadan farklarının kareli ortalamasıdır.

$$std(x) = \sqrt{\frac{\sum(x_i - \mu)^2}{n}} \quad (5)$$

2.3.5. Etkinlik (e)

Varyans değerinin kareköküdür.

$$e = \sqrt{var(x)} \quad (6)$$

2.3.6. Basıklık (b)

Aykırı-eğimli bir dağılımın nasıl olduğunun bir ölçüsüdür. Basıklık ölçüsü dağılımı aşağıdaki gibi ifade edilebilir:

$$b = \frac{E(x-\mu)^4}{\sigma^4} \quad (7)$$

Burada, μ x değerlerinin ortalamasını; σ , x değerlerinin standart sapmasını; $E(t)$; t miktar beklenen değeri temsil eder.

2.3.7. Çarpıklık (ζ)

Dağılımın asimetrik bir ölçüsüdür.

$$\zeta = \frac{\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^3}{\left(\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^2\right)^{3/2}} \quad (8)$$

2.3.8. Polinom uydurma (pol)

Bilinen n tane (x_i, y_i) noktası için en küçük kareler yöntemi kullanılarak $g(x) = a_0 + a_1x + a_2x^2$ fonksiyonunun elde edilmesi (Tapramaz, 2009);

$$e(a_0, a_1, a_2) = \sum_{i=1}^n [a_0 + a_1 * x_i + a_2 * x_i^2 - f(x_i)]^2 \quad (9)$$

Kısmi türevleri bulunup sıfıra eşitlenerek matrisel olarak ifade edilirse;

$$\begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n f(x_i) \\ \sum_{i=1}^n x_i * f(x_i) \\ \sum_{i=1}^n x_i^2 * f(x_i) \end{bmatrix} \quad (10)$$

Herhangi bir yöntemle çözülerek a_0 , a_1 ve a_2 katsayıları bulunarak:

$g(x) = a_0 + a_1x + a_2x^2$ elde edilir.

Buradaki katsayı değerleri kullanılarak öznitelik çıkartılmaktadır.

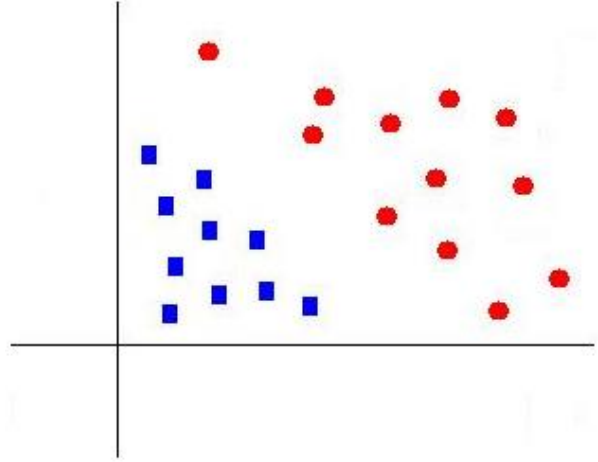
2.3. Sınıflandırma

Sınıflandırma yöntemi olarak k en yakın komşu algoritması (kNN) kullanılmıştır. kNN, eğitilmiş öğrenme algoritmasıdır ve amacı, yeni bir örnek geldiğinde var olan öğrenme verisi üzerinde sınıflandırma yapmaktır. Algoritma, yeni bir örnek geldiğinde, onun en yakın k komşusuna bakarak örneğin sınıfına karar verir (Yıldız vd., 2008).

kNN algoritmasında öncelikli olarak k değerinin belirlenmesi yapılmalıdır. k değeri belirlendikten sonra, tüm öğrenme örnekleri ile olan uzaklığının hesaplanması yapılmalıdır. Daha sonra minimum uzaklığa göre sıralama işlemi yapılır. Sıralama işleminden sonra hangi sınıf değerine ait olduğu bulunur.

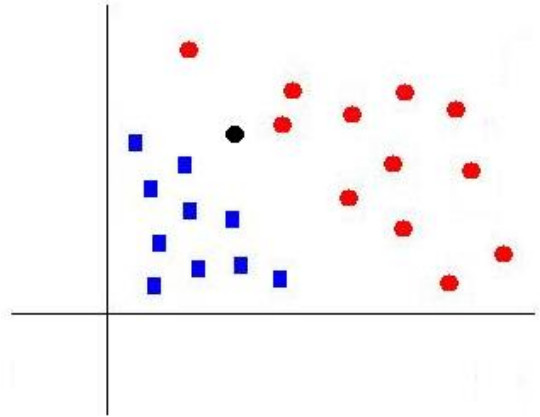
Burada k değerinin belirlenmesi, bize en yakın kaç değere bakılması gerektiğini ifade etmektedir. Örneğin k değerimiz 3 olsun. Bu durumda öğrenme kümesindeki verilerden en yakın 3 tanesi alınarak

hangi sınıfa ait olduğuna karar verilir. Aşağıdaki gibi iki boyutlu koordinat sistemine yerleştirilmiş örneklerimiz olduğunu varsayalım (Şeker, 2008):



