

ANNE KARNINDAKİ BEBEĐİN SAĐLIK DURUMU TEŐHİSİ



Elif NEBİOĐLU

TOBB Ekonomi ve Teknoloji Üniversitesi, Mühendislik
Fakültesi Bilgisayar MühendisliĐi

enebioglu@etu.edu.tr

Abstract - Bu doküman içerisinde Bil 470 Makine Öğrenmesi dersi kapsamında seçilen bir veri seti üzerine makine öğrenme algoritmaları uygulanarak kodlanan; verilen ultrason girdisine göre bebeğin sağlık durumunu tahmin eden sınıflandırıcı açıklandı.

Anahtar kelimeler – Machine Learning, classification

I. Giriş

Çocuk ölümlerinin azaltılması, Birleşmiş Milletler'in Sürdürülebilir Kalkınma Hedeflerinin birçoğunda yansıtılmaktadır ve insani ilerlemenin önemli bir göstergesidir.

BM, 2030 yılına kadar ülkelerin yeni doğan ve 5 yaş altı çocukların önlenebilir ölümlerine son vermesini ve tüm ülkelerin 5 yaş altı ölüm oranını 1.000 canlı doğumda en az 25'e düşürmeyi hedeflemesini bekliyor.

Çocuk ölümleri kavramına paralel olarak, hamilelik ve doğum sırasında ve sonrasında (2017 itibariyle) 295 000 ölüme karşılık gelen anne ölüm oranı da elbette ki. Bu ölümlerin büyük çoğunluğu (%94) düşük kaynaklı ortamlarda meydana geldi ve çoğu önlenebilirdi.

Projedeki veri seti, daha sonra üç uzman kadın doğum uzmanı tarafından 3 sınıfa sınıflandırılan Kardiyotokogram muayenelerinden çıkarılan 2126 özellik kaydını içerir:

- Normal
- Şüpheli
- Patolojik

Yukarıda bahsedilenlerin ışığında, Kardiyotokogramlar (CTG'ler), fetal sağlığı değerlendirmek için basit ve uygun maliyetli bir seçenektir ve sağlık profesyonellerinin çocuk ve anne ölümlerini önlemek için harekete geçmesine olanak tanır. Ekipmanın kendisi ultrason darbeleri göndererek ve yanıtını okuyarak çalışır, böylece fetal kalp atış hızı (FHR), fetal hareketler, uterus kasılmaları ve daha fazlasına ışık tutar.

Proje kapsamında cevap aranılan sorular;

1-Bir fetüs için sağlık teşhisini nasıl doğru bir şekilde yapabiliriz?

2- Kardiyotokogram muayenesi sonuçları fetal sağlık sonuçlarını tahmin etmeye yardımcı olabilir mi? Eğer öyleyse, fetal sağlık sonuçlarını tahmin ederken Kardiyotokogram muayenesindeki hangi ölçümler en önemlidir?

3-Fetüsün sağlığını kontrol etmek için maliyetli ve zor testler ve işlemler gerekli midir? Makine öğrenimini kullanarak geçmiş verilerden fetüs sağlığı hakkında sonuçlar çıkarabilir midir?

II. Veri Seti

Veri setinin linki

: <https://www.kaggle.com/andrewmvd/fetal-health-classification>

III. Ön İşleme Aşamaları

Proje, aşağıdaki python makine öğrenimi kitaplıkları kullanılarak uygulandı:

- Linear Algebra - numpy
- Data Processing - pandas
- Visualization - matplotlib, seaborn
- Algorithms - scikit-learn

- Input File – CSV [fetal_health.csv]

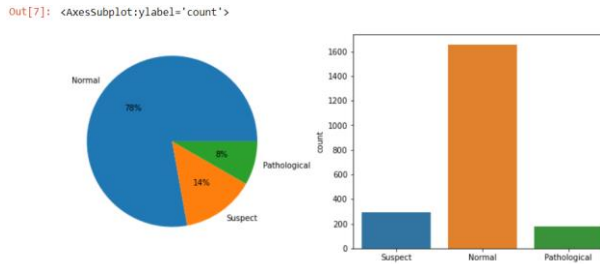
Verileri anlama sürecinde verilerin satır ve sütunları incelendi. Verilerin daha iyi anlaşılması için histogram gibi grafikler çizilip veri seti yorumlandı.

