

Meme Kanseri Erken Teşhisi için MAMA ve KTB Kullanarak Geliştirilen Model

*Makale Bilgisi / Article Info

Alındı/Received: 14.09.2023

Kabul/Accepted: 07.03.2024

Yayımlandı/Published: 29.04.2024

Enhancing Early Breast Cancer Diagnosis with MRMR and GRU-based Model

Samet AYMAZ* 

*Sağlık Bakanlığı, Trabzon İl Sağlık Müdürlüğü, Sağlık Bilgi Sistemleri Birimi, Trabzon

© Afyon Kocatepe Üniversitesi

Öz

Meme kanseri, dünya genelinde kadınlarda en sık görülen kanser türlerinden biridir ve bu hastalıkta erken teşhis hayat kurtarıcı olabilir. Bu çalışma, Wisconsin Meme Kanseri Teşhisi (WMKT) veri setine odaklanarak meme kanseri teşhisi için doğru ve güvenilir bir model geliştirme amacı gütmektedir. Çalışmada, ilk aşamada Minimum Artıklık Maksimum Alaka Düzeyi (MAMA) yöntemi kullanılarak özellik seçimi yapılmıştır. Yöntem, veri madenciliği ve özellik seçimi alanında etkili bir araç olarak kullanılmaktadır. MAMA ile özelliklerin önem sıralaması yapılarak, sadece anlamlı olanlar kullanılmıştır. Özellik seçimi, modelin karmaşıklığını azaltırken performansı artırır. Daha sonra, MAMA ile seçilen bu özellikler, meme kanseri sınıflandırması için oluşturulan Kapılı Tekrarlayan Birim (KTB) tabanlı bir sinir ağı modeli ile sınıflandırılmaktadır. KTB, tek boyutlu özellik serilerini işleme yeteneğine sahiptir ve karmaşık sınıflandırma problemlerinde etkili sonuçlar verir. Sonuçlar, bu yenilikçi yaklaşımın meme kanseri teşhisinde oldukça başarılı olduğunu göstermektedir. Yapılan değerlendirmelerde doğruluk metriği için %98.28, kesinlik metriği için %98.59, duyarlılık metriği için %98.59, özgüllük metriği için %97.67 ve F-puanı metriği için %98.59 değerleri elde edilmiştir. Sonuçlar yöntemin klinik uygulamalarda uzmanlara yardımcı olabileceğini ortaya koymaktadır. Önerilen yaklaşımın toplumun her kesimi için erişilebilirlik, basit sistemlerde bile hızlı ve yüksek doğrulukla çalışabilmek gibi önemli avantajları olduğu sonuçlardan anlaşılmaktadır.

Anahtar Kelimeler Meme kanseri; MAMA yöntemi; KTB modeli; Özellik seçimi.

Abstract

Breast cancer is one of the most common cancers in women worldwide and early detection can be life-saving. This study aims to develop an accurate and reliable model for breast cancer diagnosis by focusing on the Wisconsin Breast Cancer Diagnosis (WDBC) dataset. In the first stage, feature selection was performed using the Minimum Redundancy Maximum Relevance (MRMR) method. The method is used as an effective tool in the field of data mining and feature selection. With MRMR, the importance of the features is ranked and only the significant ones are used. Feature selection improves performance while reducing the complexity of the model. Then, these features selected by MRMR are classified by a Gated Recurrent Unit (GRU) based neural network model for breast cancer classification. The GRU is capable of handling one-dimensional feature series and gives effective results in complex classification problems. The results show that this innovative approach is highly successful in breast cancer diagnosis. In the evaluations, 98.28% for accuracy metric, 98.59% for precision metric, 98.59% for sensitivity metric, 97.67% for specificity metric and 98.59% for f-score metric were obtained. The results show that the method can help specialists in clinical practice. It is understood from the results that the proposed approach has important advantages such as accessibility for all segments of the society, fast and high accuracy even in simple systems.

Keywords Breast cancer; MRMR method; GRU model; Feature selection.

1. Giriş

Meme kanseri, dünyada kadınlar arasında en sık görülen hastalık olarak öne çıkmaktadır. Bu hastalık erkeklerde de görülebilmektedir fakat kadınlarda görülme olasılığı erkeklerde görülme olasılığından daha fazladır. 2020 yılında dünya çapında 2.3 milyon kadına meme kanseri teşhisi konulmuştur (İnt. Kyn. 1). Bu kadınlar arasında 685.000 ölüm yaşanmıştır (İnt. Kyn. 1). 2020 yılı sonu itibarıyla, son 5 yılda meme kanseri tanısı alan 7.8 milyon

kadın bulunmaktaydı (İnt. Kyn. 1). Bu sayılar meme kanserini dünyanın en yaygın kanseri haline getirmektedir. Meme kanseri dünyanın her ülkesinde ergenlik döneminden sonra her yaşta kadınlarda görülmekte, ancak ilerleyen yaşlarda artan oranlarda görülmektedir. Bu kadar sık rastlanan bu hastalık, erken teşhis sayesinde kolaylıkla tedavi edilebilmektedir. Düzenli mamografi taramaları, meme kanserini erken evrede teşhis etmek için etkili bir yöntemdir. Ayrıca,

kendini muayene etmek ve doktora düzenli olarak gitmek de erken teşhise katkı sağlayabilmektedir. Bilgisayar destekli teşhis (BDT) teknolojisi, mamografi görüntülerinin daha hassas bir şekilde incelenmesine yardımcı olan önemli bir araçtır (Jalalian vd. 2013, Guo vd. 2022, Sahu vd. 2022). Bu araç sayesinde meme kanseri vakalarının erken teşhisi mümkün olurken, tedavi şansı da büyük ölçüde artmaktadır.

BDT, mamografi görüntülerini analiz etmek için bilgisayarları kullanan bir tekniktir. Bu sistemler, meme kanserini tespit etme ve sınıflandırmada radyologlar kadar iyi olabilmektedir. BDT sistemleri, mamografi görüntülerini analiz etmek için çeşitli teknikler kullanır. Bu teknikler, görüntülerdeki kanser özelliklerini belirlemek için matematiksel modeller ve istatistiksel yöntemler kullanır. BDT sistemlerinin etkinliği, test edilen veri kümesine ve kullanılan tekniklere bağlı olarak değişebilir. Bazı çalışmalar, BDT sistemlerinin radyologlar kadar iyi performans gösterdiğini belirtmiştir. Ancak, bu sistemlerin sınırlamaları da mevcuttur (Aswathy vd. 2017, Yusoff vd. 2023). Sınırlamalardan birisi sistemlerin kanseri yanlış pozitif veya yanlış negatif olarak teşhis etmesidir. Bu durum, gereksiz ameliyatlara veya kanser belirtilerinin gözden kaçmasına yol açabilir.

