Elektrokardiyografi (EKG) Sinyallerindeki Aritmilerin Sınıflandırılması

Classification of Arrhythmia in Electrocardiography (ECG) Signals

Volkan Erol
Bilgisayar Mühendisliği Bölümü
Okan Üniversitesi
İstanbul, Türkiye
voerol@stu.okan.edu.tr

Aslı Uyar Özkaya
Bilgisayar Mühendisliği Bölümü
Okan Üniversitesi
İstanbul, Türkiye
asli.uyar@okan.edu.tr

***Özetçe*—Bu çalışmada, elektrokardiyografi (EKG) sinyallerinde gözlenen aritmiler, önceden belirlenmiş 16 sınıfın özelliklerine göre sınıflandırılmaya çalışılmıştır. Buradaki amaç kardiyag aritmilerin 2 sınıf altında toplanmasında önerilmiş bir yöntemin; Naive Bayes, k-En Yakın Komşuluk, Oylama Öznitelik Aralıkları ve PART sınıflandırıcıları ile deneylerinin yapılması ve örnek verilerin normal ve aritmi sınıflarına göre sınıflandırılmasının başarımının ölçümlenmesidir.**

Anahtar Kelimeler — Elektrokardiyografi; Aritmi; k-En Yakın Komşuluk Sınıflandırıcısı; Oylama Öznitelik Aralıkları Sınıflandırıcısı; PART Sınıflandırıcısı; Naive Bayes Sınıflandırıcısı.

***Abstract*—In this work, we tried to classify predefined electrocardiography (ECG) signal arrhythmia types into 16 classes' particularities. Our aim is to classify the example data into 2 classes as normal and arrhythmia by using Naïve Bayes, k-Nearest Neighborhood, Voting Feature Intervals and PART classifiers. We tried to measure the correctness of four techniques in classification.**

Keywords — Electrocardiography; arrhythmia; k-Nearest Neighborhood Classifier; Voting Feature Intervals Classifier; Naïve Bayesian Classifier.

978-1-4673-5563-6/13/$31.00 ©2013 IEEE

I. GİRİŞ

Aritmi, kalpteki elektriksel sinyallerin üretim ve iletimindeki anormalliklerden kaynaklanan, kalbin atış hızı ve düzenliliğindeki bozulmalardır. Bazı aritmiler tedavi gerektirmezken bazı aritmiler ölümle sonuçlanabilir. Bu yüzden aritmilerin doğru sınıflandırılması uygulanacak tedavinin belirlenmesinde yaşamsal önem taşır. Bu konuda pek çok araştırmacı, EKG sinyalinden farklı öznitelikler çıkararak farklı sınıflandırıcılarla en iyi performansı elde etmeye çalışmaktadır.

Literatürde dalgacık dönüşümü [1], Fourier dönüşümü [2] ve morfolojik özellikler [3] aritmi sınıflandırmada öznitelik çıkarmak için kullanılmıştır. Bu metodlardan elde edilen öznitelikler destek vektör makinaları [1,4], özdüzenleyici haritalar [5], bulanık mantık [6] ve yapay sinir ağları [7] gibi sınıflandıcılarla kullanılmıştır.

II. MATERYAL VE YÖNTEM

A. Verilerin Elde Edilmesi ve Hazırlanması

Bu çalışmada, UCI Arrhythmia veritabanındaki EKG kayıtlarından elde edilen 452 vuru kullanılmıştır. Her bir vuruya ait verilerin sahip olduğu 279 özellikten 103 özniteliğe

Temel Bileşen Analizi Yöntemi kullanılarak indirgenmiştir ve vuruların bu özellikleri öznitelik olarak kullanılmıştır.

Verilerin onisleme aşamasında Nümerik değerleri Nominal değerlere çevrilmesini sağlayan Weka yazılımındaki NumericToNominal yöntimi kullanılmıştır.

