

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE
SPEKTRUM KULLANIMININ TAHMİNİ

Mehmet Oğuz KELEK

YÜKSEK LİSANS TEZİ
Elektronik ve Haberleşme Mühendisliği Anabilim Dalı
Haberleşme Programı

Danışman
Prof. Dr. M. S. Ufuk TÜRELİ

Mayıs, 2024

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE SPEKTRUM
KULLANIMININ TAHMİNİ

Mehmet Oğuz KELEK tarafından hazırlanan tez çalışması 07.05.2024 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Mühendisliği Anabilim Dalı Haberleşme Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Prof. Dr. M. S. Ufuk TÜRELİ
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Prof. Dr. M. S. Ufuk TÜRELİ, Danışman
Yıldız Teknik Üniversitesi

Doç. Dr. Bülent BOLAT, Üye
Yıldız Teknik Üniversitesi

Dr. Öğr. Üyesi Nazlı TOKATLI, Üye
İstanbul Sağlık ve Teknoloji Üniversitesi

Danışmanım Prof. Dr. M. S. Ufuk TÜRELİ sorumluluğunda tarafımca hazırlanan Makine Öğrenmesi Algoritmaları ile Spektrum Kullanımının Tahmini başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Mehmet Oğuz KELEK

İmza

*Anneme
ve Aileme*



TEŐEKKÜR

Yüksek Lisans eğitimin ve Tez çalışmam boyunca bana yol gösteren, süreç hakkında bilgilendirme ve takip konularında aktif destek sağlayan değerli hocam Prof. Dr. M. S. Ufuk TÜRELİ'ye derin minnettarlığımı ifade etmek isterim.

Aynı zamanda, hayatım boyunca maddi ve manevi desteklerini esirgemeyen, fedakarlıklarıyla her zaman yanımda olan annem Kadriye KELEK'e en içten teşekkürlerimi sunarım.

Mehmet Oğuz KELEK

İÇİNDEKİLER

SİMGE LİSTESİ	vi
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ	ix
TABLO LİSTESİ	xi
ÖZET	xii
ABSTRACT	xiv
1 GİRİŞ	1
1.1 Literatür Özeti	1
1.2 Tezin Amacı ve Önemi	3
1.3 Hipotez	4
2 YÖNTEM	5
2.1 Destek Vektör Makineleri	6
2.2 Rassal Orman Algoritması	7
2.3 Uzun Kısa Süreli Bellek (LSTM)	8
3 UYGULAMA	10
3.1 Veri Seti	10
3.2 Destek Vektör Makinesi	13
3.3 Rassal Orman Algoritması	23
3.4 Uzun Kısa Süreli Bellek (LSTM)	31
4 SONUÇ	42
KAYNAKÇA	44
TEZDEN ÜRETİLMİŞ YAYINLAR	46

SİMGE LİSTESİ

W_C	Aday Değerlerin Ağırlık Matrisi
B	Bant Genişliği
b	Bias Terimi
k	Boltzmann Sabiti
b_o	Çıkış Kapısının Bias Vektörü
γ	Gama
TN	Gerçek Negatif
TP	Gerçek Pozitif
TPR	Gerçek Pozitif Oranı
W_i	Giriş Kapısının Ağırlık Matrisi
b_i	Giriş Kapısının Bias Vektörü
i_t	Giriş Kapısının Çıktısı
N	Gürültü Gücü
W	Hiper Düzlemin Normal Vektörü
\tilde{C}_t	Hücre Durumu İçin Aday Değerler
K	Kernel Değeri
\log	Logaritma Fonksiyonu
h_t	Mevcut Hücrenin Çıktısı
C_t	Mevcut Hücre Durumu
x_t	Mevcut Zaman Adımının Girdisi
T	Mutlak Sıcaklık (Kelvin)
C_{t-1}	Önceki Mevcut Hücre Durumu
h_{t-1}	Önceki Zaman Adımının Hücre Çıktısı

x	Özellik Vektörü
σ	Sigmoid Aktivasyon Fonksiyonu
SNR_{dB}	Sinyal Gürültü Oranının Desibel Cinsinden Değeri
W_f	Unutma Kapısının Ağırlık Matrisi
b_f	Unutma Kapısının Bias Vektörü
f_t	Unutma Kapısının Çıktısı
FN	Yanlış Negatif
FP	Yanlış Pozitif
FPR	Yanlış Pozitif Oranı



KISALTMA LİSTESİ

AUC	Eğrinin Altındaki Alan
C	Marjinal Hata Toleransı
CR	Bilişsel Radyo
dB	Desibel
dBm	Desibel Miliwatt
DL	Derin Öğrenme
FPR	Yanlış Pozitif Oranı
GHz	Giga Hertz
LR	Lineer Regresyon
LSTM	Uzun-Kısa Süreli Bellek
LTE	Uzun Süreli Gelişim
MHz	Mega Hertz
ML	Makine Öğrenmesi
MSE	Ortalama Hata Karesi
RBF	Radyal Temel Fonksiyon
RF	Rassal Orman Algoritması
RNN	Yenilemeli Sinir Ağı
ROC	Alıcı İşlem Karakteristikleri
SDR	Yazılım Tabanlı Radyo
SNR	Sinyal Gürültü Oranı
SVM	Destek Vektör Makinesi
TPR	Gerçek Pozitif Oranı

ŞEKİL LİSTESİ

Şekil 1.1	Spektrum kullanımı[2]	1
Şekil 2.1	SVM algoritmasının sınıflar arasındaki marjı	6
Şekil 2.2	LSTM algortimasının yapısı	8
Şekil 3.1	Veri setinin 800 MHz frekansındaki ölçümleri	12
Şekil 3.2	Veri setinin 838 MHz frekansındaki ölçümleri	12
Şekil 3.3	Destek vektör makinesi 800 MHz’de C parametresi	13
Şekil 3.4	Destek vektör makinesi 838 MHz’de C parametresi	14
Şekil 3.5	800 MHz’de destek vektör makinesi test tahminleri	14
Şekil 3.6	800 MHz’de destek vektör makinesi doğrulama tahminleri	14
Şekil 3.7	800 MHz’de destek vektör makinesi test verisi hata matrisi	15
Şekil 3.8	800 MHz’de destek vektör makinesi doğrulama verisi hata matrisi	16
Şekil 3.9	800 MHz’de destek vektör makinesi test verisi ROC eğrisi	17
Şekil 3.10	800 MHz’de destek vektör makinesi doğrulama verisi ROC eğrisi	18
Şekil 3.11	838 MHz’de destek vektör makinesi test tahminleri	19
Şekil 3.12	838 MHz’de destek vektör makinesi doğrulama tahminleri	19
Şekil 3.13	838 MHz’de destek vektör makinesi test verisi hata matrisi	20
Şekil 3.14	838 MHz’de destek vektör makinesi doğrulama verisi hata matrisi	21
Şekil 3.15	838 MHz’de destek vektör makinesi test verisi ROC eğrisi	21
Şekil 3.16	838 MHz’de destek vektör makinesi doğrulama verisi ROC eğrisi	22
Şekil 3.17	Rassal orman algoritması 800 MHz’de karar ağacı sayısı	23
Şekil 3.18	Rassal orman algoritması 838 MHz’de karar ağacı sayısı	23
Şekil 3.19	800 MHz’de rassal orman algoritması test tahminleri	24
Şekil 3.20	800 MHz’de rassal orman algoritması doğrulama tahminleri	24
Şekil 3.21	800 MHz’de rassal orman algoritması test verisi hata matrisi	25
Şekil 3.22	800 MHz’de rassal orman algoritması doğrulama verisi hata matrisi	25
Şekil 3.23	800 MHz’de rassal orman algoritması test verisi ROC eğrisi	26
Şekil 3.24	800 MHz’de rassal orman algoritması doğrulama verisi ROC eğrisi	27
Şekil 3.25	838 MHz’de rassal orman algoritması test tahminleri	28
Şekil 3.26	838 MHz’de rassal orman algoritması doğrulama verisi tahminleri	28
Şekil 3.27	838 MHz’de rassal orman algoritması test verisi hata matrisi	29
Şekil 3.28	838 MHz’de rassal orman algoritması doğrulama verisi hata matrisi	30

Şekil 3.29	838 MHz'de rassal orman algoritması test verisi ROC eğrisi	30
Şekil 3.30	838 MHz'de rassal orman algoritması doğrulama verisi ROC eğrisi	31
Şekil 3.31	800 MHz'de LSTM algoritması test tahminleri	32
Şekil 3.32	800 MHz'de LSTM algoritması doğrulama tahminleri	32
Şekil 3.33	800 MHz'de LSTM algoritması test verisi hata matrisi	33
Şekil 3.34	800 MHz'de LSTM algoritması doğrulama verisi hata matrisi	34
Şekil 3.35	800 MHz'de LSTM algoritması test verisi ROC eğrisi	34
Şekil 3.36	800 MHz'de LSTM algoritması doğrulama verisi ROC eğrisi	35
Şekil 3.37	800 MHz'de LSTM algoritması kayıp grafiği	35
Şekil 3.38	800 MHz'de LSTM tahminleri	36
Şekil 3.39	838 MHz'de LSTM algoritması doğrulama verisi tahminleri	37
Şekil 3.40	838 MHz'de LSTM algoritması test tahminleri	37
Şekil 3.41	838 MHz'de LSTM algoritması test verisi hata matrisi	38
Şekil 3.42	838 MHz'de LSTM algoritması doğrulama verisi hata matrisi	38
Şekil 3.43	838 MHz'de LSTM algoritması test verisi ROC eğrisi	39
Şekil 3.44	838 MHz'de LSTM algoritması doğrulama verisi ROC eğrisi	40
Şekil 3.45	838 MHz'de LSTM algoritması kayıp grafiği	40
Şekil 3.46	838 MHz'de LSTM tahminleri	41
Şekil 4.1	800 MHz'de algoritmaların başarımlarını parametreleri özeti	43
Şekil 4.2	838 MHz'de algoritmaların başarımlarını parametreleri özeti	43

TABLO LİSTESİ

Tablo 3.1	Kullanılan kanalların karakteri	11
Tablo 3.2	800 MHz'de SVM algoritma test verisi raporu	18
Tablo 3.3	800 MHz'de SVM algoritma doğrulama verisi raporu	19
Tablo 3.4	838 MHz'de SVM algoritma doğrulama verisi raporu	22
Tablo 3.5	838 MHz'de SVM algoritma test verisi raporu	23
Tablo 3.6	800 MHz'de rassal orman algoritma test verisi raporu	27
Tablo 3.7	800 MHz'de rassal orman algoritma doğrulama verisi raporu	27
Tablo 3.8	838 MHz'de rassal orman algoritma test verisi raporu	29
Tablo 3.9	838 MHz'de rassal orman algoritma doğrulama verisi raporu	31
Tablo 3.10	800 MHz'de LSTM algoritma test verisi raporu	36
Tablo 3.11	800 MHz'de LSTM algoritma doğrulama verisi raporu	36
Tablo 3.12	838 MHz'de LSTM algoritma test verisi raporu	41
Tablo 3.13	838 MHz'de LSTM algoritma doğrulama verisi raporu	41
Tablo 4.1	800 MHz'de algoritmaların doğrulama veri seti için doğruluk ve AUC değerleri	42
Tablo 4.2	838 MHz'de algoritmaların doğrulama veri seti için doğruluk ve AUC değerleri	43

Makine Öğrenmesi Algoritmaları ile Spektrum Kullanımının Tahmini

Mehmet Oğuz KELEK

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Prof. Dr. M. S. Ufuk TÜRELİ

Günümüzde kablosuz iletişimin artmasıyla birlikte spektrum yoğunluğu da artmaktadır. Spektrum yoğunluğunun artması aynı zamanda spektrum kıtlığına da yol açmaktadır. Ayrıca askeri alanlarda yasa dışı iletişimin tespitinde spektrum kullanımının tahmin edilmesi de önemli bir konu haline gelmiştir. Bu çalışma, Medipol Üniversitesi tarafından toplanan geniş bir veri seti üzerinde makine öğrenmesi algoritmaları uygulanarak gerçekleştirilmiştir. Çalışmanın odak noktası, belirli bir kablosuz kanalın gelecekteki güç düzeyini tahmin etmek ve bu tahminleri kullanarak dinamik spektrum erişim stratejilerini optimize etmektir. Medipol Üniversitesi'nin Agilent EXA N9010A spektrum analizörü ile yaptığı bu ölçümler, 21 saat 38 dakika 3 saniye boyunca saniyede bir frekansta sürekli olarak gerçekleştirildi ve toplam 1000 farklı frekans değeri incelendi. Kullanılan veri kümesinin %60'ı algoritmaların eğitimi için, %20 test için ve %20 algoritmaların doğrulanması için ayrılmıştır. Yoğun olarak kullanılan ve daha az kullanılan kanallarda spektrum kullanımını tahmin etmek için üç farklı algoritma kullanıldı: Rastgele Orman algoritması, Destek Vektör Makinesi algoritması ve Uzun-Kısa Süreli Bellek (LSTM) algoritması. Kullanılan algoritmalar kanalın geçmiş bilgilerini kullanarak sonuçlar üretti. Kullanılan algoritmalar birbirleriyle karşılaştırılmış ve spektrum kullanımını tahmin etmek için en ideal çözüm aranmıştır. Her üç algoritma da yoğun olarak kullanılan kanalda kötü sonuçlar vermiştir. Daha az yoğun olarak kullanılan kanalda ise Destek Vektör Makinesi ve Uzun-Kısa Süreli Bellek algoritmaları benzer sonuçlar vermiştir. Ancak LSTM algoritmasının ayırt edici özelliği biraz daha yüksekti.

