

Kompleks Değerli Yapay Sinir Ağları ile Kardiyotokogram Verisinde Morfolojik Örüntü Sınıflandırma

Eda Çapa Kızıltaş^{1*}, Ayşenur Uzun² and Ersen Yılmaz²

¹Borçelik Çelik Sanayi Ticaret A.Ş., Bursa, Turkey

²Electrical-Electronic Engineering, Bursa Uludağ University, Bursa, Turkey

*Corresponding author: eckiziltas@borcelik.com

⁺Speaker: eckiziltas@borcelik.com

Presentation/Paper Type: Oral / Full Paper

Özet – Bu çalışmada UCI Makine Öğrenmesi havuzundan alınan kardiyotokogram verisinde morfolojik örüntü sınıflandırması yapılmıştır. Sınıflandırıcı olarak kompleks değerli yapay sinir ağları kullanılmıştır. Sınıflandırıcının performansı 10-katlı çapraz doğrulama yaklaşımı kullanılarak incelenmiştir. Performans sonuçları doğruluk, duyarlılık ve özgüllük ölçütleri kullanılarak detaylı biçimde sunulmuştur.

Kardiyotokografi, Kompleks Değerli Yapay Sinir Ağları, Makine Öğrenmesi, Sınıflandırma

I. GİRİŞ

Kardiyotokografi (KTG), diğer adıyla Elektronik Fetal İzleme, genellikle gebeliğin son üç ayında fetal iyiliği veya sıkıntısını belirlemek amacıyla kullanılan, fetal kalp hızı (FKH) ve uterus kasılmalarını kaydeden bir görüntüleme cihazıdır [1]-[6]. KTG'den gelen bilgilerin uzmanlar tarafından yorumlanması fetal sıkıntısının erken tespitinde hayati bir rol oynar [7]. Doğru şartlarda gerçekleşmeyen ya da doğru yorumlanamayan KTG ölçümü fetal iyiliği durumunda gereksiz sezaryen kararı almaya ya da fetal sıkıntısını fark edememeye neden olabilir. Bu senaryolar anne ve fetal sağlığı açısından KTG verisinin kritik önemde olduğunu göstermektedir.

Var olan önem Ayres de Campos ve arkadaşlarını, KTG ölçümleri ve uzmanların yorumlarıyla ortaya çıkan sınıflardan oluşan bir veri seti hazırlamaya yöneltmiştir [8], [15]. Bu sayede KTG verisi özellikle sınıflandırma algoritması ve biyomedikal veriler üzerine yoğunlaşmış akademik çalışmalarda sıklıkla kullanılan bir veri seti haline gelmiştir. Veri seti toplam 2126 gerçek durum ölçümlerinden damıtılmış 21 nitelikten oluşmaktadır.

Tablo1. KTG Verisinin Özellikleri [12]

| Kısaltma | Açıklama |
|----------|---|
| LB | Bazal Hız Değeri |
| AC | Hızlanma |
| FM | Fetal Hareketlilik |
| UC | Uterin Yoğunluğu |
| DL | Hafif Yavaşlama |
| DS | Ciddi Yavaşlama |
| DP | Uzun Süreli Yavaşlama |
| ASTV | Anormal Kısa Dönem Değişkenliği Yüzdesi |
| MSTV | Ortalama Kısa Dönem Değişkenliği |

| | |
|----------|---|
| ALTV | Anormal Uzun Dönem Değişkenliği Yüzdesi |
| MLTV | Ortalama Uzun Dönem Değişkenliği |
| Width | Histogram Genişliği |
| Min | Histogramda en düşük frekans |
| Max | Histogramda en yüksek frekans |
| Nmax | Histogramda zirve sayısı |
| Nzeros | Histogramda sıfır frekans |
| Mode | Histogram Modu |
| Mean | Histogram Ortalaması |
| Median | Histogram Medyanı |
| Variance | Histogram Varyansı |
| Tendency | Histogram Eğimi |