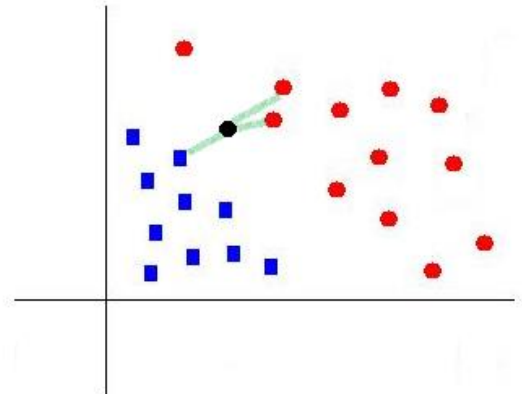
Şekil 3. İki boyutlu koordinat sisteminde örnek veriler

Şekil 3'deki örneklerden maviye mi yoksa kırmızıya mı benzediğini tespit etmemiz gereken bir de siyah örneğimizin olduğunu düşünelim.



Şekil 4. Yeni bir örnek geldiğinde hangi sınıfa dahil olması

Şekil 4'te görüldüğü gibi k değerimizi 3 aldığımızdan siyah örneğe en yakın olan 3 örneğe bakarız. k değeri 5 olsaydı en yakın 5 örneğe bakacaktık. Bu bakımdan k'nın seçimi önemlidir.



Şekil 5. Gelen örneğin sınıfının belirlenmesi

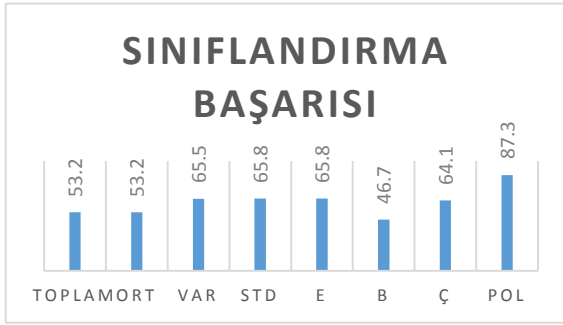
Şekil 5'te görüldüğü gibi en yakın 3 örnekten 2 tanesi kırmızı olduğundan siyah örnek için de kırmızı sınıfa ait diyebiliriz.

Yöntem uzaklıkların hesaplanmasında Öklit uzaklığı kullanılmaktadır.

3. Araştırma Bulguları

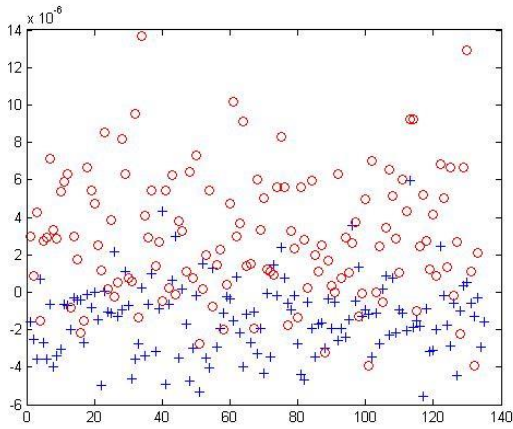
Bu veri setinin 1.elektrotlarından alınan veriler ile, yukarıda anlatılan sekiz adet öznitelik çıkarma yöntemi kullanılarak elde edilen sonuçlar grafikte gösterildiği gibidir. Görüldüğü gibi en iyi sınıflandırma başarısını polinom uydurma yönteminde elde edilmiştir. Polinom uydurma yönteminde elde edilen katsayıların ilk katsayı sınıflar arasında farklılığı açık bir şekilde göstermiştir.

Şekil 6'da görüldüğü gibi, polinom uydurma öznitelik yöntemi kullanıldığında %87.3'lük bir başarı elde edilmiştir.