Anahtar Kelimeler: Spektrum analizi, uzun-kısa süreli bellek, destek vektör makinesi, rassal orman algoritması



**YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

Estimation of Spectrum Usage with Machine Learning Algorithms

Mehmet Oğuz KELEK

Electronics and Communication Engineering
Master Thesis

Supervisor: Prof. Dr. M. S. Ufuk TÜRELİ

Nowadays, with the increase in wireless communication, spectrum density is also increasing. Increasing spectrum density also leads to spectrum scarcity. In addition, estimating spectrum usage in detecting illegal communications in military areas has become an important issue. This study was carried out by applying machine learning algorithms on a large data set collected by Medipol University. The focus of the work is to predict the future power level of a given wireless channel and optimize dynamic spectrum access strategies using these predictions. These measurements, made with Medipol University's Agilent EXA N9010A spectrum analyzer, were carried out continuously at a frequency of one per second for 21 hours 38 minutes and 3 seconds, and a total of 1000 different frequency values were examined. %60 of the dataset used is reserved for training the algorithms, %20 for testing, and %20 for validating the algorithms. Three different algorithms were used to estimate spectrum usage on heavily used and less used channels: Random Forest algorithm, Support Vector Machine algorithm, and Long-Short Term Memory (LSTM) algorithm. The algorithms used produced results using the channel's historical information. The algorithms used were compared with each other and the most ideal solution was sought to estimate spectrum usage. All three algorithms gave poor results on the heavily used channel. In the less intensively used channel, Support Vector Machine and Long-Short Term Memory algorithms gave similar results. However, the distinctiveness of the LSTM algorithm was slightly higher.

Keywords: Spectrum analysis, long-short term memory, support vector machine,

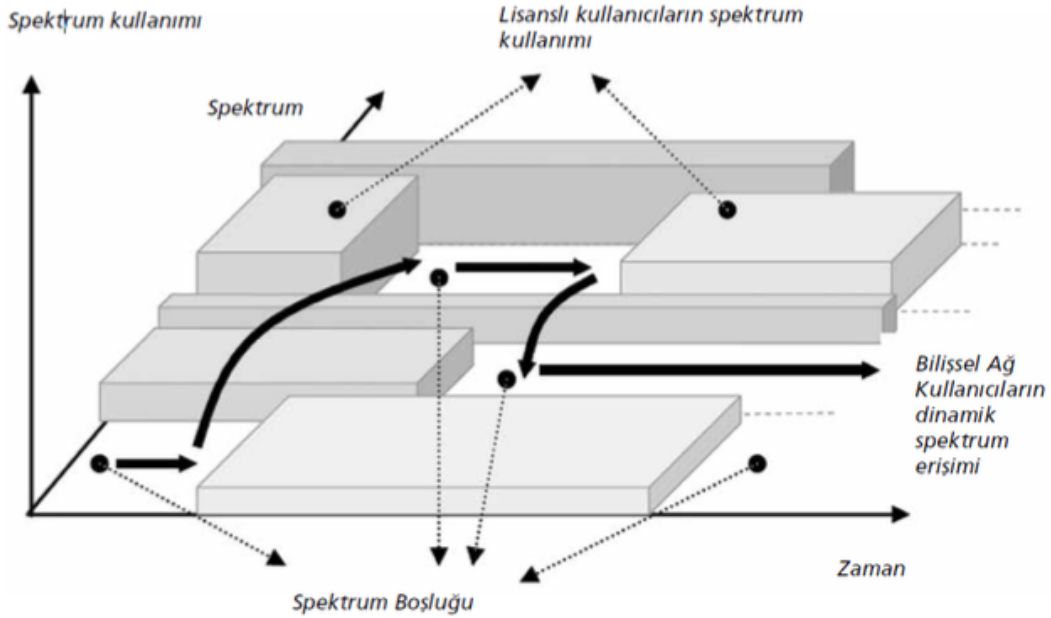
random forest algorithm



**YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING**

1.1 Literatür Özeti

Kablosuz iletişim teknolojilerinin yaygın kullanımı ile radyo frekans spektrumunda yoğunluk giderek artmaktadır. Spektrumun yoğun kullanılması zamanla spektrumun yetersiz hale gelmesine neden olur. Şekil 1.1, lisanslı kullanıcıların zaman içindeki spektrum kullanım değişiklikleri ve lisanssız kullanıcıların dinamik spektrum erişimi gösterilmektedir. Spektrum yoğunluğunun analizi temelinde yapılan çalışmalar, spektrum yetersizliğinin çok sayıda kullanıcıdan ziyade spektrumun verimsiz kullanılmasından kaynaklandığını ortaya koymuştur. Özellikle kentsel alanlarda lisanslı kullanıcıların bazı kanalları yeterince kullanmadığı gözlemlenmiştir [1].



Şekil 1.1 Spektrum kullanımı[2]

Kablosuz iletişimde spektrum kıtlığı sorununa çözüm olarak "akıllı spektrum" olarak bilinen dinamik spektrum kullanım çerçevesinin geliştirilmesi ve

uygulanması tartışılmaktadır. Gelişmiş ölçüm, modelleme, veritabanı yönetimi ve dinamik spektrum tahsisi vurgulanarak, verimli spektrum yönetiminin önemine dikkat çekilmektedir. Büyük ölçekli ölçüm cihazlarından elde edilen verilerin kullanılması ve istatistiksel modelleme uygulanmasıyla, radyo ortamına ve kullanıcı taleplerine uyum sağlayan ve spektrum verimliliğini önemli ölçüde artırmayı amaçlayan bir sistem önerilmektedir. Bu yaklaşım, sabit spektrum tahsisinden dinamik, çevre bilincine sahip spektrum kullanımına doğru bir değişikliği ve daha sürdürülebilir bir kablosuz iletişim altyapısı için uzun vadeli bir stratejiyi vurgulamaktadır [3].

Araştırma, bilişsel radyo (CR) ağlarındaki spektrum yönetiminin zorluklarını ve stratejilerini kapsamlı bir şekilde ele almaktadır. Bilişsel yetenekler ve yeniden yapılandırılabilirlik de dahil olmak üzere CR ağlarının temel yönlerini, spektrum yönetiminin ana bileşenlerini tartışır; algılama, karar verme, paylaşım ve hareketliliği özetler. Araştırma, dalgalanan kullanılabilirlik ve çeşitli hizmet kalitesi gereksinimlerini karşılamak için dinamik ve verimli spektrum kullanımının önemini vurgular. CR ağlarının mevcut ağlarla çakışmadan çalışması gerektiğini ve etkili spektrum yönetimi tekniklerinin geliştirilmesindeki devam eden araştırma zorluklarını belirtir [4].

Bu bağlamda, spektrum kullanım verimliliğini en üst düzeye çıkarmak için yeni teknolojik ve düzenleyici mekanizmaların geliştirilmesi gereklilik haline gelmiştir. Bilişsel radyo teknolojisi, spektrum boşluklarını tespit ederek ve lisanssız kullanıcıların bu boşluklarda iletişim kurmasına olanak tanıyarak spektrum verimliliğini artırmayı amaçlamaktadır [5]. Önerilen yaklaşımların temeli, lisanslı kullanıcıların frekans bandını kullanmadıkları zamanlarda lisanssız kullanıcılara geçici erişim hakkı verilmesidir. Burada kritik olan, lisanssız kullanıcılar için spektrum alanları kullanılırken lisanslı kullanıcıların haklarını korumak ve lisanslı kullanıcıların faaliyetlerine asgari düzeyde müdahale etmektir.

Askeri alanlarda iletişim güvenliğinin sağlanması, spektrum kullanımının hassas bir konusu olarak öne çıkmaktadır. Yasadışı iletişimi tespit etmek ve gerektiğinde müdahale etmek, spektrumun etkili bir şekilde izlenmesi ve yönetilmesini kritik bir mesele haline getirmiştir. Özellikle stratejik öneme sahip askeri iletişim kanalları, yetkisiz erişim ve karıştırma faaliyetlerine karşı korunmalıdır. Bu, hem sivil hem de askeri kullanımlarda spektrum tahmini ve yönetim metodolojilerinin büyük önem taşıdığı bir göstergesidir. Güvenli iletişimi sağlamak ve stratejik iletişim kaynaklarını etkili bir şekilde kullanmak için spektrumun gerçekçi bir şekilde modellenmesi esastır.

Doğru spektrum yönetimi için, lisanslı kullanıcıların kanal kullanımının gerçekçi bir şekilde modellenmesi gerekmektedir. Spektrumu doğru bir şekilde modellemek ve kanal kullanımını tahmin etmek için geliştirilen yöntemler, spektrum yönetiminde önemli bir yere sahiptir. Bu yöntemler, spektrumu daha etkili bir şekilde kullanarak kablosuz iletişim teknolojilerinin sürdürülebilirliğine katkıda bulunmayı amaçlamaktadır.

Utah, Salt Lake City’de 700 MHz ile 2.8 GHz frekans aralığında spektrum kullanımının analizi ve tahmini üzerine kapsamlı bir çalışma, büyük frekans segmentlerinin sadece %19 oranında düşük kullanımda olduğunu ortaya koymaktadır. LSTM ağları kullanılarak gelecekteki kanal işgal durumlarını tahmin eden bir sistem önermektedir ve çeşitli LSTM tabanlı modellerin performansını değerlendirmektedir. Bu modellerin, net kullanım örüntüleri gösteren frekanslarda yaklaşık %96 doğrulukla spektrum kullanımını tahmin edebildiğini göstermektedir [6].

Bilişsel radyo ağlarında kanal işgalini tahmin etmek için derin öğrenme tekniklerinin nasıl kullanılacağı araştırılmaktadır. Bu, birincil kullanıcıları rahatsız etmeden az kullanılan spektrum frekanslarını verimli bir şekilde kullanmayı amaçlamaktadır. DL yaklaşımı, mevcut yöntemlere göre tahmin doğruluğunu %5 oranında önemli ölçüde artırmaktadır [7].

Uzun Kısa Süreli Bellek (LSTM) ağları, bilişsel radyo ağlarında spektrum işgalini tahmin etmek için kullanılmıştır. LSTM tahmin performansının teorik modellenmesi ve Markov modeline dayalı spektrum verilerinin kullanımını vurgulamaktadır. Analiz, LSTM tabanlı modellerin yüksek verimlilikle gelecekteki spektrum kullanımını doğru bir şekilde tahmin edebileceğini ortaya koymaktadır. Araştırma, tahmin doğruluğunu iyileştirmek için karışım modeli yaklaşımını önererek LSTM çıktılarını modelleme üzerine bir metodolojik çerçeve geliştirmektedir [8].