Sütun türlerini, eksik değerleri, sütun-satır sayısını kontrol etme:

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2126 entries, 0 to 2125
Data columns (total 22 columns):
#   Column                                                                 Non-Null Count  Dtype
---  -
0   baseline_value                                                         2126 non-null   float
1   accelerations                                                         2126 non-null   float
2   fetal_movement                                                         2126 non-null   float
3   uterine_contractions                                                  2126 non-null   float
4   light_decelerations                                                   2126 non-null   float
5   severe_decelerations                                                  2126 non-null   float
6   prolonged_decelerations                                               2126 non-null   float
7   abnormal_short_term_variability                                       2126 non-null   float
8   mean_value_of_short_term_variability                                  2126 non-null   float
9   percentage_of_time_with_abnormal_long_term_variability              2126 non-null   float
10  mean_value_of_long_term_variability                                    2126 non-null   float
11  histogram_width                                                        2126 non-null   float
12  histogram_min                                                          2126 non-null   float
13  histogram_max                                                          2126 non-null   float
14  histogram_number_of_peaks                                              2126 non-null   float
15  histogram_number_of_zeroes                                             2126 non-null   float
16  histogram_mode                                                         2126 non-null   float
17  histogram_mean                                                         2126 non-null   float
18  histogram_median                                                       2126 non-null   float
19  histogram_variance                                                     2126 non-null   float
20  histogram_tendency                                                     2126 non-null   float
21  fetal_health                                                            2126 non-null   float
dtypes: float64(22)
memory usage: 365.5 KB
```