Literatürde WMKT (Wolberg vd. 1995) veri setini kullanan farklı çalışmalar mevcuttur. Bu çalışmalardan birisi, Rasool vd. (2022) geliştirdiği, makine öğrenimi (MÖ) tabanlı tahmini modelleri kullanarak meme kanseri tanısını iyileştirmeyi amaçlayan yöntemdir. Bu yöntem, WMKT (Wolberg vd. 1995) veri kümesini kullanarak eğitilmiş çeşitli MÖ modellerinin performansını değerlendirmektedir. Sonuçlar, MÖ modellerinin radyologlardan kadar iyi performans gösterdiğini kanıtlamıştır. MÖ modelleri, kanseri doğru ve tutarlı bir şekilde tespit edebilmiştir. Diğer bir yayın Bhardwaj vd. (2022) tarafından geliştirilen meme kanseri sınıflandırması için MÖ'nün potansiyelini gösteren bir yaklaşımdır. MÖ modelleri, meme kanserini daha doğru ve tutarlı bir şekilde sınıflandırabildiğinden radyologların performansını artırabilirler. Bu yaklaşımda, meme kanseri sınıflandırması için MÖ algoritmalarının etkinliği gösterilmekle birlikte, bazı sınırlamalarından da bahsedilmektedir. Bu sınırlamalar arasında, MÖ modellerinin veri setine bağımlı olması ve veri ön işleme ve hiperparametre ayarlama gibi tekniklerin kullanımının karmaşık olması yer almaktadır. Yine bu alanda Mohammad vd. (2022), veri madenciliği sınıflandırma ve kümeleme teknikleri kullanılarak WMKT (Wolberg vd. 1995) veri kümesi üzerinde meme kanseri patolojisinin

teşhisi için bir yaklaşım önermişlerdir. Sonuçlar, sınıflandırma algoritmalarının kümeleme algoritmalarından daha iyi performansa sahip olduğunu göstermiştir. Huang vd. (2022), Manifoldların Kararsızlığı için Görselleştirme (MKG) yöntemi ile özellik seçimi yapan, hiyerarşik kümeleme ve rastgele orman algoritması kullanarak meme kanseri sınıflandırması yapan bir yöntem önermektedir. Bu yöntem, WMKT (Wolberg vd. 1995) veri kümesini kullanarak yöntemin etkinliğini değerlendirmektedir. MKG, veri setindeki anormallikleri ve önemli özellikleri tanımlamak için kullanılır. Hiyerarşik kümeleme, verileri benzerliklerine göre kümelere ayırır. Rastgele orman algoritması (ROA) ise, bir dizi karar ağacından (KA) oluşan bir topluluk öğrenme algoritmasıdır ve özelliklerin sınıflandırılmasında kullanılır. Sonuçlar, yöntemin, kanseri doğru ve tutarlı bir şekilde tespit edebildiğini kanıtlamaktadır. Sahoo vd. (2022), Çok Katmanlı Algılayıcı (ÇKA) sinir ağı ve Öğretime dayalı Optimizasyon (ÖDO) kullanarak meme kanserinin saptanması için bir yöntem önermektedir. Bu yöntem, WMKT (Wolberg vd. 1995) veri kümesini kullanarak yöntemin etkinliğini değerlendirmektedir. Yöntem ÇKA ve Öğretme-Öğrenme Tabanlı Optimizasyon (ÖÖTO) yöntemlerinin etkili birleşiminden oluşmaktadır. ÇKA, kanseri tespit etmek için kullanılan bir sinir ağıdır. ÖÖTO ise, parametre optimizasyonu için kullanılan bir meta-sezgisel algoritmadır. Yöntemin sonuçları, oluşturulan yaklaşımın bu alanda yardımcı bir araç olarak kullanılabilirliğini göstermektedir. Suresh vd. (2022), dengesiz tıbbi hastalık verileri kümelerini sınıflandırmak için geliştirilmiş bir Çekişmeli Üretici Ağ (ÇÜA) (Goodfellow vd. 2014) yöntemini önermektedir. ÇÜA, iki ağdan oluşur: bu ağlardan birisi üretken ve diğeri ayırt edici ağdır. Üretici ağ, gerçek verilere benzeyen sentetik veri üretmekten sorumlu iken ayırt edici ağ, gerçek ve sentetik verileri ayırt etmekten sorumludur. Çalışmada, dengesiz tıbbi veri kümesi kullanılmaktadır. Dengesiz tıbbi veri kümeleri, sınıflara ait veriler arasında sayıca çok fazla fark olması anlamına gelmektedir. ÇÜA, WMKT (Wolberg vd. 1995) verilerinde %96.1'lik bir sınıflandırma doğruluğu elde etmiştir. Bu sonuç, geleneksel makine öğrenimi algoritmalarının önemli ölçüde üzerindedir. Yu vd. (2023), Pozitif-Negatif Grafik Kovaryans Sinir Ağı (PN-GKSA) adı verilen yeni bir modelin tanıtımını yapmaktadır. Bu yöntem genellikle bilgi sistemlerini sınıflandırmak için kullanılır. Çalışma, PN-GKSA'nın, Geleneksel Grafik Kovaryans Sinir Ağlarından (GKSA) daha iyi performansa sahip olduğunu göstermiştir. Albadr vd. (2023), meme kanseri tanısı için Hızlı Öğrenme Ağı (HÖA) algoritmasını kullanmakta ve algoritmanın bu alandaki potansiyelini göstermektedir. HÖA, meme kanseri tanısı için

kullanılabilecek önemli bir araçtır. Bu çalışma, meme kanseri tanısı için HÖA algoritmasının etkinliğini göstermekle birlikte, bazı sınırlamaları da vermektedir. Bu sınırlamalar arasında, HÖA algoritmasının veri setine bağımlı olması ve veri ön işleme ve hiperparametre ayarlama gibi tekniklerin kullanımının karmaşık olması yer almaktadır.

Önerilen çalışma, WMKT (Wolberg vd. 1995) veri seti için Minimum Artıklık Maksimum Alaka Düzeyi (MAMA) yöntemi ve Yinelemeli Sinir Ağının (YSA) geliştirilmiş bir modeli olan Kapılı Tekrarlayan Birim (KTB) (Cho vd. 2014) yöntemlerinin etkili birleşiminden oluşmaktadır. Bu yaklaşımda MAMA yöntemi kullanılarak, WMKT (Wolberg vd. 1995) veri setindeki özelliklerin önem oranları ortaya çıkarılmaktadır. Böylece, veri setindeki özelliklerin önem sıralaması yapılarak, modelin daha az gürültüye sahip ve daha anlamlı verilerle çalışmasına olanak tanınır. Bu da makine öğrenimi modellerinin daha iyi performans göstermesini sağlar ve gereksiz özelliklerin eklenmesini önler. Daha sonra önem oranlarına göre düzenlenen veri seti, tarafımızca probleme özgü olarak tasarlanan KTB içeren YSA yapısı ile iyi huylu ve kötü huylu olmak üzere 2 sınıfa ayrılmaktadır. KTB uzun vadeli bağımlılıkları yakalama, gradyan kaybolması sorununu daha az yaşaması, daha hızlı ve daha az kaynak gereksinimi gibi avantajlara sahiptir. Bu yapı kullanılarak, basit sistemlerle etkili kanser teşhisi yapılabilmektedir. Ayrıca, tıbbi verileri daha hassas bir şekilde analiz etme yeteneğine sahiptir ve uzun vadeli bağımlılıkları yakalayıp erken teşhislerin ve daha doğru teşhislerin yapılmasına katkıda bulunabilmektedir. Bu durum, kanserle mücadelede önemli bir rol oynayabilecektir.