1.2 Tezin Amacı ve Önemi

Bu çalışma, spektrum kullanımının kestirimi üzerine odaklanmış bir araştırmadır ve makine öğrenmesi ile derin öğrenme metodolojilerinin birleşimini kullanarak bir yaklaşım sunmaktadır. Destek Vektör Makineleri (SVM), Rassal Orman Algoritması ve Uzun Kısa Süreli Bellek (LSTM) algoritmaları, spektrum verilerinin analizi ve gelecekteki kullanım durumlarının tahmin edilmesi için kullanılmıştır. Çalışmanın ayırt edici yönlerinden biri, varsayımsal veya simülasyonlardan üretilen veri setleri yerine, gerçek dünya verilerinin toplanıp kullanılmasıdır. Bu, sonuçların doğruluğunu ve uygulanabilirliğini artırarak, spektrum tahmin

modellerinin gerçek dünya koşullarında nasıl performans göstereceğine dair daha güvenilir bilgiler sağlar. Spektrum kullanımı, telekomünikasyon sektöründe kritik bir konudur, frekans bandının verimli kullanımı, hem sivil hem de askeri iletişim sistemlerinin kalitesi ve güvenliği açısından büyük önem taşır. Spektrumun bazı kanallarının yoğun, bazılarının ise az kullanılması, genellikle kullanıcı sayısının fazlalığı veya verimsiz kullanım stratejilerinden kaynaklanır. Bu çalışma, spektrumun daha adil ve etkili bir şekilde tahsis edilmesine yardımcı olacak tahmin modelleri geliştirmeyi amaçlamaktadır. Bu metodolojilerin birleşimi, spektrum kullanımının daha etkin bir şekilde yönetilmesine olanak tanır. Özellikle askeri alanda, spektrumun kullanılmayan kanallarında potansiyel olarak gerçekleşebilecek kaçak haberleşmelerin önlenmesi, güvenlik açısından büyük önem taşır. Bu çalışma, spektrum tahmininin, bu tür güvenlik tehditlerine karşı önleyici bir önlem olarak nasıl kullanılabilmesine dair bilgiler sunar. Bu araştırma, makine öğrenmesi ve derin öğrenme algoritmalarının telekomünikasyon alanında spektrum yönetimi ve güvenliği konularına katkıda bulunabileceğini göstermektedir. Gerçek dünya verileri üzerinde yapılan bu çalışma, spektrum tahmin modellerinin doğruluğunu ve uygulanabilirliğini önemli ölçüde artırarak, spektrumun daha verimli kullanılmasına ve güvenliğinin artırılmasına yardımcı olacak yöntemler sunmaktadır.

1.3 Hipotez

Bu çalışmanın temelinde yer alan hipotez, spektrum kullanımının kestirimi için makine öğrenmesi ve derin öğrenme yöntemlerinin uygulanabilirliğini ve etkinliğini göstermektir. Özellikle, Destek Vektör Makineleri (SVM), Rastgele Orman Algoritması ve Uzun Kısa Süreli Bellek (LSTM) gibi farklı algoritmaların kullanımıyla, gerçek dünya verileri üzerinden spektrumdaki kanalların gelecekteki kullanım durumlarını tahmin etme çalışmaları yapılmaktadır. Bu yöntemlerle, spektrumun verimli kullanımı ve yönetimi üzerine önemli içgörüler elde etmek ve potansiyel olarak kaçak haberleşmeyi engellemek gibi spesifik problemlere çözüm bulmak amaçlanmaktadır. Spektrumun bazı kanallarında yoğun, bazı kanallarında ise az kullanımın sebeplerini anlamayı ve gelecekteki kullanım senaryolarını daha doğru bir şekilde kestirebilmek için gerçek verilerle çalışmanın önemini vurgular. Ayrıca, spektrum tahminlerinin, kaynakların daha etkin tahsisine ve iletişim ağlarının güvenliğinin artırılmasına olanak tanıyacağına dikkat çeker. Hem sivil hem de askeri alanda, spektrum kullanımının optimizasyonu ve güvenliği, bu çalışmanın temel hedeflerindedir.

2 YÖNTEM

Bu araştırma, kablosuz iletişim sistemlerinde spektrum verimliliğinin artırılması amacıyla Medipol Üniversitesi tarafından toplanan geniş bir veri seti üzerinde, makine öğrenmesi algoritmalarının uygulanmasıyla gerçekleştirilmiştir. Çalışmanın odak noktası, belirli bir kablosuz kanalın gelecekteki güç seviyesini tahmin etmek ve bu tahminleri kullanarak dinamik spektrum erişim stratejilerini optimize etmektir.

Araştırmada, kanalın anlık güç seviyesinin yanı sıra, sinyal gürültü oranı (SNR) hesaplanmıştır. SNR, sinyalin gücünün, arka plan gürültüsünün gücüne oranı olarak tanımlanır ve iletişimin kalitesi hakkında önemli bilgiler sunar. Belirlenen bir eşik değeri kullanılarak, SNR değerlerine göre iki kategoriye ayrılmıştır: Haberleşmenin mümkün olduğu yüksek SNR değerleri ve haberleşmenin mümkün olmadığı düşük SNR değerleri. Bu sınıflandırma, kanalın mevcut durumunu (0 veya 1) temsil eder, burada 0 spektrumun boş olduğunu, 1 ise dolu olduğunu gösterir.

Toplanan veri seti üzerinde çeşitli makine öğrenmesi algoritmaları uygulanmıştır. Bu algoritmalar, veri setindeki özellikleri kullanarak kanalın gelecekteki güç seviyesini ve dolayısıyla iletişim varlığını tahmin etmek için eğitilmiştir. Algoritmaların performansı, az yoğunluğa sahip ve yoğun kullanım gösteren iki farklı kanal üzerinde değerlendirilmiştir. Bu değerlendirme, algoritmaların çeşitli koşullar altında nasıl performans gösterdiğini anlamak için kritik öneme sahiptir.

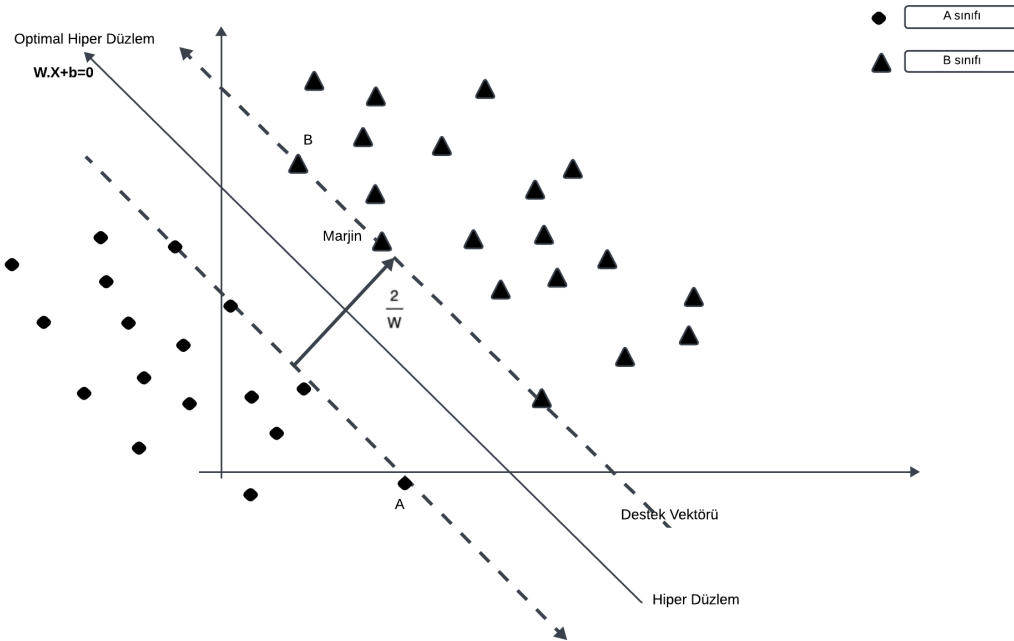
Algoritmaların karşılaştırmalı analizi, belirli koşullar altında hangi makine öğrenmesi yöntemlerinin daha etkili olduğunu ortaya koymuştur.

Kullanılan Makine öğrenmesi yöntemleri ise Destek Vektör Makineleri, Rassal Orman Algoritması ve Kısa-Uzun Süreli Bellek algoritmasıdır.

2.1 Destek Vektör Makineleri

Destek Vektör Makineleri (SVM), özellikle sınıflandırma ve regresyon görevleri için kullanılan güçlü ve esnek bir makine öğrenimi algoritmasıdır. SVM, yüksek veya sonsuz boyutlu bir özellik uzayında veri örneklerini ayırmak için optimal bir hiper-düzlem bulmayı amaçlar. Bu hiper düzlem, farklı sınıflardaki veri noktaları arasındaki en büyük marjı sağlayacak şekilde seçilir. Bu yöntem, hem lineer hem de lineer olmayan sınıflandırma problemleri için uygulanabilir.

SVM'nin temel fikri, farklı sınıflara ait veri noktalarını birbirinden ayıran bir hiper düzlem bulmaktır. Hiper düzlem, sınıflar arasındaki en büyük marjı sağlayacak şekilde seçilir. Marj, hiper düzleme en yakın noktalar arasındaki mesafedir ve SVM, bu marjı maksimize etmeyi amaçlar. Destek vektörleri, hiper düzleme en yakın olan ve dolayısıyla sınıflandırma kararını en çok etkileyen veri noktalarıdır. Bu noktalar, hiper düzlemin konumunu ve oryantasyonunu belirler. Denklem 2.1'de, w hiper düzlemin normal vektörüdür, bu vektör hiper düzlemin yönünü belirler, x özellik vektörü, b bias terimidir ve hiper düzlemin orijinden olan uzaklığını belirler. Şekil 2.1'de SVM algoritmasının ikili sınıflandırmadaki sınıfları ayırma yöntemi gösterilmiştir [9], [10], [11].



Şekil 2.1 SVM algoritmasının sınıflar arasındaki marjı

$$w \cdot x + b = 0 \quad (2.1)$$

Lineer olmayan sınıflandırma problemleri için SVM, "kernel trick" adı verilen

bir yöntem kullanır. Bu yöntem, veriyi daha yüksek boyutlu bir uzaya taşıyarak lineer olarak ayrılabilir hale getirir. Popüler kernel fonksiyonları arasında polinom, Radyal Temel Fonksiyon (RBF), ve sigmoid fonksiyonları bulunur. Denklem 2.2, $K(x_i, x_j)$ iki özellik vektörü x_i ve x_j arasındaki kernel değerini hesaplar. Burada γ pozitif bir ölçekleme faktörüdür ve $\|x_i - x_j\|^2$ ise x_i ve x_j arasındaki vektörleri arasındaki öklid mesafesinin karesini temsil eder.

$$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2) \quad (2.2)$$

2.2 Rassal Orman Algoritması

Rassal Orman (RF) algoritması, sınıflandırma ve regresyon sorunları için kullanılan güçlü bir makine öğrenmesi yöntemidir. Bu yöntem, birden fazla karar ağacının oluşturduğu bir "orman" üzerine kuruludur, ve bu ağaçların her biri, veri setinin bir önyüklemeye örneği üzerinde eğitilir. Rassal Orman Algoritmasının temel fikri, çok sayıda nispeten dekorele edilmiş modellerin topluluğunu oluşturarak, bireysel modellerin hatalarını ortadan kaldırmaktır. Bu yöntem, modelin hem varyansını hem de yanlılığını azaltmaya yardımcı olur, genel performansın artmasını sağlar.

Rassal Orman, sınıflandırma ve regresyon problemleri için kullanıldığında farklı yaklaşımlar benimser. Sınıflandırma için, her ağacın çıktısı bir sınıf etiketidir ve ormanın genel tahmini, çoğunluk oyu prensibiyle, yani en çok oy alan sınıf etiketi olarak belirlenir. Regresyon durumunda ise, her ağacın çıktısı bir sayısal değerdir ve ormanın tahmini, bu değerlerin ortalaması alınarak hesaplanır. İlginç bir şekilde, regresyon modu, olasılıkların tahmin edilmesi gibi ikili sonuçlar elde etmek için de kullanılabilir, bu da onu lojistik regresyona benzer bir araç haline getirir[12], [13].

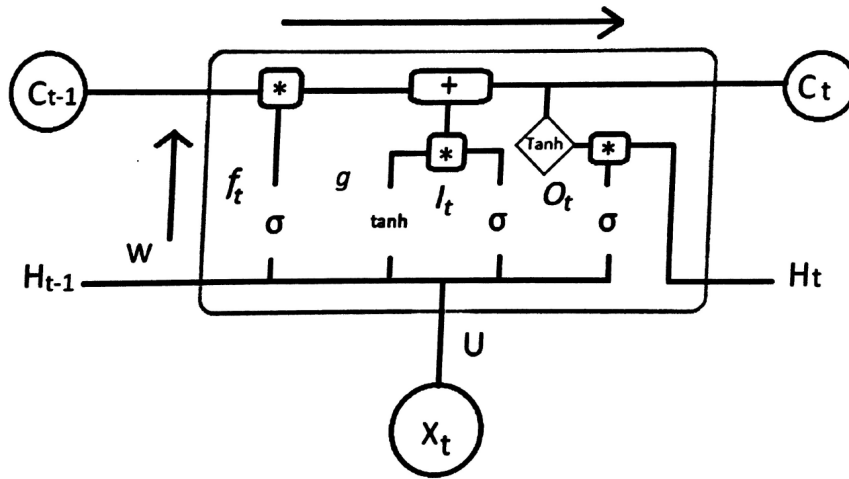
Karar ağaçlarını eğitirken, her bölünme noktasında hangi değişkenin kullanılacağına karar vermek önemlidir. Sınıflandırma durumunda, genellikle Gini endeksi kullanılır. Bu, bir düğümdeki örneklerin ne kadar iyi karıştırıldığını ölçer. Regresyon için ise, bölme kriteri olarak genellikle varyans azaltma kullanılır. Her iki durumda da, en iyi bölünme, seçilen kriteri en çok azaltan değişken ve eşik değeri kombinasyonu ile belirlenir[14].