Bu özellikler uzmanların yorumları sonucunda 3 sınıflı ve 10 sınıflı olmak üzere iki farklı grupta sınıflandırılmıştır. Fetal durumuna göre gruplanmış 3 sınıf; normal, şüpheli ve patolojik (NSP) olarak isimlendirilmiştir. Şüpheli olarak adlandırılan sınıf, uzmanların da fetal durumuna dair net bir karar veremedikleri ve fetalın sağlığını riske etmemek için ilave ölçüm yöntemlerine ya da acil sezaryene başvurdukları durumlardır. 10 sınıflı morfolojik örüntü ise FKH ve uterus kasılmalarını detaylı biçimde sınıflandırmaktadır. Tablo-2'de görüldüğü gibi 197 şüpheli (SUSP) sınıf üyesine ek olarak, sakın uyku ve şüpheli olduğuna karar verilememiş (SH) 72 adet daha veri bulunmaktadır. Bu dağılımdan anlaşılacağı üzere 10 sınıflı KTG verisinin işlenmesi, veri setinin doğası gereği güçlükler barındırmaktadır.

Tablo2. Morfolojik Desen (MD) İçin Sınıf Açıklamaları [12]

| Kısaltma | Sınıf detayı | Adet |
|----------|--------------|------|
| A | Sakin Uyku | 384 |
| B | REM Uykusu | 579 |

| | | |
|------|---|-----|
| C | Uyanık Kalma | 53 |
| D | Aktif Uyanıklık | 81 |
| SH | A veya SUSP olarak yorumlanabilir | 72 |
| AD | Hızlanma-Yavaşlama Deseni, Hareketlilik Durumu) | 332 |
| DE | Yavaşlama Deseni | 252 |
| LD | Geniş Yavaşlama Deseni | 107 |
| FS | Yatay Sinüzoidal Desen | 69 |
| SUSP | Şüpheli Desen | 197 |

Literatürde KTG verileri ile yapılan çalışmaların büyük kısmı fetal iyiliğini saptamak için 2'li (NP) ya da 3'lü (NSP) sınıflandırma üzerine yoğunlaşmıştır. Ayrıca MD sınıflandırmasını amaçlayan çalışmalar da literatürde yer almaktadır, örn. [2], [7], [14].

Baluz ve diğerleri [2] (2011), farklı makine öğrenme algoritmalarını KTG örnekleri üzerinde 2'li NP ve 10'lu MD sınıflandırma için kullanmışlardır. Çalışmalarında en iyi sınıflandırma sonuçlarını Rastgele Orman (RO) algoritması ile elde etmişlerdir. İlk sınıflandırma için %94,90 oranında, ikinci sınıflandırma için %87,30 oranında bir doğruluk belirtmişlerdir. Huang ve arkadaşları [9] (2012) Diskriminant Analizi (DA), Karar Ağacı (KA) ve Yapay Sinir Ağı (YSA) yöntemlerini kullanarak KTG verilerini NSP için sınıflandırmış ve yöntemler için sırasıyla %82,10, %86,36, %97,78 doğruluk oranı elde etmişlerdir. Yılmaz ve Kılıççer [10] (2013), İkili Karar Ağacı (İKA) ve En Küçük Kareler-Destek Vektör Makinesi (EKK-DVM) yöntemlerini birlikte kullanarak NSP sınıflandırması için %91,62 doğruluk oranına ulaşmışlardır. Ocak ve Ertunç [11] (2013) KTG'nin NP sınıflandırması için Uyarlanabilir Sinirsel Bulanık Çıkarım Sistemi (USBÇS) yöntemini kullanmış, sırasıyla patolojik ve normal sınıf için %96,60 ve %97,20 doğruluk sonuçlarına ulaşmışlardır. Aladağ ve Güven ise [12] (2014) fetal iyiliğine yoğunlaşarak inceledikleri KTG verilerini YSA ile sınıflandırmış ve sağlıklı fetalleri %97,35 oranında, patolojikleri ise %73,86 oranında tespit etmişlerdir. Şahin ve Subasi [5] (2015) 2'li NP sınıflandırması için bir çok makine öğrenmesi metodunu karşılaştırmış, en yüksek sınıflandırma doğruluğunu k-En Yakın Komşuluk (k-EYK) ve RO makine öğrenme algoritmalarını kullanarak sırasıyla %98,42 ve %99,18 olarak yakalamışlardır. Chamidah ve arkadaşları [4] (2015) Melez k-Ortalama Kümeleme (Mk-OK) metodu kullanarak özellik sayısını indirgedikleri KTG verisini DVM ile NSP'ye göre sınıflandırmış ve %90,64 doğruluk oranı elde etmişlerdir. Cömert ve diğerleri [13] (2016) YSA ve Aşırı Öğrenme Makineleri (AÖM) yöntemlerini karşılaştırmış, AÖM yönteminin %93,42'lik doğruluk oranı yanında daha hızlı olduğunu da belirtmişlerdir. MD sınıflandırma için Kamath ve Kamat [14] (2017) çalışmalarında KTG verilerine RO modelinin uygulanmasını önermişlerdir. Önerdikleri model ile %75,79 doğruluk oranı sağlamışlardır.