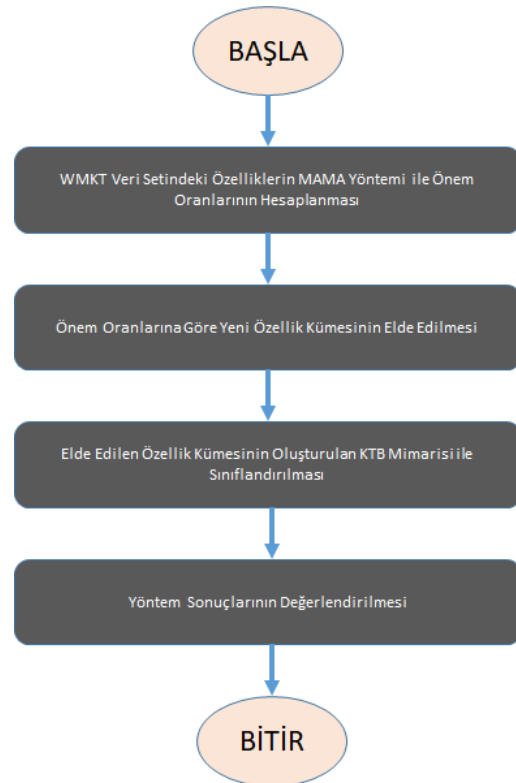
2. Materyal ve Metot

Bu çalışma, özelliklerin önemini vurgulamak ve meme kanseri teşhisini daha etkili hale getirmek amacıyla MAMA (Toğaçar vd. 2020, Ahmed vd. 2020, Billah vd. 2020) yöntemini ve KTB modelini bir araya getiren yenilikçi bir yaklaşım sunmaktadır. Bu yaklaşım, WMKT (Wolberg vd. 1995) veri seti üzerinde uygulanmıştır. Çalışma, meme kanseri teşhisinde özelliklerin önemini vurgulamak ve daha doğru teşhisler elde etmek için birleşik bir yaklaşım sunarak, hastaların erken teşhis ve tedavi alma şansını artırmayı hedeflemektedir. Bu nedenle, çalışmanın meme kanseri teşhisi alanında önemli bir katkı sağlama potansiyeli bulunmaktadır. Önerilen yaklaşıma ait akış tablosu Şekil 1 ile verilmektedir. Çalışmaya ait detaylar alt başlıklarda verilecektir.

2.1 WMKT Veri Seti Özellikleri

WMKT (Wolberg vd. 1995) veri seti, meme kanseri teşhisi ve sınıflandırması üzerine yapılan araştırmaların temel veri kaynaklarından birisidir. Bu veri seti, Wisconsin Üniversitesi Tıp Okulu'ndan Dr. William H. Wolberg tarafından toplanmıştır. Veri seti, meme kanseri teşhisi için kullanılabilecek çeşitli klinik ve hücrel özellikleri içerir. Toplamda 569 örnekten oluşan veri setinde, örneklerin 212'si kötü huylu ve 357'si iyi huylu meme tümörlerine aittir.

WMKT veri seti (Wolberg vd. 1995) her bir örnek için 30 özellikten oluşmaktadır. Bu veri setinin özellikleri, meme tümör hücrelerinin çekirdeklerinin boyutları, hücre şekli düzensizlikleri, çekirdek hücresinin büyüklüğü ve diğer morfolojik özellikleri gibi faktörleri içerir. Ayrıca, hücre çekirdeklerindeki bazı nükleer ve sitoplazmik özellikler de veri setinde yer alır. Her bir özellik, meme kanseri teşhisi için potansiyel olarak önemli bir bilgi kaynağıdır ve bu veri seti, meme kanseri teşhisinde makine öğrenimi algoritmalarının eğitilmesi ve test edilmesi için sıkça kullanılan bir kaynaktır. Veri seti, klinik olarak gerçek verilere dayanması, ileri seviye çok sayıda özellik barındırması ve geniş kullanım alanından dolayı farklı alanlarda tercih edilmektedir. Bu nedenle, hastalıkların daha erken teşhis edilmesi ve tedavi edilmesi konusunda önemli bir rol oynar.



Şekil 1. Önerilen yaklaşımın akış diyagramı

2.2 MAMA Yöntemi ile Verilerin Öneminin Ortaya Koyulması

MAMA (Ding vd. 2005, Toğaçar vd. 2020, Ahmed vd. 2020, Billah vd. 2020), veri madenciliği ve özellik seçimi alanında kullanılan bir yöntemdir. Bu yöntem, bir veri kümesindeki özelliklerin önemini belirlemek ve en fazla bilgi taşıyan özellikleri seçmek için tasarlanmıştır. Aynı zamanda, seçilen özellikler arasındaki benzerliği minimize etmeye çalışmaktadır. Bu yöntem ile en fazla bilgiyi içeren özellikler seçilerek, veri setinin içerdiği bilgiyi en iyi şekilde kullanmaya çalışmaktadır.

MAMA yöntemi, özellik seçimi aşamasında gereksiz ve tekrarlı özellikleri elemek, modelin daha temiz ve anlamlı verilerle çalışmasını sağlamak ve aşırı uydurma (overfitting) sorunlarını azaltmak için etkili bir araçtır. Bu yöntem, özellikle büyük ve karmaşık veri setlerinde kullanıldığında, modelin performansını artırabilmekte ve daha iyi sonuçlar elde etmeye yardımcı olabilmektedir. MAMA yöntemi adım adım aşağıdaki şekilde çalışmaktadır.

MAMA Adım-1 Veri Seti Hazırlanması: İlk adım, veri setinin hazırlanmasıdır. Veri seti içerisinde özellikler (bağımsız değişkenler) ve hedef değişken (sınıf etiketleri) bulunmaktadır.