B. *Naïve Bayes Sınıflandırıcısı*

Naïve Bayes Sınıflandırıcı adını İngiliz matematikçi Thomas Bayes'ten (yak. 1701 - 7 Nisan 1761) alır. Naïve Bayes Sınıflandırıcı Örüntü tanımaproblemine ilk bakışta oldukca kısıtlayıcı görülen bir önerme ile kullanılabilen olasılıkcı bir yaklaşımdır. Bu önerme örüntü tanıma da kullanılacak her bir tanımlayıcı nitelik ya da parametrenin istatistik açıdan bağımsız olması gerekliliğidir. Her ne kadar bu önerme Naive Bayes Sınıflandırıcısının kullanım alanını kısıtlandırsa da, genelde istatistik bağımsızlık koşulu esnetilerek kullanıldığında da daha karmaşık Yapay sinir ağları gibi metodlarla karşılaştırabilir sonuçlar vermektedir. Naive Bayes Sınıflandırması Makine öğreniminde öğreticili öğrenme alt sınıfındadır.

C. *K-En Yakın Komşu Sınıflandırıcısı*

En yakın k komşu algoritması örüntü tanımada çok yaygın olarak kullanılan parametrik olmayan yöntemlerden biridir. Bilinmeyen bir örüntü vektörü kendisine en yakın k komşularının çoğunluğunun sahip olduğu etiket ile etiketlendirilir. Bu yöntemin örüntü vektörünün istatistiksel dağılımından bağımsız olması da bir avantajdır. Uzaklık ölçütü bu yöntemde kullanılan değişken parametrelerden biridir. Yaygın olarak kullanılan uzaklık ölçütleri Öklit ve şehir bloklarıdır.

D. *Oylama Öznitelik Aralıkları Sınıflandırıcısı*

Oylama Öznitelik Aralıkları (VFI) algoritması arttırımsal olmayan bir sınıflandırma algoritmasıdır[8]. Her bir örnek ve özelliği, çalışma değerlerinin bir vektör ayrıca örneğin sınıfı temsil eden bir etiket olarak temsil edilir. Eğitim örneklerinden, VFI algoritması her öznitelik için öznitelik aralıkları belirler. İki komşu aralık farklı sınıf kümelerine dahil olabilirler. Her özellik sınıfları arasında gerçek sayılarla oy dağıtarak sınıflandırma katılır. VFI algoritması genelde Naive Bayes algoritması ile öznitelikleri teker teker ele aldığı için kıyaslanmaktadır.

E. *PART Sınıflandırıcısı*

PART, Weka yazılımı içerisinde PART karar listesini oluşturmak için yazılmış bir modüldür. Böl ve yönet yaklaşımında göre hareket etmektedir. Her iterasyonda C4.5 algortimasını kullanan kısmi bir karar ağacı oluşturmaktadır

F. *Değerlendirme*

Verilerin işlenmesi ve sınıflandırılması için Weka programı kullanılmıştır. Sınıflandırıcının performansı aşağıdaki ölçütlere göre belirlenmiştir:

$$Duyarlilik = \frac{TP}{TP+FN} x100\% \quad (1)$$

$$Belirlilik = \frac{TN}{TN+FP} x100\% \quad (2)$$

$$Seçicilik = \frac{TP}{TP+FP} x100\% \quad (3)$$

$$GenelBasarim = \frac{TP+TN}{TP+TN+FP+FN} x100\% \quad (4)$$

Burada, TP sınıflandırıcının doğru olarak sınıflandırdığı ilgilenilen aritmi tipine ait vuru sayısı, TN sınıflandırıcının doğru olarak sınıflandırdığı diğer tip aritmik vuruların sayısı, FP sınıflandırıcının yanlış olarak sınıflandırdığı ilgilenilen aritmi tipine ait vuru sayısı, FN ise sınıflandırıcının yanlış olarak sınıflandırdığı diğer tip aritmik vuruların sayısıdır.