II. METOT

A. Kompleks Değerli Yapay Sinir Ağları (KD-YSA)

Karmaşık sayı uzayı ile işlem yapan iki boyutlu sinir ağı modelleri kompleks değerli sinir ağları olarak adlandırılır

[16]. Bu modeller, geleneksel sinir ağı modellerinden giriş değerleri, nöron ağırlıkları ve aktivasyon fonksiyonları ile ayrışır [16].

KD-YSA kompleks değerli girişlere ihtiyaç duyarlar. Eğer giriş değerleri kompleks değilse, reel sayıları kompleks sayı uzayına taşımak gerekmektedir [17]-[18]. Bu çalışmada [18]'de önerilen yaklaşım kullanılmıştır.

Veri setindeki her bir özellik için;

$$X = (x_1, x_2, x_3, \dots, x_m) \quad X \in \mathfrak{R}^m \quad (1)$$

$\mathfrak{R}^m \rightarrow \mathbb{C}^m$ dönüşümünün yapılması gerekir.

$\forall x_j \quad j \in \{1, 2, \dots, m\}$ ve $x_j \in [a, b]$ ve $a, b \in \mathfrak{R}$ için

$$\varphi_j = \frac{\pi(x_j - a)}{(b - a)}, \quad \varphi_j \in [0, \pi] \quad (2)$$

olarak ifade edilebilir.

Dönüştürülen değerler, kompleks değerli ağırlık (w_j) ve yan (b_k) değerleri kullanılarak (3) denklemi ile k . çıkış aktarılır.

$$z_k = \left(\left(\sum_{j=1}^m w_j^R x_j^R - w_j^S x_j^S \right) + b_k^R \right) + i \left(\left(\sum_{j=1}^m w_j^S x_j^R + w_j^R x_j^S \right) + b_k^S \right) \quad (3)$$

Çıkış katmanında değerlerin kompleks değerden reel değere dönüşmesi için [18]'te verilen ve (4) denklemi ile gösterilen aktivasyon fonksiyonu kullanılmıştır.

$$f_{act}(z_k) = (f_a(z_k^R) - f_a(z_k^S))^2 \quad (4)$$

Burada $f_a(x)$ sigmoid fonksiyonudur.

$$f_a(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

Bu çalışmada kullanılan ağ modeli Geriye Doğru Yayılım algoritması ile eğitilmiştir. Ağırlık katsayıları ve yan değerlerinin güncellenmesinde Gradyen inişi kullanılmıştır.