MAMA Adım-2: Özelliklerin İlgililik Skorlarının Hesaplanması: MAMA, her özelliğin hedef değişken ile ilişkisini ölçen bir ilgililik skoru hesaplamaktadır. İlgililik skoru, özelliğin hedef değişkeni ne kadar iyi açıkladığını göstermektedir. İlgililik skoru denklem 1 (Toğaçar vd. 2020) ile verilmektedir;

$$R(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (1)$$

Denklem 1'de özellik değerleri X ile, hedef değişken değerleri Y ile, X'in ortalama değeri μ_X ile, Y'nin ortalama değeri μ_Y ile, X'in standart sapması σ_X ile Y'nin standart sapması σ_Y ile temsil edilmektedir.

MAMA Adım-3: Ardından, seçilen özellikler arasındaki redundans (artıklık) ölçülmektedir. Bu işlem, seçilen özelliklerin birbirine ne kadar benzediğini belirlemek için yapılmaktadır. Redundansı azaltmak, gereksiz tekrarlılığı önlemeye yardımcı olmaktadır.

MAMA Adım-4 MAMA Skorunun Hesaplanması: İlgililik ve redundans bilgileri kullanılarak, her özellik için bir MAMA skoru hesaplanmaktadır. Hesaplanan skor, bir özelliğin

hem ilgi çekici hem de diğer özelliklerle benzer olmayan bir özelliği temsil etme yeteneğini yansıtır. Skor hesabı denklem 2 (Toğaçar vd. 2020) ile yapılmaktadır.

$$MAMA(X) = R(X, Y) - \frac{1}{K} \sum_{i=1}^K (X, X_i) \quad (2)$$

Denklem 2'de değerlendirilen özellik vektörü X ile hedef değişken değerleri Y ile, diğer özelliklerin sayısı ise K ile temsil edilmektedir.

MAMA Adım-5 Sonuç Veri Seti Oluşturulması: MAMA skorları, özelliklerin sınıflandırma başarısına katkısını göstermektedir. Bu nedenle önerilen çalışmada, orijinal veri setindeki özellik vektörleri ile yöntem sonucu elde edilen önem oranları çarpılarak yeni, daha az gürültüye sahip bir veri seti elde edilir.

2.3 Oluşturulan KTB Mimarisi ile Sınıflandırma

KTB (Cho vd. 2014, Chen vd. 2019, Shadid vd. 2020, Saijad vd. 2020, Shewalkar vd. 2019), bir boyutlu özellik vektörü sınıflandırma için önemli bir YSA türüdür. Özellik sınıflandırma görevlerinde kullanılmasının nedeni, veri setinde oluşabilecek örüntüleri tahmin etme yeteneğine sahip olmasıdır. KTB, veri setindeki örüntü öğrenme yeteneği ile özellikler arasındaki önemli bağımlılıkları yakalayabilmektedir. Standart YSA'lara (Sherstinsky vd. 2020, Kanita vd. 2019, Cho vd. 2014) göre KTB, uzun vadeli bağımlılıkları daha iyi ifade edebilmektedir. Bu durum, daha önceki iterasyonlardaki bilgilere dayalı olarak gelecekteki örüntüleri daha iyi tahmin etmeye yardımcı olmaktadır. Bu çalışmada, bir boyutlu özellik vektörlerinden oluşan WMKT (Wolberg vd. 1995) veri seti kullanıldığından KTB kullanılması oldukça uygundur. Bu çalışma için KTB uygulaması adım adım aşağıdaki gibidir;

KTB Adım-1 Veri Hazırlığı: Çalışma WMKT (Wolberg vd. 1995) kanser verilerinin sınıflandırılmasına odaklanmaktadır. Veri setlerinde her bir özelliğin sınıflandırmaya katkısı farklı olabilmektedir. Bu nedenle çalışmada ilk olarak veri setindeki özelliklerin önem oranları MAMA yöntemi ile belirlenmekte daha sonra orijinal veri seti bu önem oranlarına göre yeni bir veri setine dönüştürülmektedir.

KTB Adım-2 Model Mimarisi Seçimi: Sınıflandırma problemlerine en önemli aşamalardan birisi mimariye karar verme aşamasıdır. Bu çalışmada probleme özgü olarak yeni bir mimari tasarımı oluşturulmuştur. Bu mimari, giriş, KTB katmanı, aktivasyon katmanı, tam-bağlantılı katman ve sınıflandırma katmanından oluşmaktadır. Giriş katmanı, veri setinde her bir örnek 30 özellikten oluştuğundan 30x1'lik bir vektördür. Giriş

katmanından sonra bir KTB katmanı gelmektedir. KTB katmanı, parametre olarak gizli katman sayısı içermektedir. Bu çalışmada gizli katman sayısına Izgara Arama (Nurdian vd. 2016) yaklaşımı ile karar verilmektedir. Bu yaklaşım sonrası ilk KTB katmanı için gizli katman sayısı 15 olarak bulunmuştur. Gizli katman sonrası, bir adet aktivasyon katmanı gelmektedir. Aktivasyon katmanı olarak ReLU (Agarap vd. 2018, Shewalkar vd. 2019) seçilmiştir. ReLU katmanı (Agarap vd. 2018, Shewalkar vd. 2019), gradyan kaybolmasını engelleyerek ağı eğitime katkı sağlamaktadır. Daha sonra Seyreltme katmanı (Nitish vd. 2014, Shewalkar vd. 2019) ile ilk blok tamamlanır ve bu katman ile ağı ezberlemesinin önüne geçilir. Bu katmanlar sonrası bir KTB, bir aktivasyon ve bir seyreltme katmanı (Nitish vd. 2014, Shewalkar vd. 2019) daha gelmektedir. Ağı derinliği denemeler sonucu bulunmuştur. Ağı derinliği arttıkça daha zor verilerin öğrenilmesi sağlanarak eğitim başarısı artırılmaktadır. 2. KTB katmanı için gizli katman sayısı 35 olarak bulunmuştur. Tüm bu katmanlar sonrası iki adet tam-bağlantılı katman gelmektedir. Peş peşe tam-bağlantılı katmanlar, özellikler arasındaki karmaşık ilişkileri ifade ederek sinir ağına doğrusallık olmayan veriler ekler ve daha büyük veri setleri ve karmaşık görevler için daha fazla öğrenme kapasitesi sağlar. Ayrıca bu katmanlar, derinliği artırarak ağı daha fazla bilgiyi modelleyebilmesine yardımcı olmaktadır. Son olarak sınıflandırma katmanı kullanılarak, giriş katmanından verilen özelliğin iyi huylu ya da kötü huylu olarak sınıflandırılması sağlanmaktadır. Oluşturulan KTB Mimarisi detayları Çizelge 1 ile verilir.