III. UYGULAMA VE BAŞARIMLAR

Naive Bayes, k-En Yakın Komşuluk, VFI ve PART Sınıflandırıcıları ile 10-Katlı Çapraz Geçerlilik ve 2/3 Öğrenme 1/3 Test Kümesi Dağılımı Uygulanarak çeşitli deneyler yapılmıştır.

Naive Bayes 10-Katlı Çapraz Geçerlilik (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	69	88,2	79,1	72,1
Aritmi	79,1	53,1	69	72,1

Tablo 1- Naive Bayes Sınıflandırıcısı ile 10-Katlı Çapraz Geçerlilik Uygulanarak Elde Edilen Sonuçlar

Naive Bayes 2/3 Öğrenme Test Kümesi Dağılımı (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	70,8	90,9	80,5	73,3
Aritmi	80,5	50	70,8	73,3

Tablo 2- Naive Bayes Sınıflandırıcısı ile 2/3 Öğrenme 1/3 Test Kümesi Dağılımı Uygulanarak Elde Edilen Sonuçlar

k-En Yakın Komşuluk 10-Katlı Çapraz Geçerlilik (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	62,6	91,4	77,7	65,7
Aritmi	77,7	35,3	62,6	65,7

Tablo 3- k-En Yakın Komşuluk Sınıflandırıcısı ile 10-Katlı Çapraz Geçerlilik Uygulanarak Elde Edilen Sonuçlar

k-En Yakın Komşuluk 2/3 Öğrenme Test Kümesi Dağılımı (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	61,6	96,6	81,3	63,6
Aritmi	81,3	19,7	61,6	63,6

Tablo 4- k-En Yakın Komşuluk Sınıflandırıcısı ile 2/3 Öğrenme 1/3 Test Kümesi Dağılımı Uygulanarak Elde Edilen Sonuçlar

VFI 10-Katlı Çapraz Geçerlilik (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	69,1	61,2	59,6	64,1
Aritmi	59,6	53,1	69,1	64,1

Tablo 5- VFI Sınıflandırıcısı ile 10-Katlı Çapraz Geçerlilik Uygulanarak Elde Edilen Sonuçlar

VFI 2/3 Öğrenme Test Kümesi Dağılımı (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	70,3	72,7	61,9	66,8
Aritmi	61,9	59,1	70,3	66,8

Tablo 6- VFI Sınıflandırıcısı ile 2/3 Öğrenme 1/3 Test Kümesi Dağılımı Uygulanarak Elde Edilen Sonuçlar

PART 10-Katlı Çapraz Geçerlilik (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	61,6	94,3	81,8	65
Aritmi	81,8	30,4	61,6	65

Tablo 7- PART Sınıflandırıcısı ile 10-Katlı Çapraz Geçerlilik Uygulanarak Elde Edilen Sonuçlar

PART 2/3 Öğrenme Test Kümesi Dağılımı (yüzde)				
Sınıf	***Duyarlılık***	***Seçicilik***	***Belirlilik***	***Genel Başarım***
Normal	66,9	94,3	83,3	70,1
Aritmi	83,3	37,9	66,9	70,1

Tablo 8- PART Sınıflandırıcısı ile 2/3 Öğrenme 1/3 Test Kümesi Dağılımı Uygulanarak Elde Edilen Sonuçlar

IV. SONUÇLAR VE TARTIŞMA

Bu çalışmada, UCI Arrhythmia veritabanındaki EKG kayıtlarından elde edilen aritmik vurular J48 ve Naive Bayes sınıflandırıcıları kullanılarak sınıflandırılmıştır. Normal ve 15 çeşit aritmiye ait olmak üzere toplam 452 vuru kullanılmıştır. Her bir vuruya ait verilerin sahip olduğu 279 özellikten 103'ü öznitelik olarak kullanılmıştır.