Rassal Orman, modelin tahminlerindeki değişkenlerin önemini belirlemek için de kullanılır. Geleneksel bir yaklaşım, Gini endeksindeki ortalama düşüşü kullanmaktır; ancak, bu yöntem bazı durumlarda taraflı olabilir. Daha objektif bir alternatif, değişkenlerin permütasyon önemidir. Bu yaklaşımda, bir değişkenin değerleri tüm ağaçlarda rastgele karıştırılır ve modelin performansındaki değişiklik ölçülür. Bir değişkenin önemi, permütasyon öncesi ve sonrası performans

arasındaki fark olarak hesaplanır. Bu yöntem, modelin tahmin performansına gerçekten katkıda bulunan değişkenleri belirlemeye yardımcı olur[15].

2.3 Uzun Kısa Süreli Bellek (LSTM)

Uzun Kısa Süreli Bellek (LSTM) ağları, özellikle zaman serileri verileri, doğal dil işleme ve diğer sıralı veri türlerinde yaygın olarak kullanılan bir tür yapay sinir ağıdır. LSTM'ler, geleneksel geri yayılım algoritmalarının karşılaştığı uzun bağımlılıkların öğrenilmesi konusundaki zorlukların üstesinden gelmek için tasarlanmıştır. Şekil 2.2'de LSTM algoritmasının yapısı gösterilmiştir [16], [17], [18], [19].



Şekil 2.2 LSTM algoritmasının yapısı

LSTM ağları, temelde üç kapıdan oluşur: unutma kapısı, giriş kapısı ve çıkış kapısı. Bu kapılar, hücrenin bilgiyi nasıl saklayacağını, unutacağını veya bilgiyi çıktı olarak vereceğini kontrol eder.

Unutma Kapısı: Hücrenin eski bilgileri unutulmasını sağlar. Önceki durumdan ve mevcut girdiden gelen bilgiler bir sigmoid fonksiyonu tarafından işlenir. Bu, hangi bilginin tutulup hangisinin unutulacağını belirler. Denklem 2.3, f_t unutma kapısının çıktısını, σ sigmoid aktivasyon fonksiyonunu, W_f unutma kapısının ağırlık matrisini, h_{t-1} önceki zaman adımının hücre çıktısını, x_t mevcut zaman adımının girdisini, b_f unutma kapısının bias vektörünü temsil etmektedir.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.3)$$

Giriş Kapısı: Yeni bilginin hücre durumuna ne kadar ekleneceğini belirler. Yeni giriş, bir sigmoid fonksiyonu ile hangi değerlerin güncelleneceğine karar verirken, bir tanh fonksiyonu yeni aday değerleri oluşturur. Denklem 2.4 ve 2.5, i_t giriş kapısının çıktısını, σ sigmoid aktivasyon fonksiyonunu, W_i giriş kapısının ağırlık matrisini, h_{t-1} önceki zaman adımının hücre çıktısını, x_t mevcut zaman adımının girdisini, b_i giriş kapısının bias vektörünü, \tilde{C}_t hücre durumu için aday değerleri, W_C aday değerlerin ağırlık matrisini, b_C bias vektörünü temsil etmektedir.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.4)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2.5)$$

Çıkış Kapısı: Hücrenin sonraki katmana ne kadar bilgi göndereceğini belirler. Hücre durumunun bir kısmı, bir sigmoid ve bir tanh fonksiyonu ile işlendikten sonra çıktı olarak verilir. Denklem 2.7 ve 2.8, o_t çıkış kapısının çıktısını, σ sigmoid aktivasyon fonksiyonunu, W_o çıkış kapısının ağırlık matrisini, h_{t-1} önceki zaman adımının hücre çıktısını, x_t mevcut zaman adımının girdisini, b_o çıkış kapısının bias vektörünü, h_t mevcut hücrenin çıktısını, C_t mevcut hücre durumunu, C_{t-1} önceki hücre durumunu temsil etmektedir. Denklem 2.6, önceki hücre durumunun unutmaya kapısı ile nasıl modifiye edildiğini ve yeni bilginin giriş kapısı ve aday değerler aracılığıyla nasıl eklendiğini gösterir.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.6)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.7)$$

$$h_t = o_t * \tanh(C_t) \quad (2.8)$$

3

UYGULAMA

Elde edilen verilerin analizi için Destek Vektör Makineleri (SVM), Rassal Orman algoritması, Uzun Kısa Süreli Bellek (LSTM) ağırları yöntemleri kullanılmıştır. SVM algoritmasının seçilmesinin sebebi kendinden önceki durumların birbirinden bağımsız olma durumunu incelemek ve çok boyutlu özelliklerde hızlı öğrenme süresine sahip olması, RF algoritmasının seçilmesinin sebebi ise yine özelliklerin birbirinden bağımsız olma durumu, problemin bir sınıflandırma problemi olması ve sınıflandırma için bir eşik değerinin olma durumu, LSTM algoritmasının seçilmesi ise veri setinin zaman serisi olması ve kendinden önceki durumlara bağlı olma durumundan dolayı bu üç algoritmalar seçilmiştir.

3.1 Veri Seti

Spektrum kullanımını tahmin etmek üzere, Medipol Üniversitesi tarafından Agilent EXA N9010A spektrum analizörü ile Medipol Üniversitesi kampüsünde yapılan bu ölçümler, 21 saat 38 dakika 3 saniye boyunca kesintisiz olarak saniyede bir frekansta gerçekleştirilmiş ve 800 MHz - 2.6 GHz arasında toplamda 1000 farklı kanal olacak şekilde incelenmiştir. Bant genişliği 1.6 MHz, kullanılan cihazın noise figure değeri 8 – 12 dB arasında kullanılan kanalın merkez frekansına bağlı olarak değişmektedir. Bu süre zarfında, çeşitli sinyallerin güç seviyeleri detaylı bir şekilde kaydedilmiştir. Şekil 3.2 ve Şekil 3.1’de örnek olarak kanalların güç ölçümleri gösterilmiştir. Uygulamada kullanılacak kanallar LTE band20 olan 800 MHz ve 838 MHz frekanslı kanalların ölçümleri alınmıştır. LTE band20’de 800 MHz downlink, 838 MHz uplink frekansdır. 838 MHz uplink frekansında kanalın gücünün az görünmesinin sebebi vericinin gücünün az olması, ölçüm için kullanılan antenin kapalı bir alanda olması, antenin yüksekte olmaması, vericilerin antene uzaklığı, ek kayıplar ve vericilerin bina içerisinde olmasıdır. Vericinin gücünün az olması sinyalin yol kaybına olan hassasiyetini artırır, ölçüm yapılan antenin kapalı bir alanda olması yine yol kaybını artıracaktır, antenin yüksekte olmama durumu direkt olarak sinyali almasına engel olacaktır ve gelen sinyaller

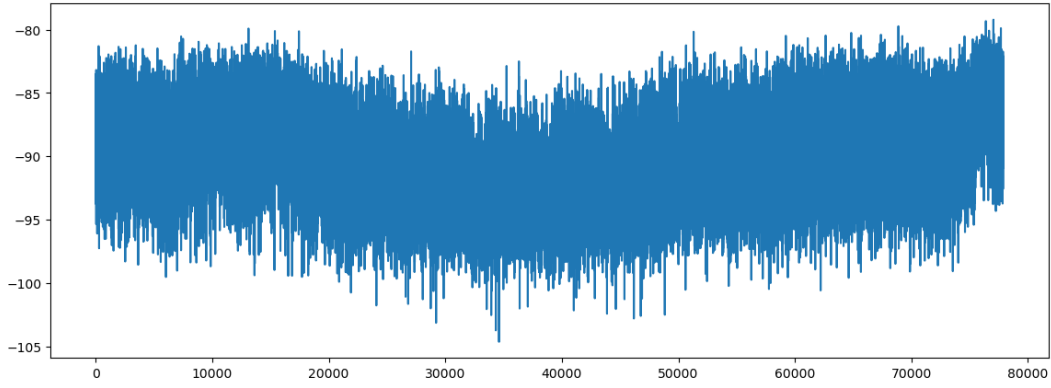
daha fazla yol kaybına uğrayacaktır, vericilerin gücü az olduğu için antenden uzaklaştıkça yol kaybına hassasiyeti artacaktır, vericilerin bina içerisinde olması sinyalin antene ulaşana kadar bir çok duvar ve kattan geçmesi gerekeceğinden ekleme kaybına sebep olacaktır ve sinyal fazla zayıflayacaktır. Ekleme kaybı, sinyalin iletilirken sinyalin ek duvarlardan geçme durumunda oluşan toplam kayıptır. Bu sebeplerden sadece lokal bölgelerdeki sinyalleri alacağı için bu kanalda haberleşme yokmuş gibi görünecektir. 800 MHz downlink frekansının verici gücü yüksek olduğundan ve sadece binanın içinden değil dışından da sinyal alabildiği için kalabalık gibi görünecektir. Bu sebepten downlink kalabalık, uplink kanalı ise boş kanal olarak varsayılmıştır. Kanalin o andaki gücü düşük ise boş, yüksek ise doludur denmiştir. Bu varsayıma göre kanalların durumunu Tablo 3.1'de de gösterildiği üzere 800 MHz bandındaki kanalın güç dağılımı yüksek iken 838 Mhz bandındaki kanalın güç dağılımı düşüktür. 800 MHz bandındaki kanal, kullanımı yoğun kanal olarak ve 838 Mhz bandındaki kanal, ise az kullanılan kanal olarak modellenmiştir. Sivil alanlarda spektrum kısıtlılığının yaşanabileceği kanalların modellenmesi için 800 MHz frekanslı kanal seçilmiştir. Askeri alanlarda spektrum taranırken çok kullanılan kanallar yoğun bir şekilde incelenirken az kullanılan kanallar yoğun bir şekilde incelenemeyeceğinden burada yapılan kaçak haberleşmelerin tespiti için az yoğun kanalın modellenmesi için 838 MHz merkez frekanslı kanal seçilmiştir. Bu çalışmanın amacı, sivil ve askeri alanlarda spektrumu kullanımı tahmini olduğundan sivil alanlarda spektrumun verimliliğini arttırmak için yoğun kanallardan biri, askeri alanlarda spektrumda gerçekleşebilecek illegal haberleşmelerin tespit edilmesi için yoğunluğu çok az olan kanal tercih edilmiştir. Gerçek bir askeri alan probleminin çözülebilmesi için askeri bölgelerden veri seti oluşturulmalıdır.

Tablo 3.1 Kullanılan kanalların karakteri

Kanal Bilgisi	800 MHz	838 MHz
Zaman	77882	77882
Ortalama	-90.17 dBm	-96.18 dBm
Standart Sapma	3.19 dBm	1.68 dBm
Minimum Güç	-104.62 dBm	-108.33 dBm
Maksimum Güç	-79.22 dBm	-73.13 dBm
% 25	-92.49 dBm	-97.11 dBm
% 50	-90.41 dBm	-96.01 dBm
% 75	-87.93 dBm	-95.05 dBm

Kanallardaki gürültü, 3.1 ve 3.2 denklemleri kullanılarak kanallardaki gürültünün gücü hesaplanmıştır.