Hedef etiket olan 'fetal_health' sütununa bakıldığında fetüslerin büyük çoğunluğunun sağlıklı olduğu görülüyor.



Veri hazırlama aşamasında, modelleme için kullanılacak verileri oluşturan tüm etkinlikler bulunmaktadır. Projenin veri anlama kısmında verilerin çok karmaşık olmadığı gözlemlendi. Sütun başlıkları gibi bazı yerlerde sorunlar tespit edildi. Veri hazırlama aşamasında veriler temizlendi ve veriler makine öğrenmesi yaklaşımlarına uyarlandı.

Veri temizleme:

- Kayıp değerlerin kontrolü

- Verilerdeki null değerleri (varsa) NaN değerlerine çevirme işlemi.

- Tekrarlanan satırların kontrolü
 - Veri setinde tekrarlanan 13 satır olduğu görüldü. Yinelenen satırlar silindikten sonra 2113 satır kalmıştır.

- Sütun başlıklarını yeniden adlandırma

- Veri seti kullanılmadan önce, sütun adları çok uzun olduğu için sütun adları sembollerle değiştirildi.

- Bütün sütunlarda aykırı değer kontrolü

- Nzeros, Mode, Ortalama, Medyan, Varyans gibi sütunlarda aykırı değerler gözlemlendi. Aykırı değeri olan satırlar silindikten sonra veri seti 1843 satıra indirildi.

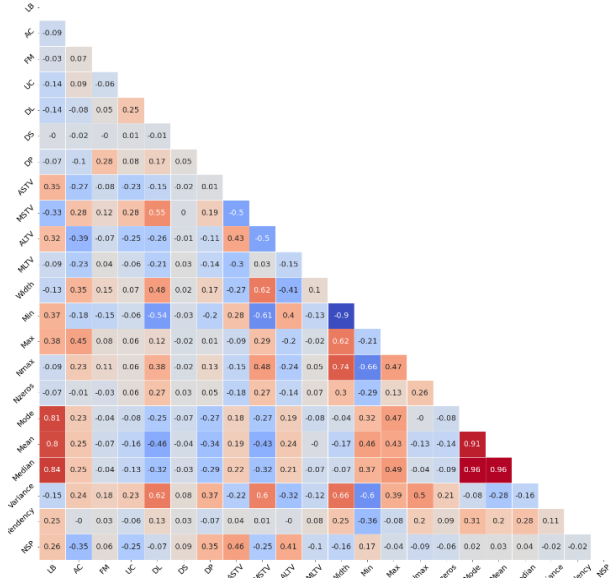
- Korelasyon

Python kütüphanesi Pandas tarafından sağlanan data.corr () ile özellikler ve fetal sağlık durumları arasındaki ilişkiyi kontrol edildi.

Bu değerler arasındaki ilişkiyi incelemek için bir korelasyon haritası çizildi.

Fetal sağlık koşulları ile özellikler arasındaki ilişkiye bakıldığında birçok özelliğin önemsiz olduğu görüldü.

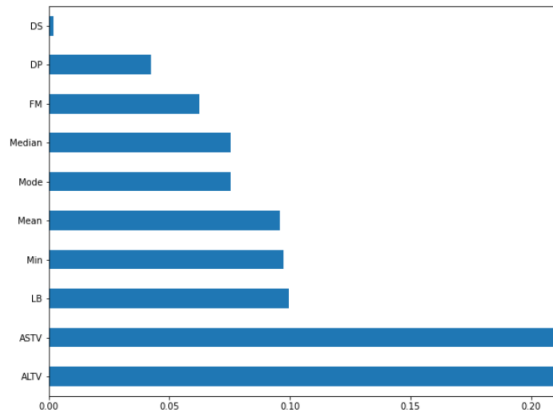
Bu işlemlerden sonra düşük korelasyonlu sütunlar veri setinden çıkarıldı. Veri setinin son durumuna bakıldığında hedef sütunla birlikte toplamda 11 sütun olduğu görüldü.



• Özellik önemi

Oluşturulan bu modelin amacı daha az zaman ve maliyetle fetal sağlığı tespit etmek olduğu için özellik sayısını model için gerekli sayıya indirilmeye çalışıldı. Bunu yapmanın yolu, kalan özelliklerden hangisinin model için daha önemli olduğunu kontrol etmektir. Scikit-learn kütüphanesinde ExtraTreesClassifier ile özelliklerin önemi test edilmeye çalışıldı.

Özellik önemine bakıldığında, DS sütununun model için gerekli olmadığına karar verildi ve bu sütun veri kümesinden kaldırıldı. Veri setinin son versiyonuna bakıldığında, hedef sütun ile birlikte toplamda 10 sütun olduğu görüldü.



• Veri kontrolü

Veri setinin son hali analiz edildiğinde, FM, DP sütunlarının çok fazla 0 içerdiği görüldü. Bu tespiti ardından FM, DP unsurları yüzde olarak incelendi. Bu işlemten sonra DP'nin %94'ünün 0'dan oluştuğu görüldü.