Çizelge 1. Oluşturulan KTB mimarisi

Katmanlar	Parametreler
Giriş Katmanı	Özellik Vektörü Boyutu= 30x1
KTB Katmanı 1	Gizli Katman Sayısı=15
Aktivasyon Katmanı	ReLU
Seyreltme Katmanı	0.2
KTB Katmanı 2	Gizli Katman Sayısı=35
Aktivasyon Katmanı	ReLU
Seyreltme Katmanı	0.2
Tam-bağlantılı Katman1	Çıkış Boyutu=190
Tam-bağlantılı Katman 2	Çıkış Boyutu=Sınıf Sayısı
Sınıflandırma Katmanı	

KTB Adım-3: Model eğitimi ve hiperparametrelerin belirlenmesi, makine öğrenimi projelerinde kritik bir öneme sahiptir. Doğru model eğitimi, modelin verileri öğrenmesi ve gelecekteki tahminleri yapabilmesi için

temel bir adımdır. Hiperparametrelerin uygun bir şekilde ayarlanması, modelin performansını optimize etmek için gereklidir. Bu süreçler, modelin başarısını belirlemede ve genelleme yeteneğini artırmada belirleyici bir rol oynar. Çalışmada model eğitimi k-katlamalı çapraz doğrulama kullanılarak gerçekleştirilmektedir. K değeri belirlenirken farklı değerler denenmiş ve en iyi sonucu k=5 değerinin verdiği görülmüştür. Ayrıca, hiperparametreler belirlenirken ızgara arama yaklaşımı kullanılmıştır. Bu yaklaşım sonucu elde edilen parametreler Çizelge 2 ile verilmektedir.

Çizelge 2. Oluşturulan mimari için belirlenen parametreler

Hiperparametreler	Değerler
k-katlama	k=5
Optimizasyon Yöntemi	Adam (Kingma vd. 2014)
Öğrenme Katsayısı	0.001
Maks. Tur Sayısı	20
Min. Parti Boyutu	4

KTB Adım-4 Model Değerlendirmesi: Eğitilen KTB modeli, test verileri üzerinde değerlendirilmektedir. Modelin performansı, sınıflandırma doğruluğu, hassasiyet, duyarlılık, f-puan ölçütlerine göre değerlendirilmektedir. KTB, özellik vektörü sınıflandırma görevlerinde başarılı bir şekilde kullanılabilir güçlü bir modeldir. İlgili verileri analiz etmek ve değişiklikleri yakalamak için bu tür bir modelin seçilmesi, daha doğru sınıflandırmalar yapılmasına yardımcı olmaktadır.

3. Bulgular

Bu çalışma, meme kanserinin bilgisayar destekli teşhisi için yeni bir yaklaşım sunmaktadır. Çalışmanın amacı, sınıflandırma başarısına etkisi az olan özelliklerin baskılanarak oluşturulan özgün KTB mimarisi ile kanserli hücrelerin tespitinin sağlanmasıdır. Bu çalışmanın değerlendirilmesinde karışıklık (Sandra vd. 2019, Lewis vd. 2001) matrisinden hesaplanan doğruluk, kesinlik, duyarlılık, özgüllük ve f-puanı (Roujol vd. 2014, Goutte vd. 2005) metrikleri kullanılmaktadır. Metriklere ait denklemler sırasıyla Denklem 3-7 ile verilmektedir.

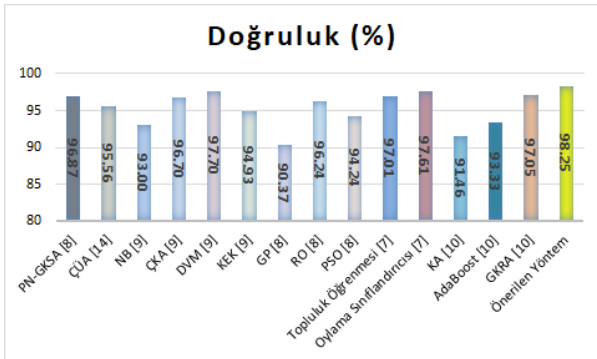
Literatür bu alanda çok fazla çalışma içermektedir. Bu nedenle çalışmada son yıllarda yaygın olarak kullanılan yaklaşımlar ile karşılaştırmalar yapılmaktadır. Bu yöntemler, PN-GKSA (Bhardwaj vd. 2022), ÇÜA (Suresh vd. 2022), Naif Bayes (NB) (Mohammad vd. 2022), ÇKA (Mohammad vd. 2022), Destek Vektör Makineleri (DVM) (Mohammad vd. 2022), K-En Yakın Komşu (KEK) (Mohammad vd. 2022), Genetik Programlama (GP) (Bhardwaj vd. 2022), Rastgele Ormanlar (RO) (Bhardwaj vd. 2022), Parçaçık Sürü Optimizasyonu (PSO) (Bhardwaj

vd. 2022) yöntemi, Topluluk Öğrenmesi (Rasool vd. 2022) yöntemi, Oylama Sınıflandırıcısı Yöntemi (Rasool vd. 2022), Karar Ağaçları (KA) (Huang vd. 2022), AdaBoost (Freund vd. 1995, Huang vd. 2022) yöntemi, Lojistik Regresyon (LR), Çekirdek-Temel Bileşen Analizi (Ç-TBA) (Mohammad vd. 2022), DVM ile Stokastik Gradyan İnişi (SGİ) yöntemleri birleşimi, Doğrusal Ayırma Analizi (DAA) ve Rastgele yöntemlerinin birleşimi ve Gizli Koşullu Rastgele Alan (GKRA) (Huang vd. 2022) yöntemleridir.

Doğruluk metriği, kanser teşhisi gibi önemli bir sınıflandırma görevinde model performansının değerlendirilmesine katkı sağlar. Bu metrik, bir modelin doğru tahminlerinin toplam örnek sayısına oranını yansıtır ve bu oran genellikle yüzde cinsinden ifade edilir. Doğruluk metriği denklem 3 ile hesaplanmaktadır.

$$\text{Doğruluk} = \frac{\text{Doğru Örnek Sayısı}}{\text{Toplam Örnek sayısı}} \times 100 \quad (3)$$

Doğruluk metriği, teşhisin doğruluğunu değerlendirerek, hastaların daha hızlı ve daha doğru bir şekilde tedavi edilmesine yardımcı olur. Bu çalışmada, önerilen yaklaşım literatürdeki farklı yaklaşımlar ile karşılaştırılmıştır. Karşılaştırma sonuçları Şekil 2 ile verilmektedir. En iyi sonuç %98.25 ile önerilen yöntemdedir.