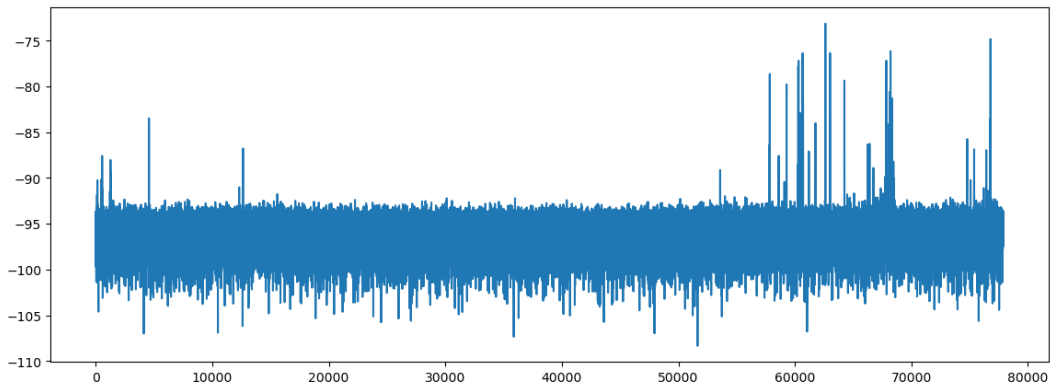
$$N = kTB \quad (3.1)$$



Şekil 3.1 Veri setinin 800 MHz frekansındaki ölçümleri

$$N_{dBm} = 10 \log_{10}(N \times 1000) \quad (3.2)$$

Kanallardaki gürültü güçleri hesaplandıktan sonra eşik değeri belirlenmiştir. Denklem 3.3 ve 3.4'den yararlanılarak SNR değeri belirlenmiştir. Değişken veri tiplerine göre çeşitli minimum SNR değerleri vardır. 20 dB SNR değeri bir çok veri tipi için minimum değer olarak düşünülebilir. Minimum SNR değerinde haberleşmenin olabileceği düşüncesinden dolayı eşik değeri için 20 dB SNR değerine sahip sinyal gücü belirlenmiştir. Eşik değerinin 20 dB seçilmesi sınırdaki güçlerin sınıflandırmasını güçleştirecektir ve bu algoritmaları zorlu bir veri seti ile karşı karşıya bırakacaktır. Bu eşik değerinin altında kalan güç değerleri 0 olarak, üzerinde kalan güç değerleri ise 1 olarak etiketlenmiştir. Spektrumda kullanıcının olmadığı durumlar 0, kullanıcının var olduğu durumlar 1 olarak belirtilmektedir. Bir başka deyişle spektrum dolu ise 1 boş ise 0 sonucunu almaktadır. Böylelikle veri seti ikili sınıflandırmaya uygun hale getirilmiştir. Algoritmaların tahminleri belirlenen eşik değerlerine göre sınıflandırılmış olacaktır.



Şekil 3.2 Veri setinin 838 MHz frekansındaki ölçümleri

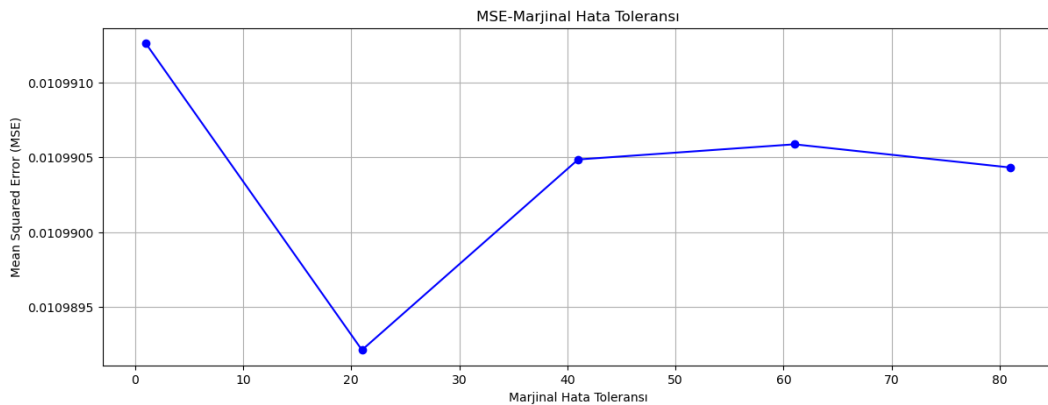
$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{\text{Sinyal Gücü}}{\text{Gürültü Gücü}} \right) \quad (3.3)$$

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{10^{\left(\frac{\text{Sinyal Gücü}_{\text{dBm}}}{10}\right)}}{10^{\left(\frac{\text{Gürültü Gücü}_{\text{dBm}}}{10}\right)}} \right) \quad (3.4)$$

3.2 Destek Vektör Makinesi

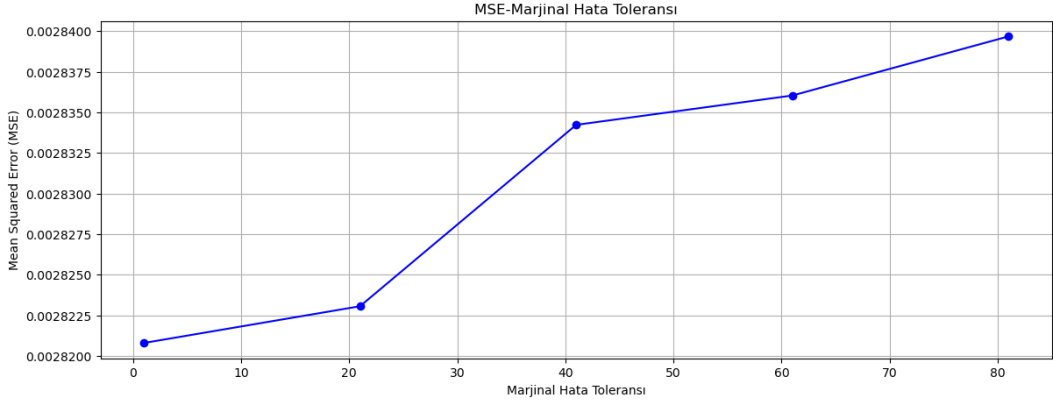
Bu çalışmada, algoritmalar için seçilen yoğunluğu az olan kanal ile yoğun olan kanal üzerinde Destek Vektör Makinesi kullanılmıştır. SVM algoritmasında veri seti düzenlenmesi yapılmıştır. Tahmin edilen $t + 1$. durum kendinden önceki 10 saniye (durum) ile ilişkilendirilmiştir. $t - 9$. andan $t - 9$. ana kadar olan durumlar ile ilişkilendirilmiştir. RBF kernel, verileri daha yüksek boyutlu bir uzaya taşıyarak doğrusal olmayan ilişkileri modelleyebilir, γ parametresi ile modelin esnekliği ayarlanabilir. Yüksek bir γ değeri, modelin veriye daha duyarlı olmasına, düşük bir γ değeri ise daha genelleştirici bir model oluşturulmasına yardımcı olur. RBF, farklı türdeki ve yapıdaki verilerde etkili bir şekilde çalışabilir. Bu sebeplerden dolayı RBF kerneli tercih edilmiştir.

Veri seti eğitim, test ve doğrulama seti olarak üçe ayrılmıştır. Veri setinin %60'nın eğitim seti, %20'si test seti ve %20'si de doğrulama seti olarak ayrılmıştır. SVM algoritmasının parametrelerinden biri olan marjinal hata toleransı (C) ile ortalama hata karesi ilişkisi incelenmiş ve en optimum parametre seçilmiştir. Şekil 3.3 ve 3.4'da görüldüğü üzere uygun C parametresi seçilmiştir.

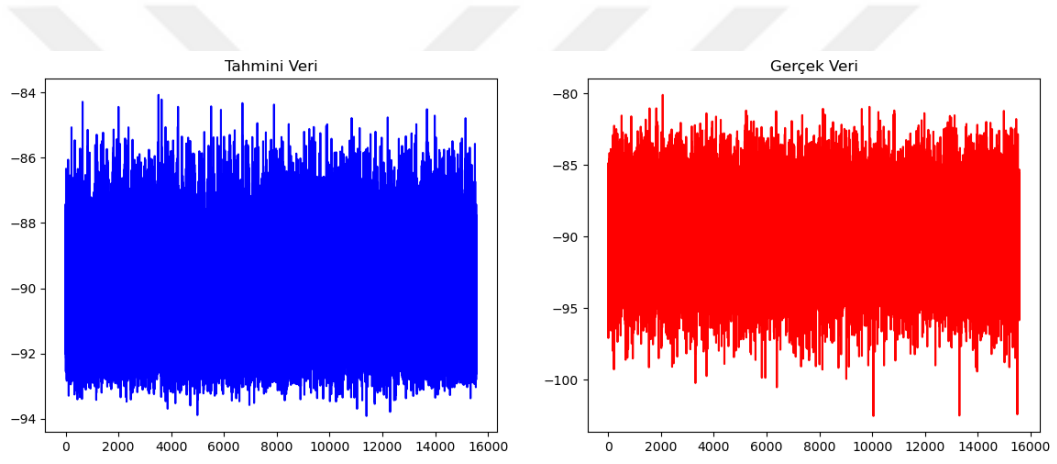


Şekil 3.3 Destek vektör makinesi 800 MHz'de C parametresi

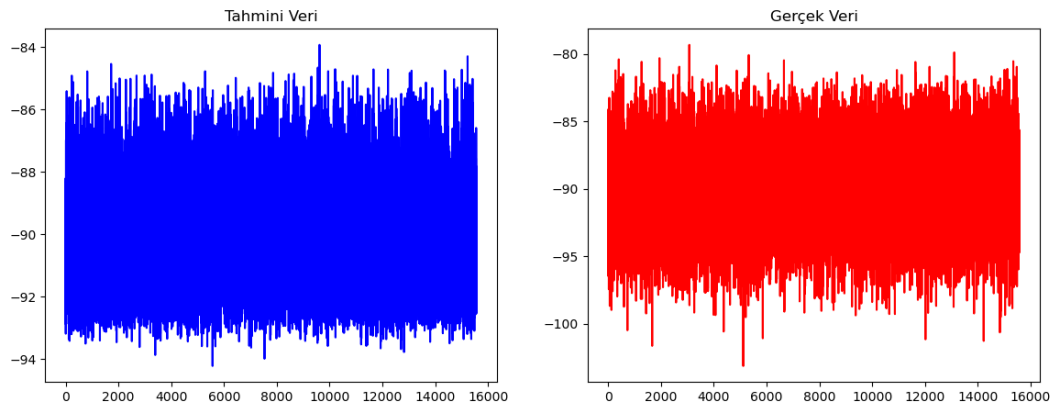
Belirlenen optimum C parametresi ile SVM algoritması ile 800 MHz frekansındaki kanalın gelecek durumu ile ilgili kestirimler yapılmıştır. Destek Vektör Makinesinin test ve doğrulama kestirimleri şekil 3.5 ve 3.6'de gösterilmiştir.