Modelde yanlış sonuçlara sebebiyet vermemesi için bu sütun veri setinden çıkarıldı. Verilerin son haline bakıldığında 22 sütundan oluşan veri seti 9 sütuna inmiş oldu.

- Verilerin eğitim-test olarak ayrılması

Labeled Data	
Train Data	Test Data
%80	%20

- Veriler yukarıda görülen oranda train-test olarak ayrıldı.

- Verilerin normalleştirilmesi

Normalleştirme, özelliklerin [0, 1] aralığına yeniden ölçeklendirilmesini ifade eder. Verileri normalleştirmek için min-max ölçeklendirme uygulanabilir. Projede de verileri normalleştirmek adına MinMax ölçeklendirme kullanılmıştır.

IV. Model Seçimi

Modelleme aşamasında hazırlanan veri seti üzerinde makine öğrenmesi yaklaşımlarını kullanarak tahminler yapıldı. Bu yaklaşımlarda kullanılan veri seti eğitim setidir.

Denetimli Öğrenme

Denetimli bir öğrenme algoritması bilinen bir girdi verisi seti ve verilere bilinen yanıtları alır, ardından yeni verilere yanıt için makul tahminler oluşturmak üzere bir modeli eğitir. Tahmin etmeye çalışılan

çıktı için bilinen veriler varsa denetimli öğrenme kullanılabilir.

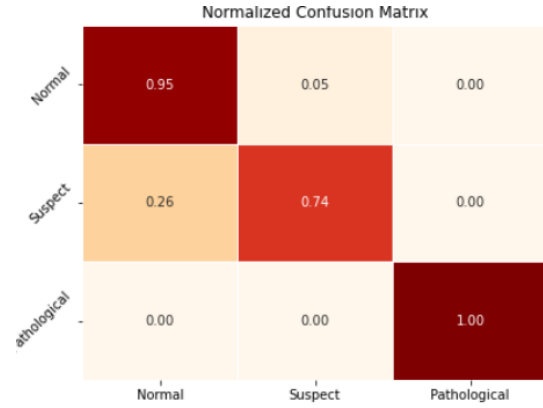
Model seçimi aşamasında birçok denetimli öğrenme tekniği kullanıldı. Bu teknikler Random Forest Classifier, K-Nearest Neighbors, Decision Tree Classifier, Support Vector Machine'tir.

❖ Random Forest Classifier

Algoritma ,birden fazla karar ağacı üretetek sınıflandırma işlemi esnasında sınıflandırma değerini yükseltmeyi hedefler. Random forest algoritması birbirinden bağımsız olarak çalışan birçok karar ağacının bir araya gelerek aralarından en yüksek puan alan değerin seçilmesi işlemidir.

Model Random Forest Classifier ile eğitildiğinde %92,4 doğruluk puanı elde edildi. Karışıklık matrislerine bakıldığında, Normal için 16, Şüpheli için 20 ve Patolojik için 0 yanlış tahmin olduğu görüldü. Random Forest ile Normal için %92, Şüpheli için %92 ve Patolojik için %100 doğruluk oranına ulaşıldığı görüldü.

Accuracy Score : 0.924119				
Classification	Report: precision	recall	f1-score	support
Normal	0.96	0.95	0.95	307
Suspect	0.71	0.74	0.73	50
Pathological	1.00	1.00	1.00	12
accuracy			0.92	369
macro avg	0.89	0.90	0.89	369
weighted avg	0.93	0.92	0.92	369
AUC_ROC Score: 0.9057401170363456				
Accuracy of Normal = 0.9241				
Accuracy of Suspect = 0.9241				
Accuracy of Pathological = 1.0000				

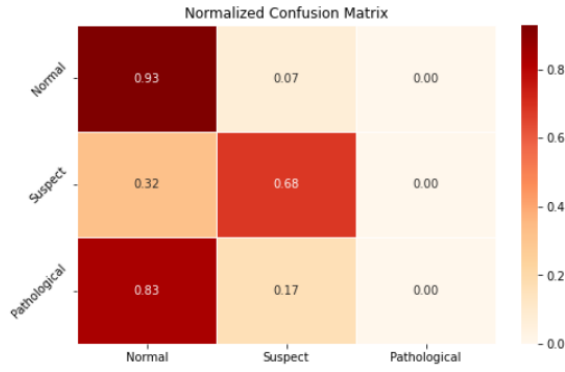


❖ Decision Tree Classifier

Bir karar ağacı, çok sayıda kayıt içeren bir veri kümesini, bir dizi karar kuralları uygulayarak daha küçük kümelerle bölmek için kullanılan bir yapıdır. Yani basit karar verme adımları uygulanarak, büyük miktarlardaki kayıtları, çok küçük kayıt gruplarına bölerek kullanılan bir yapıdır.