B. k-Katlı Çapraz Doğrulama

k kez yinelenen adımlarda k eş parçaya bölünmüş verinin denenmemiş bir kısmı test verisi, kalan kısmı ise öğrenme verisi olarak kullanılır [6]. Test verileri kullanılarak bulunan tahminleme doğruluklarının ortalaması modelin veri setini temsil etmedeki başarısını gösterir.

C. Doğruluk – Duyarlılık - Özgüllük

Tahminleme doğruluğu ya da sınıflandırma doğruluğu, bir modelin sınıf üyelerini doğru atayabilme gücü olarak ifade edilebilir ve (6) denklemi ile hesaplanır.

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Burada TP, TN, FP ve FN gerçek pozitif, gerçek negatif, yanlış pozitif ve yanlış negatif sayılarıdır.

Duyarlılık gerçek pozitif oranı, Özgüllük ise gerçek negatif oranı olarak tanımlanır. Duyarlılık ve Özgüllük (7) ve (8) denklemleri kullanılarak hesaplanır [6].

$$Duyarluluk = \frac{TP}{TP + FN} \quad (7)$$

$$Özgüllük = \frac{TN}{TN + FP} \quad (8)$$

Tablo3. Morfolojik Örtüntü İçin KD-YSA Sonuçları

| Öğrenme Oranı | Gizli Katman Nöron Sayısı | İterasyon | Ortalama Tahminleme Doğruluğu | Özgüllük ve Duyarlılık | | | | | | | | | | |
|---------------|---------------------------|-----------|-------------------------------|------------------------|-------|-------|-------|--------|--------|--------|--------|--------|----------|-------|
| | | | | A (%) | B (%) | C (%) | D (%) | SH (%) | AD (%) | DE (%) | LD (%) | FS (%) | SUSP (%) | |
| 0,00010 | 10 | 50000 | 81,28% | Duyarlılık | 88,28 | 94,13 | 42,33 | 70,56 | 48,93 | 91,92 | 85,02 | 58,09 | 35,71 | 62,50 |
| | | | | Özgüllük | 94,57 | 97,74 | 99,13 | 99,03 | 98,79 | 97,27 | 97,55 | 98,62 | 98,35 | 97,48 |
| 0,00030 | 10 | 50000 | 81,81% | Duyarlılık | 88,81 | 97,07 | 33,00 | 73,06 | 50,71 | 89,50 | 84,23 | 65,55 | 34,29 | 61,08 |
| | | | | Özgüllük | 93,60 | 98,01 | 99,08 | 99,02 | 98,30 | 98,72 | 98,46 | 98,47 | 98,45 | 97,11 |

III. DENEMELER

Çalışmamızda bir gizli katmana sahip KD-YSA kullanılmıştır. KD-YSA modelinde 21 giriş nöronu, 10 gizli nöron ve 10 çıkış nöronu bulunmaktadır.

Modelin performans değerlendirmesi yapılırken 10-katlı Çapraz Doğrulama yapılmıştır. KD-YSA'nın eğitilmesinde öğrenme oranı olarak 0.0001 ve 0.003 kullanılmış ve 50000 iterasyon sonucunda ağırlık katsayıları ve yan değerleri belirlenmiştir.

Sınıflar bazında duyarlılık ve özgüllük oranları ile ortalama tahminleme doğruluğu sonuçları Tablo 3'te detaylı bir biçimde sunulmuştur.

IV. SONUÇLAR

Daha önce Aşırı Öğrenme Makineleri [7] uygulanarak sınıflandırılmış KTG verisinde performans değerlerinin KD-YSA modeline göre daha yüksek olduğu gözlenmiştir. KD-YSA modelinde on sınıfın beş tanesi için tahminleme doğruluğu düşük çıkmıştır. Bu beş sınıf içinde C, LD, SH ve FS sınıflarının ortak özelliği toplam örnek sayısının azlığı nedeniyle veri seti içerisinde iyi temsil edilememiş olmalarıdır. Ayrıca SUSP ve SH sınıfları şüpheli olunan örnekleri temsil etmektedir. Bu nedenlerle belirtilen beş sınıf için elde edilen performans değerleri diğer sınıflara göre daha düşük gerçekleşmiştir.