Şekil 3.4 Destek vektör makinesi 838 MHz’de C parametresi

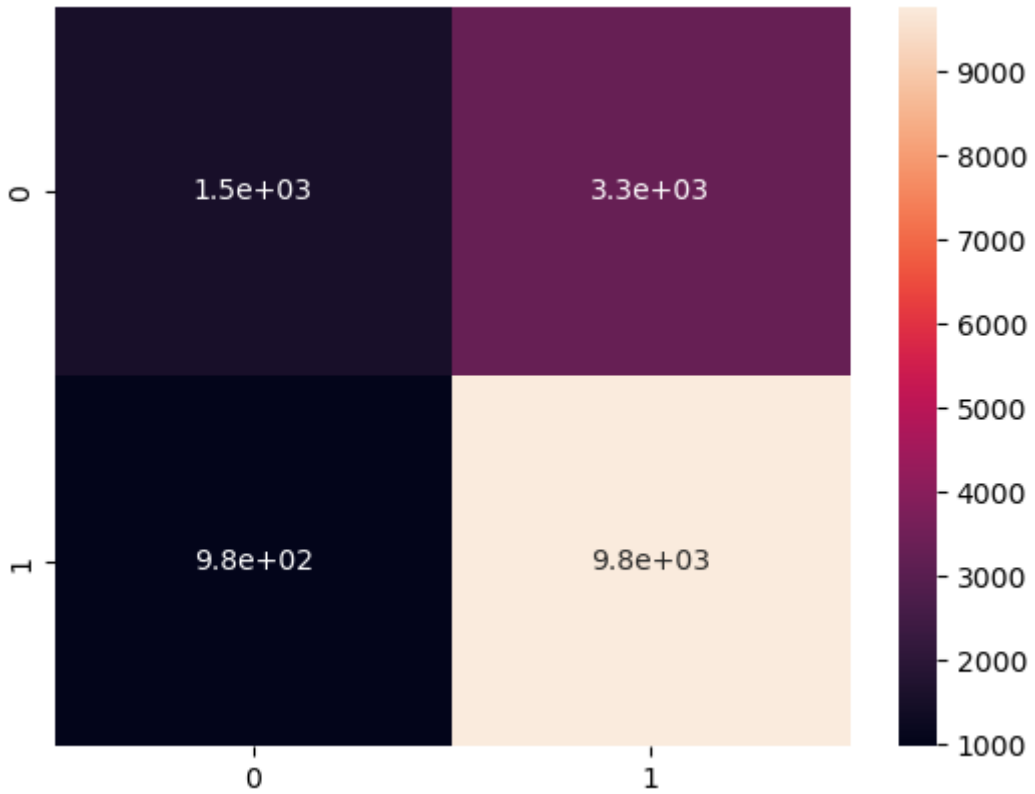


Şekil 3.5 800 MHz’de destek vektör makinesi test tahminleri



Şekil 3.6 800 MHz’de destek vektör makinesi doğrulama tahminleri

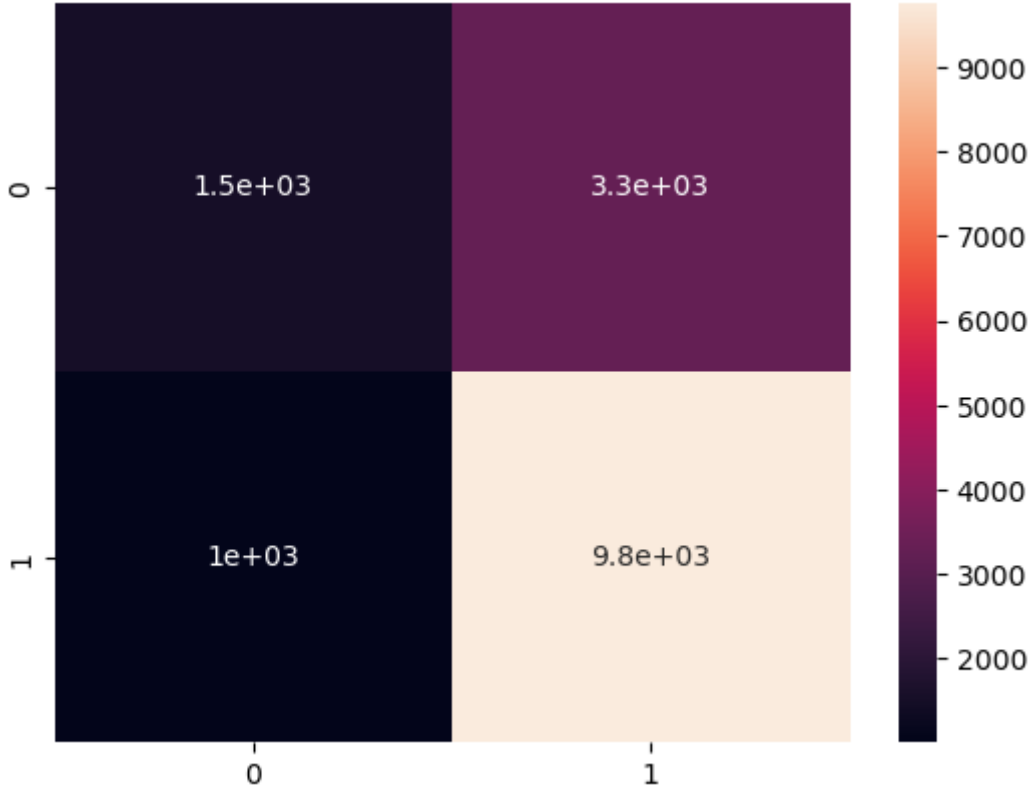
Kestirilen deęerler eřik deęerinden geęirilerek, kanalın mevcut durumunu (0 veya 1) belirlenir, burada 0 spektrumun boş olduęunu, 1 ise dolu olduęunu gösterir. Őekil 3.7’de gösterildięi üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduęu durum sayısı 4828’dir. 1518 adet boş olan durum boş olarak kestirilmiř, 3310 adeti ise dolu olarak kestirilmiřtir. Kanalın dolu olduęu durum sayısı 10747’dir. 9780 adet dolu olan durum dolu olarak kestirilmiř, 967 adeti ise boş olarak kestirilmiřtir Algoritmanın boş olarak kestirdięi 2480 durumun 1500 durumu boş 980 durumu ise dolu olarak kestirilmiřtir. Algoritmanın dolu olarak kestirdięi 13100 durumun 9800 durumu dolu 3300 durumu ise boş olarak kestirilmiřtir. Algoritmanın boş olarak kestirdięi 2480 durumun 980 durumu dolu 1500 durumu ise boş olarak kestirilmiřtir.



Őekil 3.7 800 MHz’de destek vektör makinesi test verisi hata matrisi

Őekil 3.8’de gösterildięi üzere doęrulama veri setinde belirlenen kanal durumlarının kanalın boş olduęu durum sayısı 4801’dir. 1500 adet boş olan durum boş olarak kestirilmiř, 3301 adeti ise dolu olarak kestirilmiřtir. Kanalın dolu olduęu durum sayısı 10747’dir. 9752 adet dolu olan durum dolu olarak kestirilmiř, 995 adeti ise boş olarak kestirilmiřtir Algoritmanın boş olarak kestirdięi 2500 durumun 1500 durumu boş 1000 durumu ise dolu olarak kestirilmiřtir. Algoritmanın dolu olarak kestirdięi 13100 durumun 9800 durumu dolu 3300 durumu ise boş olarak kestirilmiřtir. Algoritmanın boş olarak kestirdięi 2500 durumun 1000 durumu dolu

1500 durumu ise boş olarak kestirilmiştir.



Şekil 3.8 800 MHz’de destek vektör makinesi doğrulama verisi hata matrisi

Algoritmanın başarımlı oranı test veri setinde %72, doğrulama veri setinde %73’dir. Kanalın boş ve dolu olma durumu dengeli bir dağılıma sahip olmadığı için sadece başarımlıma bakılması eksik olacaktır. Bu sebepten ayrıca ROC eğrisi de incelenmelidir. Denklem 3.5 denklemlerindeki TPR hassasiyet değeridir.

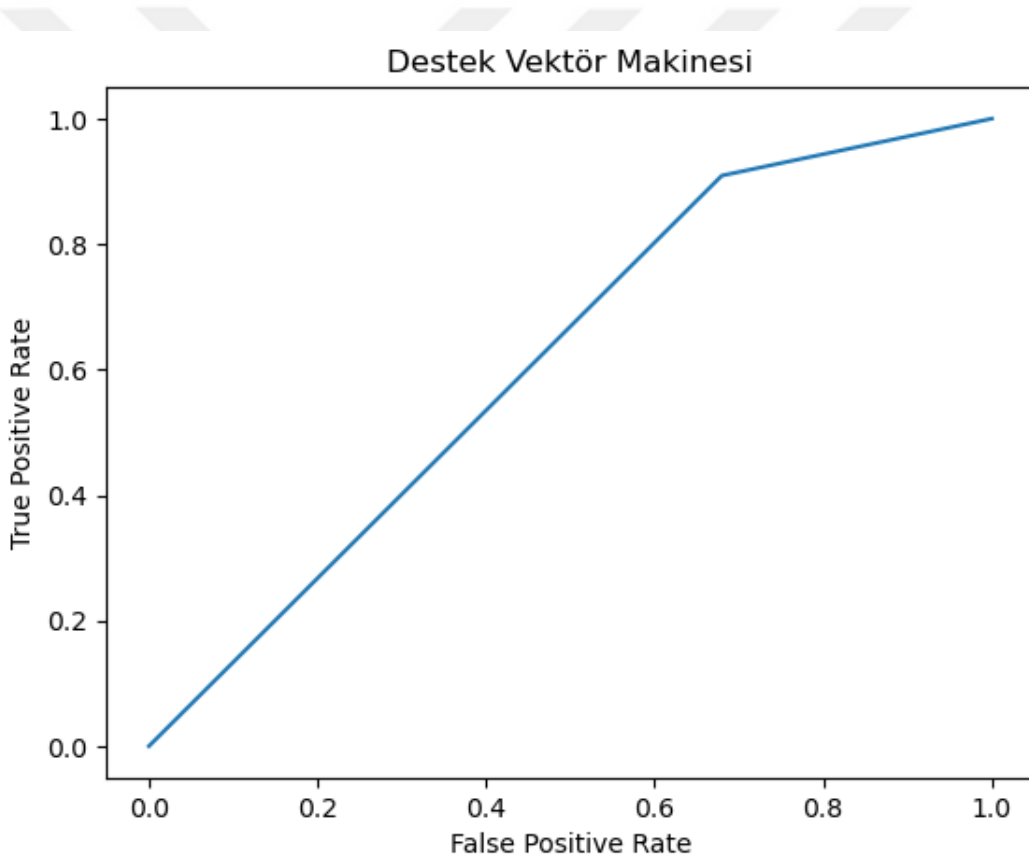
$$TPR = \frac{TP}{TP + FN} \quad (3.5)$$

$$FPR = \frac{FP}{FP + TN} \quad (3.6)$$

TPR (True Positive Rate) VE FPR (False Positive Rate) değerleri ile eğri çizilmektedir. TPR, diğer adıyla duyarlılık, pozitif olarak sınıflandırılması gereken öğelerin ne kadarının doğru olarak pozitif olarak sınıflandırıldığını gösterir. TP (True Positive), Gerçekten pozitif olan ve model tarafından pozitif olarak sınıflandırılan durum sayısı, FN (False Negative) gerçekten pozitif olan ancak model tarafından yanlışlıkla negatif olarak sınıflandırılan durum sayısı gösterir. FPR, modelin negatif olarak sınıflandırılması gereken öğeleri yanlışlıkla ne oranda

pozitif olarak sınıflandırdığını gösterir. FP (False Positive), gerçekten negatif olan ancak model tarafından yanlışlıkla pozitif olarak sınıflandırılan durum sayısı, TN (True Negative) gerçekten negatif olan ve model tarafından doğru olarak negatif olarak sınıflandırılan durum sayısını gösterir. AUC, ROC eğrisinin altında kalan alanı ifade eder. Eğrinin altında kalan alan ne kadar 1'e yakınsa algoritmanın ayırt ediciliği o kadar artmaktadır. Kesinlik, pozitif olarak tahminlenen değerlerin gerçekten kaç adedinin pozitif olduğunu göstermektedir. Duyarlılık ise pozitif olarak tahmin edilmesi gereken işlemlerin ne kadarının pozitif olarak tahmin edildiğini gösteren bir metriktir. f1 Score değeri bize Kesinlik ve Duyarlılık değerlerinin harmonik ortalamasını göstermektedir.

Şekil 3.9'de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.61 olarak hesaplanmıştır.

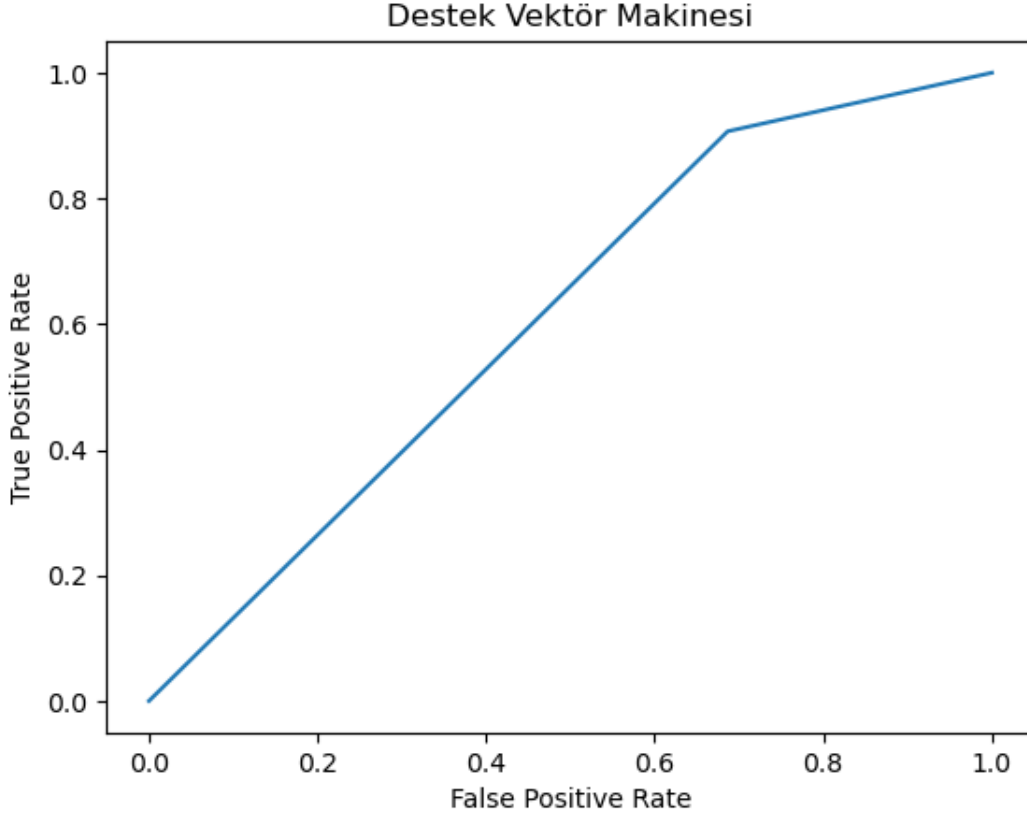


Şekil 3.9 800 MHz'de destek vektör makinesi test verisi ROC eğrisi

Şekil 3.9'de gösterildiği gibi doğrulama verisinin ayırt edicilik değeri olan AUC değeri 0.615 olarak hesaplanmıştır.