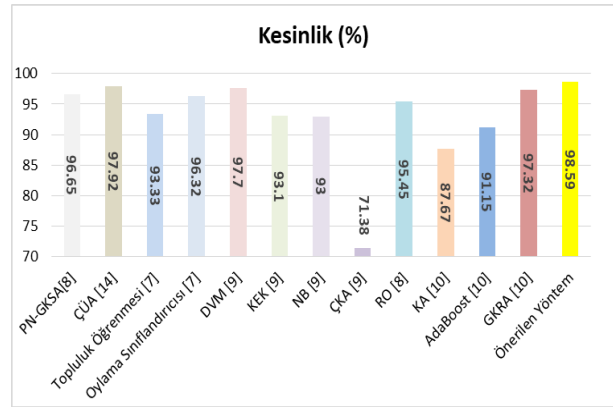
Şekil 2. Test veri seti için doğruluk metriğine göre karşılaştırma sonuçları

Kesinlik değeri, bir yöntemin ne kadar hassas ve kesin sonuçlar üretme yeteneği olduğunu yansıtmaktadır. Metrik, denklem 4 ile hesaplanmaktadır. Denklemde doğru pozitif ve yanlış pozitifler kullanılır. Doğru pozitifler (DP), modelin doğru olarak pozitif tahminlerini, yanlış pozitifler (YP) modelin yanlış olarak pozitif tahminlerini göstermektedir.

$$\text{Kesinlik} = \frac{DP}{DP+YP} \times 100 \quad (4)$$

Metrik sonuçları, yanlış pozitif sonuçların minimize edildiği ve sadece gerçek pozitif sonuçların kabul edildiği anlamına gelmektedir. Bu bağlamda, yüksek kesinlik

değeri, yöntemin veri analizi veya teşhis gibi uygulamalarda gereksiz hataları önlemeye yardımcı olduğunu göstermektedir. Özellikle sağlık alanında, yüksek kesinlik değeri, yanlış teşhislerin ve gereksiz tıbbi müdahalelerin azaltılmasına katkı sağlar. Bu durum, hastaların yanlış teşhis sonucu olumsuz etkilenme riskini düşürerek, yöntemin güvenilirliğini ve hasta güvenini artırır. Bu çalışmada, önerilen yaklaşım literatürdeki farklı yaklaşımlar (Rasool vd. 2022, Bhardwaj vd. 2022, Mohammad vd. 2022, Huang vd. 2022, Suresh vd. 2022) ile karşılaştırılmıştır. Karşılaştırma sonuçları Şekil 3 ile verilmektedir. En iyi sonuç %98.59 ile önerilen yaklaşımdadır.



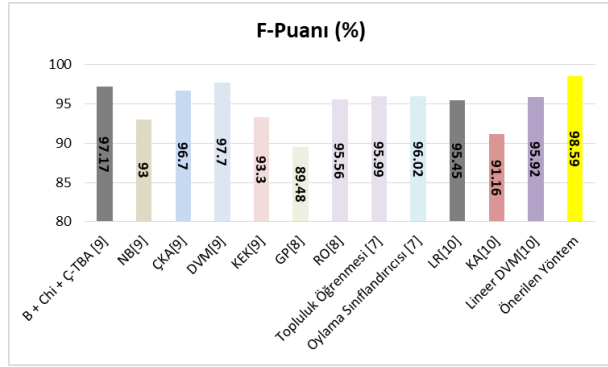
Şekil 3. Test veri seti için kesinlik metriğine göre karşılaştırma sonuçları

F-puanı, bir sınıflandırma modelinin performansını ölçen ve kesinlik ile duyarlılık metriklerinin birleşimini temsil eden bir ölçüm metriğidir. Genellikle bilgisayar bilimleri, veri madenciliği ve sınıflandırma problemlerinin değerlendirilmesinde kullanılmaktadır. Kesinlik Metriği denklem 4 ile verilmektedir. Hassasiyet ve F-puanı metrikleri sırasıyla denklem 5 ve 6 ile verilmektedir. Denklemlerde, yanlış negatifler (YN) yanlış olarak sınıflandırılan negatif sınıfları ifade etmektedir.

$$\text{Hassasiyet} = \frac{DP}{DP+YN} \times 100 \quad (5)$$

$$F - \text{Puanı} = \frac{2 \times \text{Hassasiyet} \times \text{Kesinlik}}{\text{Hassasiyet} + \text{Kesinlik}} \times 100 \quad (6)$$

F-puanı, özellikle hastalıkların erken teşhisi için önemlidir. Erken teşhis, tedaviye daha erken başlama ve hastalığın ilerlemesini önleme şansını artırabilir. Yüksek F-puanına sahip bir teşhis yöntemi, hastalar arasında güven oluşturmaktadır. Bu çalışmada, önerilen yaklaşım literatürdeki farklı yaklaşımlar (Rasool vd. 2022, Bhardwaj vd. 2022, Mohammad vd. 2022, Huang vd. 2022) ile karşılaştırılmıştır. Karşılaştırma sonuçları Şekil 4 ile verilmektedir. En iyi sonuç %98.59 ile önerilen yöntemdedir.



Şekil 4. Test veri seti için f-puanı metriğine göre karşılaştırma sonuçları

Duyarlılık ve özgüllük, sınıflandırma problemlerini ve test değerlendirmesini yaparken birlikte değerlendirildiklerinde, birçok avantaj sunmaktadır. Yüksek duyarlılık, gerçek pozitif sonuçların kaçırılma riskini azaltırken, yüksek özgüllük yanlış pozitif sonuçları minimize eder. Bu iki metrik bir arada kullanıldığında, bir testin veya modelin performansını geniş bir perspektiften değerlendirmemize olanak tanır. Bu da, testin veya modelin hem doğrulukla ilgili hassasiyetini hem de yanlış teşhisleri önleme yeteneğini hesaba katan daha dengeli ve güvenilir sonuçlar elde etmemize yardımcı olur. Duyarlılık metriği denklem 5 ile hesaplanmaktadır. Özgüllük metriği ise denklem 7 ile hesaplanmaktadır. Denklemlerde, doğru negatifler, negatif sınıflara ait doğru tahmin sayısını vermektedir.

$$\text{Özgüllük} = \frac{DN}{DN+YP} \times 100 \quad (7)$$

Bu çalışmada, önerilen yaklaşım literatürdeki farklı yaklaşımlar (Bhardwaj vd. 2022, Mohammad vd. 2022, Huang vd. 2022, Alshayegi vd. 2022) ile karşılaştırılmıştır. Karşılaştırma sonuçları Çizelge 3 ile verilmektedir. En iyi sonuç duyarlılık metriği için %98.59 ve özgüllük metriği için %97.67 sonucu ile önerilen yöntemdedir.

Çizelge 3. Önerilen yöntemin duyarlılık ve özgüllük metriklerine göre karşılaştırma sonuçları

Yöntemler	Metrikler	
	Duyarlılık (%)	Özgüllük (%)
B + Chi + Ç-TBA [9]	97.72	94.23
ÇÜA [14]	93.62	94.52
DAA+Rastgele [30]	95.6	95.7
KA [10]	90	92.33
AdaBoost[10]	91.1	94.61
GKRA [10]	94.77	98.41
ÇKA [9]	72.46	70.62
Önerilen Yöntem	98.59	97.67

Önerilen çalışma, zaman metriği kullanılarak diğer çalışmalarla karşılaştırılmaktadır. Zaman metriği, çalışmaların test veri setini ne kadar sürede

sınıflandırdığını göstermektedir. Karşılaştırma sonuçları Çizelge 4 ile verilmektedir. Çizelge incelendiğinde, zaman metriği açısından en iyi değer 0.05 ile önerilen yöntemle ait olduğu görülmektedir.