Modeli Karar Ağacı Sınıflandırıcısı ile eğitildiğinde %86,4 doğruluk puanı elde edildi. Karışıklık matrislerine baktığımızda, Normal için 23, Şüpheli için 18 ve Patolojik için 12 yanlış tahmin olduğu görüldü. Karar Ağacı Sınıflandırıcısı ile Normalde %88, Şüpheli için %90 ve Patolojik için %97 doğruluk oranına ulaşıldığı görüldü.

Accuracy Score : 0.864499				
Classification	Report: precision	recall	f1-score	support
Normal	0.92	0.93	0.92	307
Suspect	0.59	0.68	0.63	50
Pathological	0.00	0.00	0.00	12
accuracy			0.86	369
macro avg	0.50	0.54	0.52	369
weighted avg	0.84	0.86	0.85	369
AUC_ROC Score: 0.6856248022979065				
Accuracy of Normal = 0.8699				
Accuracy of Suspect = 0.8916				
Accuracy of Pathological = 0.9675				



❖ Support Vector Machines

Bu algoritmada, her bir veri maddesini belirli bir koordinatın değeri olan her özelliğin değeri ile birlikte n-boyutlu boşluğa (burada n sahip olduğunuz özelliklerin sayısı) bir nokta olarak çizilir. Ardından, iki sınıftan oldukça iyi ayırım yapan hiper-düzlemi bularak sınıflandırma gerçekleştirilir.

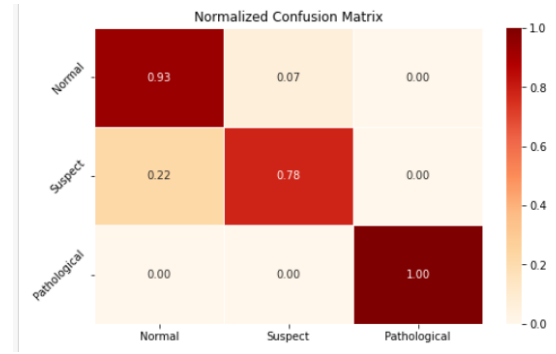
Model Destek Vektör Makinesi ile eğitildiğinde %91,3 doğruluk puanı elde edildi. Karışıklık matrislerine bakıldığında, Normal için 24 yanlış tahmin, Şüpheli için 15 ve Patolojik için 0 olduğu görüldü. Destek Vektör Makinesi ile Normalde %91,5, Şüpheli için %91,5 ve Patolojik için %100 doğruluk oranına ulaşıldığı görüldü.

Accuracy Score : 0.913279

Classification Report:				
	precision	recall	f1-score	support
Normal	0.96	0.93	0.95	307
Suspect	0.66	0.78	0.72	50
Pathological	0.92	1.00	0.96	12
accuracy			0.91	369
macro avg	0.85	0.90	0.87	369
weighted avg	0.92	0.91	0.92	369

AUC_ROC Score: 0.9114466151922388

Accuracy of Normal = 0.9133
Accuracy of Suspect = 0.9160
Accuracy of Pathological = 0.9973



❖ K-Nearest Neighbors

KNN ile temelde yeni noktaya en yakın noktalar aranır. K, bilinmeyen noktanın en yakın komşularının miktarını temsil eder. Sonuçları tahmin etmek için algoritmanın k miktarını (genellikle bir tek sayı) seçilir.

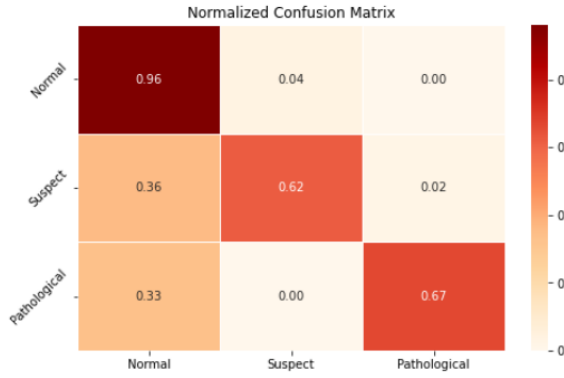
Model K-Nearest Neighbors ile eğitildiğinde %90,5 doğruluk puanı elde edildi. Karışıklık matrislerine bakıldığında, Normal için 15, Şüpheli için 25 ve Patolojik için 4 yanlış tahmin olduğu görüldü. K-Nearest Neighbors ile, Normal için %90, Şüpheli için %91 ve Patolojik için %98 doğruluğa ulaşıldığı görüldü.