800 MHz frekanslı kanalın Test verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %61, duyarlılığı %32, f1 score değeri ise %42, kanalın dolu olma durumunun kesinliği %75, duyarlılığı %91, f1 score değeri ise %82 olarak

hesaplanmıştır. Kanalin gerçekte boş olan durumlarının %32'i kestirilmiş, boş olarak tahmin edilen durumların %61'i gerçekten boş olma durumudur. Kanalin gerçekte dolu olan durumlarının %91'i kestirilmiş, dolu olarak tahmin edilen durumların %75'i gerçekten dolu olma durumudur.



Şekil 3.10 800 MHz'de destek vektör makinesi doğrulama verisi ROC eğrisi

Tablo 3.2 800 MHz'de SVM algoritma test verisi raporu

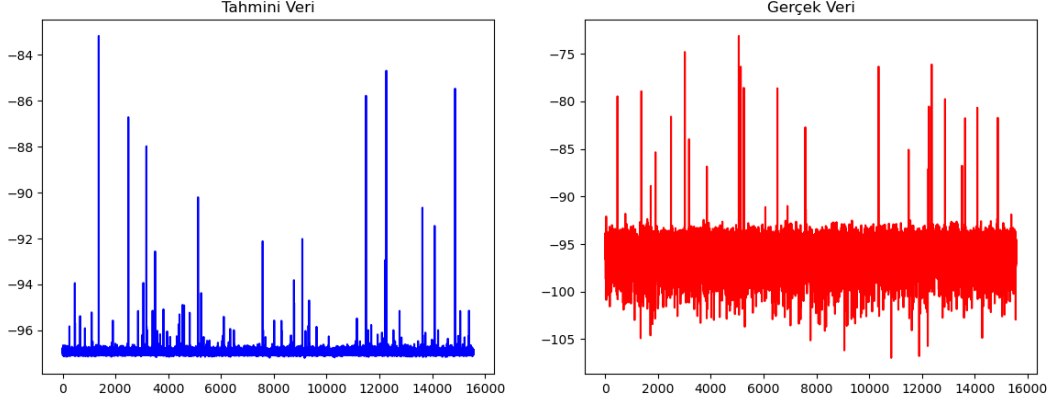
Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.61	0.32	0.42	4828
Dolu	0.75	0.91	0.82	10747

800 MHz frekanslı kanalın doğrulama verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %60, duyarlılığı %31, f1 score değeri ise %41, kanalın dolu olma durumunun kesinliği %75, duyarlılığı %91, f1 score değeri ise %82 olarak hesaplanmıştır. Kanalin gerçekte boş olan durumlarının %31'i kestirilmiş, boş olarak tahmin edilen durumların %60'ı gerçekten boş olma durumudur. Kanalin gerçekte dolu olan durumlarının %91'i kestirilmiş, dolu olarak tahmin edilen durumların %75'i gerçekten dolu olma durumudur.

Belirlenen optimum C parametresi ile SVM algoritması ile 838 MHz frekansındaki kanalın gelecek durumu ile ilgili kestirimler yapılmıştır. Destek Vektör Makinesinin test ve doğrulama kestirimleri şekil 3.11 ve 3.12'de gösterilmiştir.

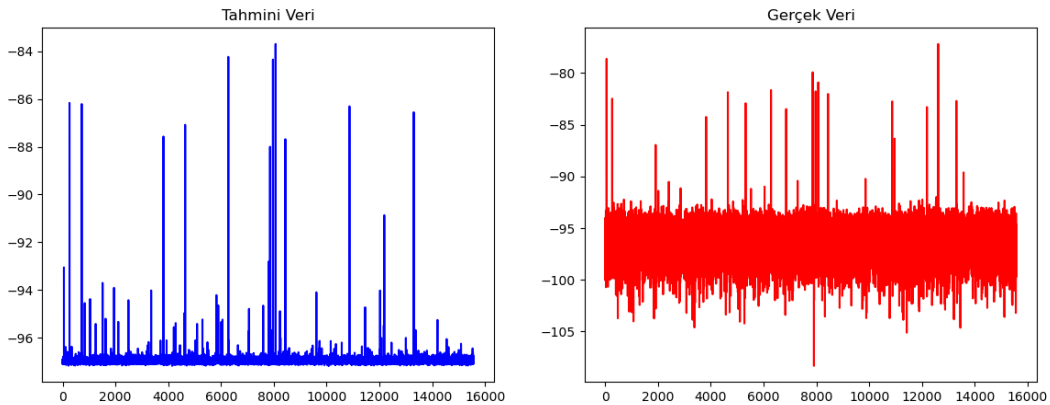
Tablo 3.3 800 MHz’de SVM algoritma doğrulama verisi raporu

Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.60	0.31	0.41	4801
Dolu	0.75	0.91	0.82	10747



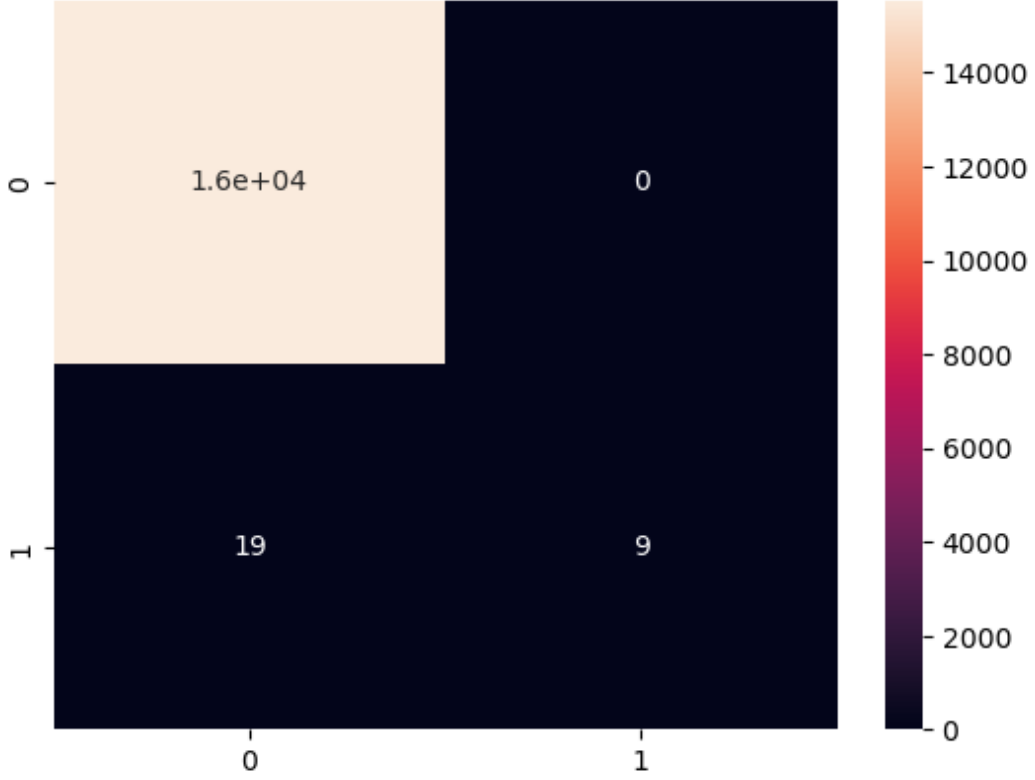
Şekil 3.11 838 MHz’de destek vektör makinesi test tahminleri

Kestirilen değerler 838 MHz frekansındaki kanal için belirlenen eşik değerinden geçirilerek, kanalın mevcut durumunu (0 veya 1) belirlenir, burada 0 spektrumun boş olduğunu, 1 ise dolu olduğunu gösterir. Şekil 3.13’de gösterildiği üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 15547’dir. 15547 adet boş olan durum boş olarak kestirilmiş, 9 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 28’dir. 9 adet dolu olan durum dolu olarak kestirilmiş, 19 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 15566 durumun 15547 durumu boş 19 durumu ise dolu olarak kestirilmiştir. Algoritmanın dolu olarak kestirdiği 9 durumun 9 durumu dolu 0 durumu ise boş olarak kestirilmiştir.



Şekil 3.12 838 MHz’de destek vektör makinesi doğrulama tahminleri

Algoritmanın dolu olarak kestirdiği 9 durumun 9 durumu dolu 0 durumu ise boş olarak kestirilmiştir.



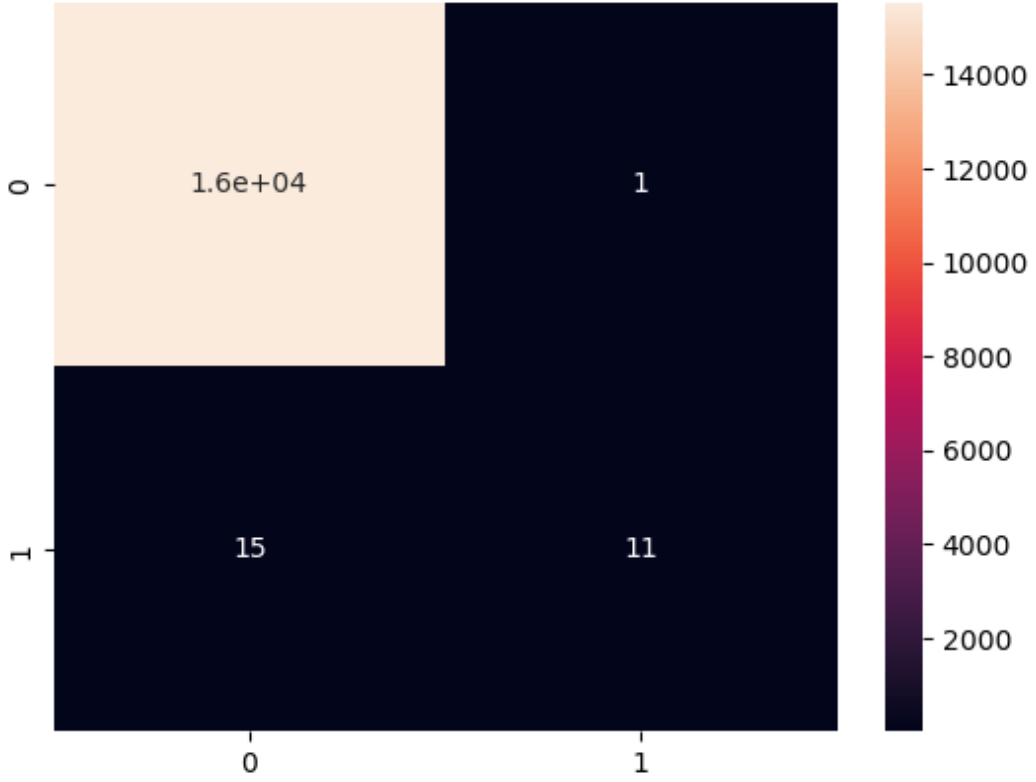
Şekil 3.13 838 MHz’de destek vektör makinesi test verisi hata matrisi

Şekil 3.14’de gösterildiği üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 15547’dir. 15547 adet boş olan durum boş olarak kestirilmiştir, 9 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 28’dir. 9 adet dolu olan durum dolu olarak kestirilmiştir, 19 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 15566 durumun 15547 durumu boş 19 durumu ise dolu olarak kestirilmiştir. Algoritmanın dolu olarak kestirdiği 12 durumun 11 durumu dolu 1 durumu ise boş olarak kestirilmiştir.

