KOCAELİ ÜNİVERSİTESİ

MÜHENDİSLİK FAKÜLTESİ

ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ

MAKİNE ÖĞRENMESİ PROJESİ

MAKİNE ÖĞRENİMİ KULLANILARAK MEME KANSERİ VAKALARININ TAHMİNİ

İLAYDA KIRIMLI

DANIŞMAN: DR. ÖĞR. ÜYESİ AYHAN KÜÇÜKMANİSA

KOCAELİ 2022

# ÖNSÖZ VE TEŞEKKÜR

Araştırmamdaki her aşamada bana yardımcı olan ve eğitim alanında dersleriyle bize vizyon katan çok değerli hocamız Dr. Öğr. Üyesi Ayhan KÜÇÜKMANİSA’ya ve lisasn eğitimim boyunca benden desteklerini esirgemeyen aileme teşekkürlerimi sunarım.

İlayda KIRIMLI

KOCAELİ, 2022

# İÇİNDEKİLER

# ŞEKİLLER DİZİNİ

# TABLOLAR DİZİNİ

# SİMGELER VE KISALTMALAR DİZİNİ

# MAKİNE ÖĞRENİMİ KULLANILARAK MEME KANSERİ VAKALARININ TAHMİNİ

## ÖZET

Bu tez çalışmasının amacı “breast cancer wisconsin (diagnostic) data set” kullanılarak makine öğrenimi ile meme kanserini iyi huylu ya da kötü huylu olarak sınıflandırmaktır. Bu çalışmada kendi algoritmamızı üreterek diğer algoritmalarla karşılaştırma yoluna gidilmiştir.

Bu tez çalışmasında ilk olarak veriseti uygulamaya hazır hale getirilmiştir yani projenin ihtiyacı olmayan parametler temizlenmiştir ve veriseti pythonda işlenebilecek hale getirilmiştir. Daha sonra verisetindeki parametlerin tabloda sütunlara ayrılması için “Pandas” kütüphanesi kullanılmıştır. Sunumda kullanılmak üzere parametreler görselleştirilmiştir. Veriseti test ve eğitim olarak ikiye bölünmüştür. Kanserin iyi veya kötü huylu olarak ayrılması için algoritma geliştirilmiştir ve diğer makine öğrenimi algoritmalarıyla sonuçlar karşılaştırılmıştır. Böylece farklı bir algoritma hesaplaması hedeflenmiştir.

Bu çalışma python dilinde ve Conda ile gerçekleştirilen geliştirme ortamında yapılmıştır. Conda ile oluşturulan pakette başlıca yüklenen paketler; yapayzeka kütüphanesi olarak tensorflow, kütüphanenin kullanımı kolayşatırmak amaçlı keras, dosyalardan veriseti okumak için pandas, n boyutlu matrislerin kullanımını sağlamak için numpy, makine öğrenmesinde sınıflandırma için sckitlearn olarak sıralanabilir.

**Anahtar Kelimeler:** Makine öğrenmesi, veriseti, python,meme kanseri, algoritma

# PREDICTION OF BREAST CANCER CASE USING MACHINE LEARNING

## ABSTRACT

The aim of this thesis is to classify breast cancer as benign or malignant using machine learning using the “breast cancer wisconsin (diagnostic) data set”. In this study, we produced our own algorithm and compared it with other algorithms.

In this thesis, first the dataset was made ready for application, that is, the parameters that were not needed by the project were cleaned and the dataset was made to be processed in python. Then, the "Pandas" library was used to divide the parameters in the dataset into columns in the table. Parameters are visualized to be used in the presentation. The dataset is divided into test and training. An algorithm has been developed to distinguish cancer as benign or malignant, and the results are compared with other machine learning algorithms. Thus, it is aimed to calculate a different algorithm.

This work was done in python language and development environment with Conda. The main installed packages in the package created with conda; tensorflow as an artificial intelligence library, keras to make the library easy to use, pandas to read dataset from files, numpy to enable the use of n-dimensional matrices, sckitlearn for classification in machine learning.