Şekil 6. Sınıflandırma başarısı

Şekil 7'de görüldüğü gibi iki sınıf verileri arasında gözle görülür bir farklılık vardır. Polinom uydurma ile yapılan öznitelik sonucunda şekilden de görüldüğü gibi mavi ile gözükken sınıf negatif değerler, kırmızı ile gösterilen değerler ise pozitif değerler almıştır. Bu şekilde iyi bir sınıflandırma başarısı elde edilmiş olur.



Şekil 7. Polinom uydurma özniteliği sonucunda elde edilen sınıf farklılığı

4. Tartışma ve Sonuç

Çoklu verilerle işlem yaparken EEG verisi ile alınan verilerden hangi elektrotların bizler için daha anlamlı olduğu uygulamalarda büyük bir problemdir. Kullanılan öznitelik yöntemi uygulamalardan uygulamalara farklılık gösterebilir. Bu tarz çalışmalarda önemli olan birden çok öznitelik bulup bunların beraber kullanımını yapabilmektir. Bu şekilde daha iyi bir sınıflandırma başarısı elde edilebilir. Örneğin bu uygulamada etkinlik ve basıklık beraber kullanıldığında ayrı ayrı kullanımlarından daha iyi bir sonuç elde edildiği görülmektedir.

Bu veri seti için polinom uydurma özniteliği yanında başka bir öznitelik ile beraber kullanılarak başarımlar artırılabilir. Aynı zamanda sınıflandırma için kullanılan kNN yöntemi yerine başka yöntemler kullanılarak sınıflandırma başarısı değişebilir.

Teşekkür

Çalışmalarımız sırasında yardımlarını esirgemeyen Dr. Önder Aydemir'e desteklerinden dolayı çok teşekkür ederiz.

Conflict of Interest

No conflict of interest was declared by the authors.

5. Kaynaklar

Blankertz, B., Dornhege, G., Schafer, C., Krephi, R., Kohlmorgen, J., Müller, K.R., Kunzmann, V., Losch, F., and Curio, G., 2003. Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis, IEEE Trans. Neural Sys. R. Eng., 11(2), 127-131.

Blankertz, B., Müller, K-R., Curio, G., Vaughan, T.M., Schalk, G., Wolpaw, J.R., Schlögl, A., Neuper, C., Pfurtscheller, G., Hinterberger, T., Schröder, M., Birbaumer, N., 2004. The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG Single Trials., IEEE Transaction on Biomedical Engineering

Chen, G. 2014. Are electroencephalogram (EEG) signals pseudo-random number generators? Journal of Computational and Applied Mathematics, 268, 1-4.

Collura, T. F. 1993. History and evolution of electroencephalographic instruments and techniques. Journal of Clinical Neurophysiology, 10(4), 476-504.

Kayikcioglu, T., Aydemir, O. 2010. A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data, Pattern Recognition Letters, 31, 1207-1215.

Kotchoubey, B., Schleichert, H., Lutzenberger, W., and Birbaumer, N., 1997. A new method for self-regulation of slow cortical potentials in a timed paradigm, *Appl. Psychophysiol. Biofeedback*, 22,(2), 77-93.

Şeker, Ş.E., 2008. KNN (K Nearest neighborhood, en yakın k komşu). <http://bilgisayarkavramlari.sadievrenseker.com/2008/11/17/knn-k-nearest-neighborhood-en-yakin-k-komsu/> (Erişim Tarihi:26.08.2014).

Tapramaz, R., 2009. Sayısal Çözümleme, Literatür Yayıncılık.

Xu, B-G., Song, A-G., 2008. Pattern recognition of motor imagery EEG using wavelet transform, *J. Biomedical Science and Engineering*, 1, 64-67.

Vaughan, T. M., Miner, L.A., McFarland, D.J., and Wolpaw, J.R., 1998. EEG-based communication: analysis of concurrent EMG activity, *Electroencephalogr. Clin. Neurophysiol*, 107, 428-433.

Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M., Brain-computer interfaces for communication and control. *Clin. Neurophysiol*, 113, 767-791.

Yavuz, S., Deveci, M., 2013. İstatiksel Normalizasyon Tekniklerinin Yapay Sinir Ağın Performansına Etkisi. *Erciyes Üniversitesi İİBF dergisi*, 40, 167-187.

Yıldız, T., Yıldırım, S., Altılar, T., 2008. İstenmeyen İletilerin Paralelleştirilmiş KNN Algoritması ile Tespiti. *Akademik Bilişim*, 2008, Çanakkale.

Yüksek, Ahmet G., Bircan H., Zontul, M. ve Kaynar O., 2007. Sivas İlinde Yapay Sinir Ağları ile Hava Kalitesi Modelinin Oluşturulması Üzerine Bir Uygulama. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 8, 97-112.