Naive Bayes Sınıflandırıcısı ile yapılan deneylerde 10-katlı çapraz geçerlilik uygulandığında genel başarım oranı yüzde 72,1 olmuştur. 2/3 Öğrenme 1/3 Test Kümesi Dağılımı kullanıldığında bu oran yüzde 73.3'e çıkmıştır. Her iki yöntemde de bu sınıflandırıcının duyarlılık, seçicilik ve belirlilik değerlerinin normal sınıfı için yaklaşık yüzde 80 ve üzerinde olduğunu görmekteyiz.

K-En Yakın Komşuluk Sınıflandırıcısı ile yapılan deneylerde 10-katlı çapraz geçerlilik uygulandığında genel başarım oranı yüzde 65,7 olmuştur. 2/3 Öğrenme 1/3 Test Kümesi Dağılımı kullanıldığında bu oran yüzde 63,6 olmuştur. Aritmi sınıfı için elde edilen seçicilik değerleri oldukça düşüktür.

Oylama Öznitelik Aralıkları Sınıflandırıcısı ile yapılan deneylerde 10-katlı çapraz geçerlilik uygulandığında genel başarım oranı yüzde 64,1 olmuştur. 2/3 Öğrenme 1/3 Test Kümesi Dağılımı kullanıldığında bu oran yüzde 66,8 olmuştur. Her iki sınıf içinde elde edilen değerler yüzde 50 ile 70 arasından yer almaktadır.

PART Sınıflandırıcısı ile yapılan deneylerde 10-katlı çapraz geçerlilik uygulandığında genel başarım oranı yüzde 65,1 olmuştur. 2/3 Öğrenme 1/3 Test Kümesi Dağılımı kullanıldığında bu oran yüzde 70,1 olmuştur. Normal sınıfı için elde edilen değerler oldukça yüksek olmasına rağmen aritmi sınıfının değerleri daha düşüktür.

Her 4 sınıflandırıcıda özellikle aritmi sınıfı için edilen başarım oranlarının çok yüksek olmadığını görüyoruz. Bu problem ve veri kümesi için Naive Bayes sınıflandırıcısı elde ettiğimiz sonuçlara göre daha uygun gözükmektedir. Bu sınıflandırıcıda Aritmi sınıfı için de elde edilen sonuçlar daha yüksek olmuştur.

KAYNAKÇA

[1] Song , M. H., Lee, J., Cho, S.P., Lee, K.J. ve Yoo, S.K. "Support vector machine based arrhythmia classification by using reduced features", *International Journal of Control, Automation and Systems*, 3(4):571-579, 2005.

[2] Minami, K., Nakajima, H. Ve Toyoshima, T. "Real time discrimination of ventricular tachyarrhythmia with fourier transform neural network", *IEEE Transection on Biomedical Engineering, Vol.46*, 1999, p.179-185.

[3] Palreddy, S., Hu, Y.H. ve Tompkins, W.J. "A patient adaptable ecg beat classifier using a mixture of experts approach", *IEEE Tarnsection on Biomedical Engineering*, 44(9):891-900, 1997.

[4] Acır, N., "Classification of ECG beats by using a fast least square support vector machines with a dynamic programming feature selection algorithm", *Neural Computing and Applications*, 14(4):299-309, 2005.

[5] Braccini, G., Edenbrandt, L.,Lagerholm, M., Petrson, C., Rauer, O., Rittner, R. Ve Sörnmo, L.,"Clustering of ecg complexes using hermite functions and self organizing maps", *IEEE Transection on Biomedical Engineering*, 47(7):838:848, 2000.

[6] Osowski, S. Ve Linh, T.R., "Ecg beat recognition using fuzzy hybrid neural network", *IEEE Transection on Biomedical Engineering*, 48(11):1265:1271, 2001.

[7] Kutlu, Y., Kuntalp, M. Ve Kuntalp, D., *"Effect of the input window size in arrhythmia classification with multilayer perceptron network structures", IEEE 15.Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Eskişehir,* 11-13 Haziran 2007

[8] Demiröz, G., Güvenir, H.A. *"Classification by Voting Feature Intervals", 9th European Conference on Machine Learning Prague, Czech Republic, April 23–25, 1997 Proceedings*