Accuracy Score : 0.905149

Classification Report:				
	precision	recall	f1-score	support
Normal	0.93	0.96	0.95	307
Suspect	0.74	0.62	0.67	50
Pathological	0.80	0.67	0.73	12
accuracy			0.91	369
macro avg	0.82	0.75	0.78	369
weighted avg	0.90	0.91	0.90	369

AUC_ROC Score: 0.8087758349315771

Accuracy of Normal = 0.9079
Accuracy of Suspect = 0.9187
Accuracy of Pathological = 0.9837



V. Değerlendirme

Değerlendirme aşaması, oluşturulan modellerin kalitesini değerlendirir ve soruna tam ve uygun şekilde yanıt verip vermediklerini kontrol eder. Bunu yapmak, tahmin edilen model için bir test seti kullanan tablolar ve grafikler gibi diğer çıktılar gibi çeşitli teşhis yöntemlerinin hesaplanmasını gerektirir.

Değerlendirme metrikleri:

- **Train / Test Accuracy :** Elde edilen sonuçlardaki doğru tahminlerin yüzdesidir.
- **Classification Report :** Kesinlik, geri çağırma, f1-skor gibi değerleri değerlendirmek içindir.
- **Confusion Matrix :** Bir sınıflandırma modelinin performansını ve sonucu ne kadar doğru sınıflandırdığını açıklayan tablodur.
- **Roc Curve :** Modelin sınıfları ne kadar ayırt edebildiğini açıklar.

Fetal sağlığı tahmin etmek amacıyla modeli oluşturmak üzere 4 farklı sınıflandırma modeli kullanıldı. Bu modeller hem doğrusal hem de Ağaç tabanlı modelleri içerir. Model oluşturmada parametre ayarının çok

önemli olduğu görüldü ve Grid Search yaparak veriler için en uygun parametreler belirlendi. Cross validation ise bu aşamada yapıldı. Bu modeller yukarıdaki değerlendirme metrikleri ile test edildiğinde en iyi sonucun Random Forest Classifier ile ulaşıldığı görüldü.

Bunun sebebinin Random Forest algoritmasının bir çeşit optimizasyon yapmasından kaynaklanabileceği düşünüldü. (Birçok karar ağacını bir araya getirilmesiyle olan optimizasyon)

Bu çalışma sonucunda cevap aranan problemlerin cevaplarına ulaşıldı. Makine öğrenmesi yardımıyla çok maliyetli testlerden kaçınılarak anne karnındaki bir bebeğin sağlık durumu hakkında bilgi sahibi olunabildiği görüldü.

Modelin doğruluğunu artırmak için birçok yol denendi, ancak düşük fetal sağlık verileri nedeniyle derin öğrenme yöntemlerini denemedi. %92 doğruluk yeterli görünmese de verinin artmasıyla birlikte derin öğrenme yöntemleriyle bu oranı %98-%99'a çıkarmak hedeflendi.

VI. Referanslar

<https://www.kaggle.com/andrewmvd/fetal-health-classification>

<https://medium.com/analytics-vidhya/classification-fetal-health-bd578beae25>

Video Sunum linki:

<https://www.youtube.com/watch?v=wjh2bqyz8q4>