Çizelge 4. Önerilen yaklaşım ve literatürdeki yaklaşımlara ait test süreleri

Önerilen Yöntem ve Literatürdeki Yöntemlere Ait Test Süreleri (sn.)	DVM					Önerilen Yöntem
	LR	-SGİ	ÇKA	RO	NB	
Zaman	0.34	0.13	3.10	0.06	0.08	0.05

Sonuç olarak, önerilen yaklaşımın, basit sistemlerde yüksek doğrulukla çalışabilen bir sistem olduğunu vurgulamak önemlidir. Bu yaklaşım, kanser teşhisi sürecini hızlandırırken, düşük maliyetli ve kolayca uygulanabilir bir çözüm sunar. Basit sistemlerde bile yüksek doğruluk sağlaması, bu yöntemin tıbbi bakımdan kaynak sıkıntısı yaşanan bölgelerde veya erişim sorunları olan bölgelerde büyük bir potansiyele sahip olabileceğini göstermektedir. Bu yaklaşım, kanser teşhisinin erken ve etkili bir şekilde yapılmasını kolaylaştırarak hastaların daha iyi sonuçlar elde etmelerini sağlar. Çizelge 5, önerilen yaklaşımın koşullu sisteme ait Merkezi İşlem Birimi (MİB), Rastgele Erişimli Bellek (REB), Depolama Alanı ve İşletim Sistemine ait detayları vermektedir.

Çizelge 5. Oluşturulan çalışmanın test edildiği sisteme ait bilgiler

Sistem Bilgileri	
MİB	Intel i5-10400
REB	16GB
Depolama	256GB SSD
İşletim Sistemi	Windows 10

4. Sonuçlar ve Tartışma

Bu çalışma, meme kanseri teşhisinde yeni bir yaklaşım sunmaktadır. WMKT (Wolberg vd. 1995) veri seti üzerinde uygulanan MAMA yöntemi ile özelliklerin önem dereceleri belirlenmiş ve veri seti daha anlamlı ve etkili hale getirilmiştir. MAMA tarafından sıralanan özellikler, gereksiz karmaşıklığı azaltarak, modelin daha kesin sonuçlar üretmesine katkı sağlamıştır. Daha sonra, bu özelliklerle oluşturulan veri seti, KTB içeren özel bir YSA yapısı ile sınıflandırılmıştır. KTB, uzun vadeli bağımlılıkları yakalama yeteneği, gradyan kaybını azaltma ve daha hızlı işleme gibi avantajlara sahiptir. Bu sayede, meme kanseri iyi huylu ve kötü huylu olarak etkili olarak sınıflandırılabilir. Sonuçlar, önerilen yaklaşımın meme kanseri teşhisinde oldukça başarılı olduğunu göstermektedir. Bu yöntem, tıbbi verileri daha hassas bir

şekilde analiz etme yeteneği sunmaktadır. Bu durum erken teşhislerin ve daha doğru teşhislerin yapılmasına yardımcı olabilir. Böylece hastaların daha erken müdahale ile başarılı tedavi olma olasılığı artar, bu da kanserle mücadelede önemli bir adım olabilir. Sonuç olarak, bu çalışma, MAMA ve KTB yöntemlerin birleşimi ile meme kanseri teşhisinin iyileştirilmesine odaklanmaktadır.

Etik Standartlar Bildirgesi

Yazarlar tüm etik standartlara uyduklarını beyan ederler.

Yazarlık Katkı Beyanı

Yazar: Kaynaklar, Araştırma, Deney, Yazma – orijinal taslak
Görselleştirme, Yazma – orijinal taslak

Çıkar Çatışması Beyanı

Yazarların bu makalenin içeriğiyle ilgili olarak beyan edecekleri hiçbir çıkar çatışması yoktur.

Verilerin Kullanılabilirliği

Bu çalışma sırasında oluşturulan veya analiz edilen tüm veriler, yayınlanan bu makaleye dahil edilmiştir.

5. Kaynaklar

Agarap, A. F. (2018). Deep learning using rectified linear units (relu). ArXiv Preprint ArXiv:1803.08375.

Ahmed, Y. A., Koçer, B., Huda, S., Al-Rimy, B. a. S., & Hassan, M. M. (2020). A system call refinement-based enhanced Minimum Redundancy Maximum Relevance method for ransomware early detection. *Journal of Network and Computer Applications*, 167, 102753.
<https://doi.org/10.1016/j.jnca.2020.102753>

Albadr, M. A. A., Ayob, M., Tiun, S., AL-Dhief, F. T., Arram, A., & Khalaf, S. (2023, April 27). Breast cancer diagnosis using the fast learning network algorithm. *Frontiers in Oncology*, 13.
<https://doi.org/10.3389/fonc.2023.1150840>

Alshayegi, M. H., Ellethy, H., Abed, S., & Gupta, R. (2022, January). Computer-aided detection of breast cancer on the Wisconsin dataset: An artificial neural networks approach. *Biomedical Signal Processing and Control*, 71, 103141.
<https://doi.org/10.1016/j.bspc.2021.103141>

Aswathy, M., & Mani, M. P. (2017). Detection of breast cancer on digital histopathology images: Present status and future possibilities. *Informatics in Medicine Unlocked*, 8, 74–79.
<https://doi.org/10.1016/j.imu.2016.11.001>

Bhardwaj, A., Bhardwaj, H., Sakalle, A., Uddin, Z., Sakalle, M., & Ibrahim, W. (2022). Tree-Based and Machine

learning algorithm analysis for breast cancer classification. *Computational Intelligence and Neuroscience*, 2022, 1–6.

<https://doi.org/10.1155/2022/6715406>

Billah, M., & Waheed, S. (2020). Minimum redundancy maximum relevance (mRMR) based feature selection from endoscopic images for automatic gastrointestinal polyp detection. *Multimedia Tools and Applications*, 79(33–34), 23633–23643.
<https://doi.org/10.1007/s11042-020-09151-7>

Chen, J., Jing, H., Chang, Y., & Liu, Q. (2019). Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process. *Reliability Engineering & System Safety*, 185, 372–382.
<https://doi.org/10.1016/j.res.2019.01.006>

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv (Cornell University).
<https://doi.org/10.48550/arxiv.1406.1078>

Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259.

Ding, C., and H. Peng. "Minimum redundancy feature selection from microarray gene expression data." *Journal of Bioinformatics and Computational Biology*. Vol. 3, Number 2, 2005, pp. 185-205.