REFERANSLAR

- [1] Z. Alfirevic, D. Devane, G.M. Gyte, A. Cuthbert "Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour," *Cochrane Database Syst Rev*, vol. 2, 2017
- [2] R.A.R.S. Baluz and C.N. dos Santos, "Applying machine learning approaches to assess Cardiotocography exams," in *6th Iberian Conference on Information Systems and Technologies (CISTI 2011)*, 2011, pp. 1-6.
- [3] S.G. Jacob and R.G. Ramani, "Evolving efficient classification rules from cardiotocography data through data mining methods and techniques," in *European Journal of Scientific Research*, 2012, vol. 78, pp. 468-480.
- [4] N. Chamidah, and I. Wasito, "Fetal state classification from cardiotocography based on feature extraction using hybrid K-Means and support vector machine," in *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2015, pp. 37-41
- [5] H. Sahin and A. Subasi, "Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques," *Applied Soft Computing*, vol. 33, pp. 231-238, 2015.
- [6] E. Yılmaz, "Fetal State Assessment from Cardiotocogram Data Using Artificial Neural Networks," *Journal of Medical and Biological Engineering*, vol. 36, pp. 820-832, 2016.
- [7] A. Uzun, E.C. Kızıltas and E. Yılmaz, "Cardiotocography Data Set Classification with Extreme Learning Machine," in *International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES'18)*, 2018
- [8] (2007) UCI machine learning repository. [Online] Available: <https://archive.ics.uci.edu/ml/index.php>
- [9] M.L. Huang and Y.Y. Hsu, "Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network," *Journal of Biomedical Science and Engineering*, vol. 5.09, p. 526, 2012.
- [10] E. Yılmaz and C. Kılıkçier, "Determination of fetal state from cardiotocogram using LS-SVM with particle swarm optimization and binary decision tree," *Computational and mathematical methods in medicine*, 2013.
- [11] H. Ocak and H.M. Ertunc, "Prediction of fetal state from the cardiotocogram recordings using adaptive neuro-fuzzy inference systems," *Neural Computing and Applications*, vol. 23.6, pp. 1583-1589, 2013.
- [12] S. Aladağ, and A. Güven, "Kardiyotokogram Verilerinin Yapay Sinir Ağları ile Sınıflandırılması Classification of Cardiotocogram Data with Using Artificial Neural Networks," in *Tıp Teknolojileri Ulusal Kongresi'14*, 2014
- [13] Z. Cömert, A.F. Kocamaz, S. Güngör, "Cardiotocography signals with artificial neural network and extreme learning machine," in *24th Signal Processing and Communication Application Conference (SIU)*, May 2016, pp. 1493-1496.
- [14] R.S. Kamath and R. K. Kamat, "Modelling Fetal Morphologic Patterns Through Cardiotocography Data: Decision Tree Based Approach," *Journal of Pharmacy Research*, vol.12.1 pp. 9-12, 2017
- [15] D. Ayres de Campos, J. Bernardes, A. Garrido, J. Marques-de-Sa and L. Pereira-Leite, "SisPorto 2.0 :A Program for Automated Analysis of Cardiotocograms," *Journal of Matern Fetal Med*, vol. 9.5, pp.311-318, 2000
- [16] A. Hirose, *Complex-Valued Neural Networks: Theories and Applications*, World Scientific Publishing Co. Pte. Ltd. 5 Toh Tuck Link, Singapore, 2003, pp.1-6.
- [17] H. Gürüler, "A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method," *Neural Computing and Applications*, vol.28.7, pp. 1657-1666, 2017
- [18] M.F. Amin and K. Murase, "Single-layered complex-valued neural network for real-valued classification problems." *Neurocomputing*, vol.72.4-6, pp. 945-955, 2009