**Keywords:** Machine Learning, dataset, breast cancer, algorithm

# GİRİŞ

Meme kanseri kadınlar arasında ikinci en ölümcül kanser türüdür. Son elli yılda dünyadaki en önemli sağlık problemlerinden birisi haline gelmiş olup etkisi giderek artmaktadır. Erken teşhis, meme kanserinin tedavisinde ve yönetiminde önemli bir yer tutmaktadır. Bilgisayar destekli teşhis sistemleri (CAD), meme kanserinin erken teşhisinde ve meme kanseri sebebiyle gerçekleşen ölüm oranının azalmasında kullanılmak üzere geliştirilen bilgisayar yazılımlarıdır. Erken teşhis tüm kanser türlerinde olduğu gibi meme kanseri tedavisinde çok önemlidir. Bu yazılımlar geliştirilirken görüntü işleme yöntemleri, yapay zeka yöntemleri başta olmak üzere bir çok yöntem kullanılmakta olup insan gözüyle algılanamayan problemli durumların tespit edilmesi amaçlanmaktadır. Dünya Sağlık Örgütü‘ nün (WHO) istatistiklerine göre, 2012 yılında 8.2 milyon insan kanser nedeniyle hayatını kaybetmiştir, bununla beraber kanser ölümlerinin miktarında artış beklenmekte olup önümüzdeki 20 yıl içerisinde kanserden kaynaklı ölümlerin sayısının 14 milyonu bulabileceği tahmin edilmektedir [1]. Tüm kanser türleri arasında meme kanseri kadınlarda görülen ve tüm kanser türleri arasında en çok ölüme sebep olan ikinci kanser türü olarak ön plana çıkmaktadır [2]. Yayınlanan istatistiklere göre geçtiğimiz 50 yıl içerisinde meme kanseri gelişmiş ve gelişmekte olan ülkeler için en büyük sağlık problemi olduğu görülmekte olup son yıllarda artış göstermektedir. Amerika Birleşik Devletleri (ABD, Amerika Birleşik Devletleri)’nde Ulusal Kanser Esntitüsü (NCI, National Cancer Institue) istatistiklerine göre 2014 yılında 40,000 kişinin meme kanseri sebebiyle hayatını kaybedeceği, 232,670 kişinin ise meme kanserine yakalanacağı tahmin edilmektedir. Meme kanseri tüm kanser türleri arasında 2014 tahminlerine göre %14 ile ikinci sırada yer alması yönüyle kanser türleri arasında en çok rastlanan ikinci kanser türüdür. Orta yaş gurubunda olan kadınlar (55-64 yaş aralığı) %61 ile en yüksek risk gurubudur. 2006-2010 verilerine dayanarak yapılan istatistiğe göre meme kanseri sebebiyle ölen kadınların %68‘i yine orta yaş gurubundandır. Gelecekte meme kanserini engellemek için etkili bir yol görülmemektedir, çünkü bu hastalığın kesin sebebi bilinememektedir. Yapılan bir araştırmada yaş ailesel meme kanseri öyküsü, kişisel meme kanseri öyküsü, çocuk doğurma yaşı, menstrual öykü ve beden yapısı faktörleri açısından bir analiz yapılmış olup ailesel kanser öyküsü diğer faktörler arasında diğer faktörlere nazaran daha yüksek risk olarak öne çıkmaktadır [3].

Buna karşın erken safhalarda tedavi edilmesi kadınlar için hastalığın tam olarak tedavi edilmesi açısından büyük bir şans yaratmaktadır. Bu yüzden meme kanserinin erken teşhisi hastalık ve ölüm oranlarının azalmasında büyük rol oynamaktadır.

## Literatür Çalışması

# KAYNAKLAR

[1] http://www.who.int/mediacentre/factsheets/fs297/en/, EriĢim Tarihi:15 Haziran.

[2] ALTHUIS, J., DOIZER, W., ANDERSON, SS., DEVESA, LA., BRINGTON, M., Global trends in breast cancer incidence andmortality 1973–1997, Int.J. Epidemiol,2(34):405-412,2005.

[3] ASLAN, A., GÜRKAN, FE., Kadınlarda meme kanseri risk düzeyi. The Journal of Breast Health, 3(2):63-68, 2007.