Freund, Y., Schapire, R.E. (1995). A decision-theoretic generalization of on-line learning and an application to boosting. In: Vitányi, P. (eds) *Computational Learning Theory*. EuroCOLT 1995. Lecture Notes in Computer Science, vol 904. Springer, Berlin, Heidelberg.
https://doi.org/10.1007/3-540-59119-2_166

Goutte, C., & Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In Springer eBooks (pp. 345–359).
https://doi.org/10.1007/978-3-540-31865-1_25

Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). *Generative Adversarial Nets* (PDF). *Proceedings of the International*

- Conference on Neural Information Processing Systems (NIPS 2014). pp. 2672-2680.
- Guo, Z., Xie, J., Wan, Y., Zhang, M., Qiao, L., Yu, J., Chen, S., Li, B., & Yao, Y. (2022). A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis. *Central European Journal of Biology*, 17(1), 1600–1611.
<https://doi.org/10.1515/biol-2022-0517>
- Huang, Z., & Chen, D. (2022). A breast cancer diagnosis method based on VIM feature selection and hierarchical clustering Random Forest algorithm. *IEEE Access*.
<https://doi.org/10.1109/access.2021.3139595>
- Jalalian, A., Mashohor, S., Mahmud, R., Saripan, M. I., Ramli, A. R., & Karasfi, B. (2013). Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. *Clinical Imaging*, 37(3), 420–426.
<https://doi.org/10.1016/j.clinimag.2012.09.024>
- Karita, S., Chen, N., Hayashi, T., Hori, T., Inaguma, H., Jiang, Z., Someki, M., Soplein, N. E. Y., Yamamoto, R., Wang, X., Watanabe, S., Yoshimura, T., & Zhang, W. (2019). A Comparative Study on Transformer vs RNN in Speech Applications. 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU).
<https://doi.org/10.1109/asru46091.2019.9003750>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Lewis, H. G., & Brown, M. (2001). A generalized confusion matrix for assessing area estimates from remotely sensed data. *International Journal of Remote Sensing*, 22(16), 3223–3235.
<https://doi.org/10.1080/01431160152558332>
- Mohammad, W. T., Teete, R., Al-Aaraj, H., Rubbai, Y., & Arabyat, M. M. (2022b). Diagnosis of Breast Cancer Pathology on the Wisconsin Dataset with the Help of Data Mining Classification and Clustering Techniques. *Applied Bionics and Biomechanics*, 2022, 1–9.
<https://doi.org/10.1155/2022/6187275>
- Nitish, S., Geoffrey, H., Alex, K., Ilya, S., and Ruslan, S. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1 (January 2014), 1929-1958.
- Nurdian, S. W., Adu, N., Palupi, I. R., & Raharjo, W. (2016). Comparison tomography relocation hypocenter grid search and guided grid search method in Java island. *Journal of Physics: Conference Series*, 776, 012113.
<https://doi.org/10.1088/1742-6596/776/1/012113>
- Rasool, A., Bunternghit, C., Luo, T., Islam, M. R., Qu, Q., & Jiang, Q. (2022). Improved Machine Learning-Based Predictive Models for breast cancer diagnosis. *International Journal of Environmental Research and Public Health*, 19(6), 3211.
<https://doi.org/10.3390/ijerph19063211>
- Roujol, S., Weingärtner, S., Foppa, M., Chow, K., Kawaji, K., Ngo, L., Kellman, P., Manning, W. J., Thompson, R. B., & Nezafat, R. (2014). Accuracy, Precision, and reproducibility of four T1 mapping sequences: A Head-to-Head comparison of MOLLI, ShMOLLI, SASHA, and SAPPHERE. *Radiology*, 272(3), 683–689.
<https://doi.org/10.1148/radiol.14140296>
- Sahoo, S. K., Saha, A. K., Ezugwu, A. E., Agushaka, J. O., Abuhaija, B., Alsoud, A. R., & Abualigah, L. (2022). Moth Flame Optimization: Theory, modifications, hybridizations, and applications. *Archives of Computational Methods in Engineering*, 30(1), 391–426.
<https://doi.org/10.1007/s11831-022-09801-z>
- Sahu, Y., Tripathi, A., Gupta, R. K., Gautam, P., Pateriya, R. K., & Gupta, A. (2022). A CNN-SVM based computer aided diagnosis of breast Cancer using histogram K-means segmentation technique. *Multimedia Tools and Applications*, 82(9), 14055–14075.
<https://doi.org/10.1007/s11042-022-13807-x>
- Sajjad, M., Khan, Z. A., Ullah, A., Hussain, T., Ullah, W., Lee, M. Y., & Baik, S. W. (2020). A novel CNN-GRU-Based hybrid approach for Short-Term residential load Forecasting. *IEEE Access*, 8, 143759–143768.
<https://doi.org/10.1109/access.2020.3009537>
- Sendra, A. L., Carrasco, A., Martín, A. J., & De Las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 216–231.
<https://doi.org/10.1016/j.patcog.2019.02.023>
- Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos Solitons & Fractals*, 140, 110212.
<https://doi.org/10.1016/j.chaos.2020.110212>
- Shewalkar, A. N., Nyavanandi, D., & Ludwig, S. A. (2019). Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU.

Journal of Artificial Intelligence and Soft Computing Research, 9(4), 235–245.

<https://doi.org/10.2478/jaiscr-2019-0006>

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306.

<https://doi.org/10.1016/j.physd.2019.132306>

Suresh, T., Brijet, Z., & Subha, T. D. (2022, November 2). Imbalanced medical disease dataset classification using enhanced generative adversarial network. *Computer Methods in Biomechanics and Biomedical Engineering*, 1–17.

<https://doi.org/10.1080/10255842.2022.2134729>

Toğaçar, M., Ergen, B., & Cömert, Z. (2020). Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks. *Biocybernetics and Biomedical Engineering*, 40(1), 23–39.

<https://doi.org/10.1016/j.bbe.2019.11.004>

Wolberg, W., Mangasarian, Olvi, Street, Nick, and Street, W.. (1995). Breast Cancer Wisconsin (Diagnostic). UCI Machine Learning Repository. <https://doi.org/10.24432/C5DW2B>.

Yusoff, M., Haryanto, T., Suhartanto, H., Mustafa, W. A., Zain, J. M., & Kusmardi, K. (2023). Accuracy Analysis of deep learning methods in breast Cancer Classification: A Structured review. *Diagnostics*, 13(4), 683.

<https://doi.org/10.3390/diagnostics13040683>

Yu, B., Xie, H., & Xu, Z. (2023b). PN-GCN: Positive-negative graph convolution neural network in information system to classification. *Information Sciences*, 632, 411–423.

<https://doi.org/10.1016/j.ins.2023.03.013>

İnternet kaynakları

1. Breast Cancer Dataset, <https://www.who.int/news-room/fact-sheets/detail/breast-cancer> (12.07.2023.).