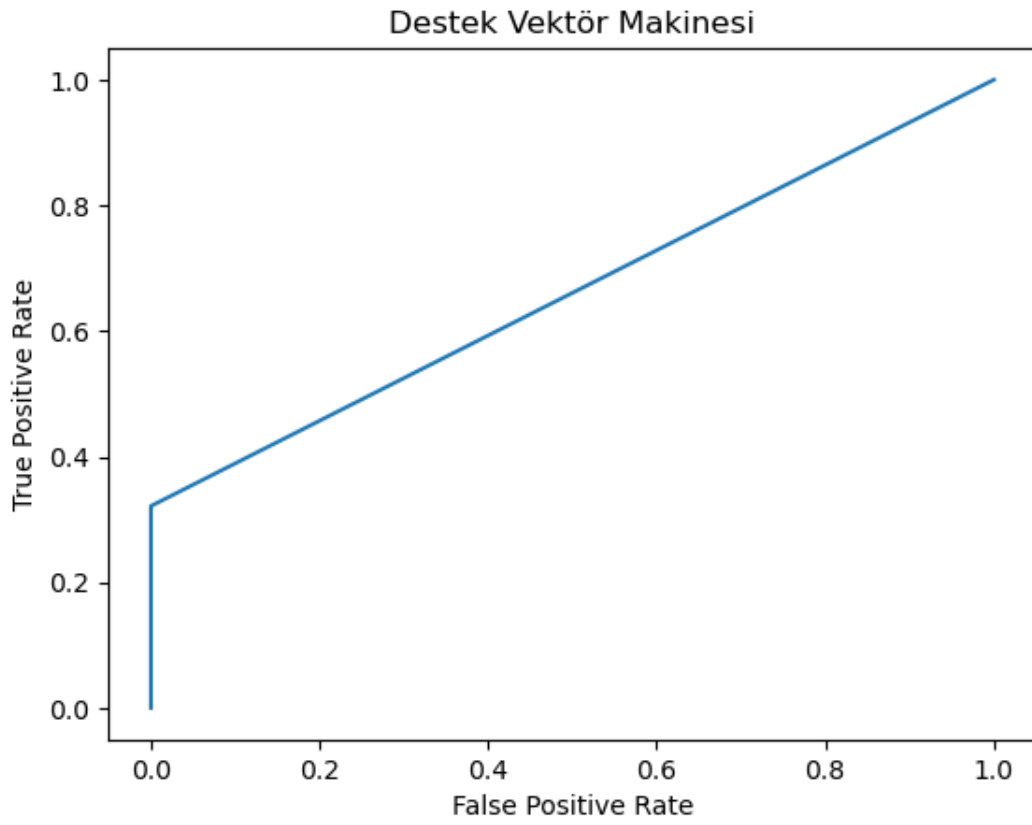
Algoritmanın başarı oranı test veri setinde %99, doğrulama veri setinde %99’dür. Kanalın boş ve dolu olma durumu dengeli bir dağılıma sahip olmadığı için sadece başarıma bakılması eksik olacaktır. Bu sebepten ayrıca ROC eğrisi de incelenmelidir. Denklem 3.5 denklemlerindeki TPR hassasiyet değeridir.

Şekil 3.15’de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.66 olarak hesaplanmıştır. Şekil 3.16’de gösterildiği gibi doğrulama verisinin ayırt edicilik değeri olan AUC değeri 0.71 olarak hesaplanmıştır.

838 MHz frekanslı kanalın test verisinin sonuçları, kanalın boş olma durumunun kesinliği %99, duyarlılığı %99, f1 score değeri ise %99, kanalın dolu olma durumunun kesinliği %99, duyarlılığı %32, f1 score değeri ise %58’dir.

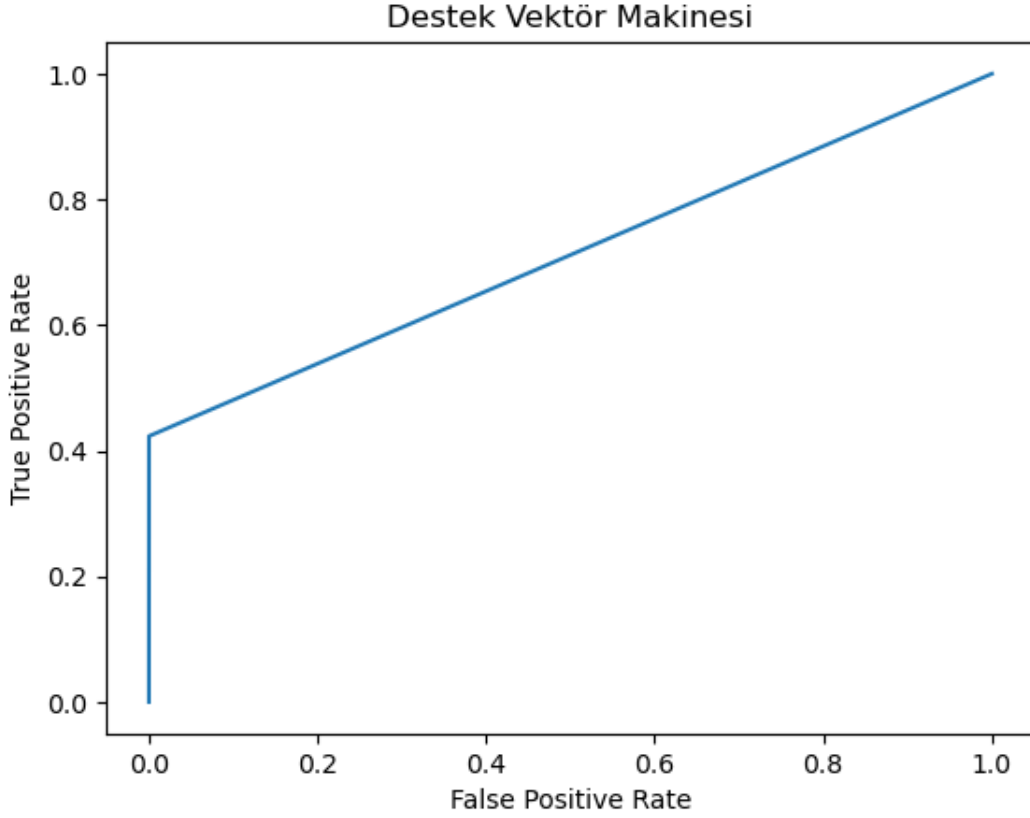


Şekil 3.14 838 MHz’de destek vektör makinesi doğrulama verisi hata matrisi



Şekil 3.15 838 MHz’de destek vektör makinesi test verisi ROC eğrisi

Kanalın gerçekte boş olan durumlarının %100'ü kestirilmiş, boş olarak tahmin edilen durumların %99'ü gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %32'i kestirilmiş, dolu olarak tahmin edilen durumların %99'ü gerçekten dolu olma durumudur.



Şekil 3.16 838 MHz'de destek vektör makinesi doğrulama verisi ROC eğrisi

838 MHz frekanslı kanalın doğrulama verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %100, duyarlılığı %100, f1 score değeri ise %100, kanalın dolu olma durumunun kesinliği %92, duyarlılığı %42, f1 score değeri ise %58 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %100'ü kestirilmiş, boş olarak tahmin edilen durumların %100'ü gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %42'si kestirilmiş, dolu olarak tahmin edilen durumların %92'si gerçekten dolu olma durumudur.

Tablo 3.4 838 MHz'de SVM algoritma doğrulama verisi raporu

Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.99	0.99	0.99	15548
Dolu	0.92	0.42	0.58	26

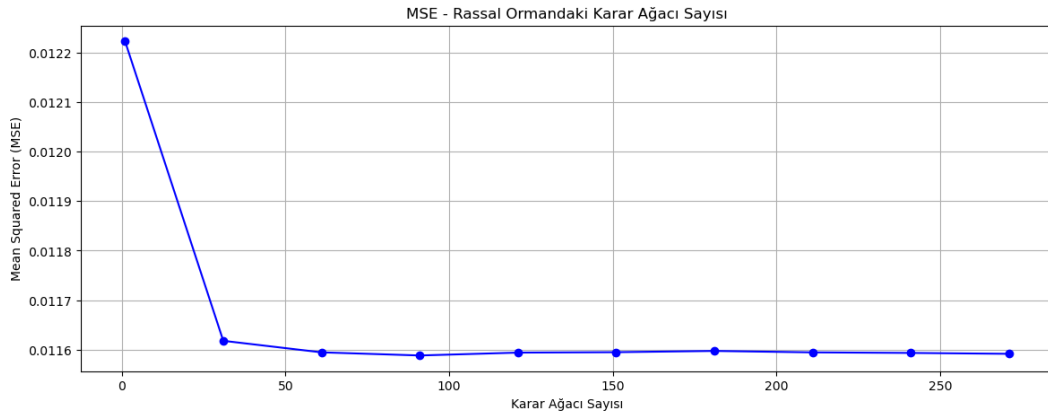
Tablo 3.5 838 MHz'de SVM algoritma test verisi raporu

Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.99	0.99	0.99	15547
Dolu	0.99	0.32	0.58	28

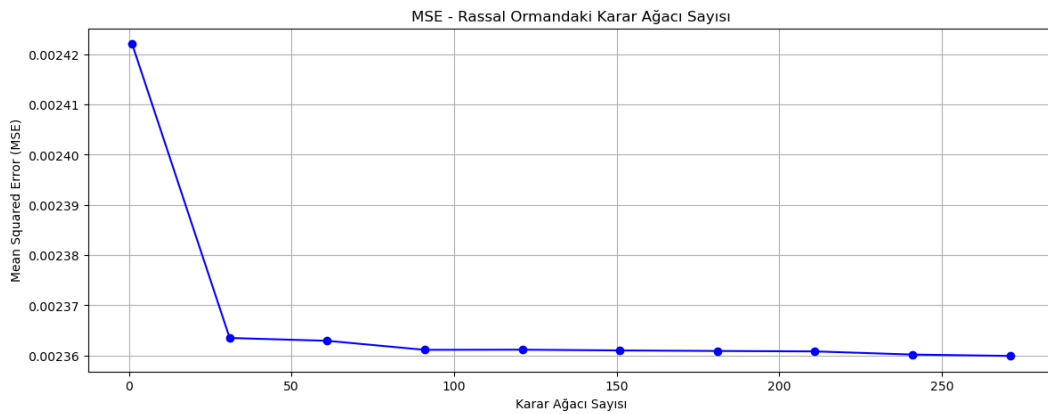
3.3 Rassal Orman Algoritması

Rassal Orman algoritmasında veri seti düzenlenmesi yapılmıştır. Tahmin edilen $t + 1$. durum kendinden önceki 10 saniye (durum) ile ilişkilendirilmiştir. $t - 9$. andan $t - 9$. ana kadar olan durumlar ile ilişkilendirilmiştir.

Veri seti eğitim, test ve doğrulama seti olarak üçe ayrılmıştır. Veri setinin %60'nın eğitim seti, %20'si test seti ve %20'si de doğrulama seti olarak ayrılmıştır. Rassal Orman algoritmasının parametrelerinden biri olan Karar Ağacı Sayısı ile ortalama hata karesi ilişkisi incelenmiş ve optimum karar ağacı sayısı belirlenmiştir. Şekil 3.17 ve 3.18'de görüldüğü üzere uygun karar ağacı sayısı belirlenmiştir.



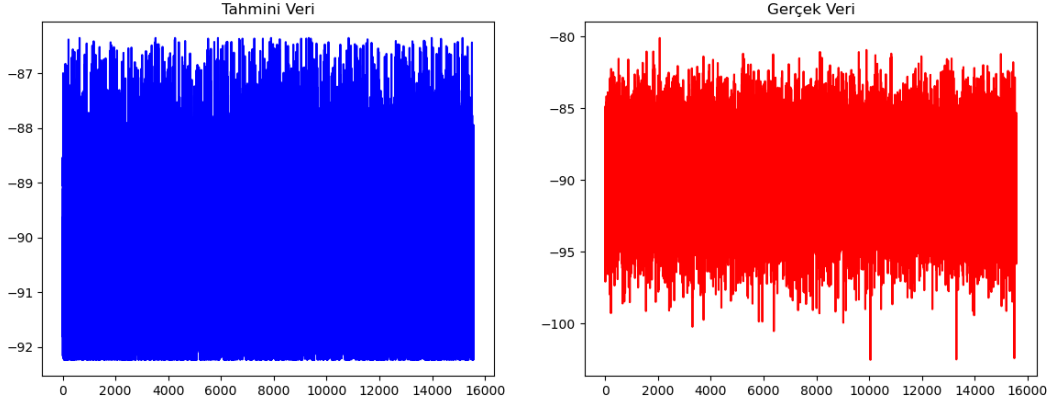
Şekil 3.17 Rassal orman algoritması 800 MHz'de karar ağacı sayısı



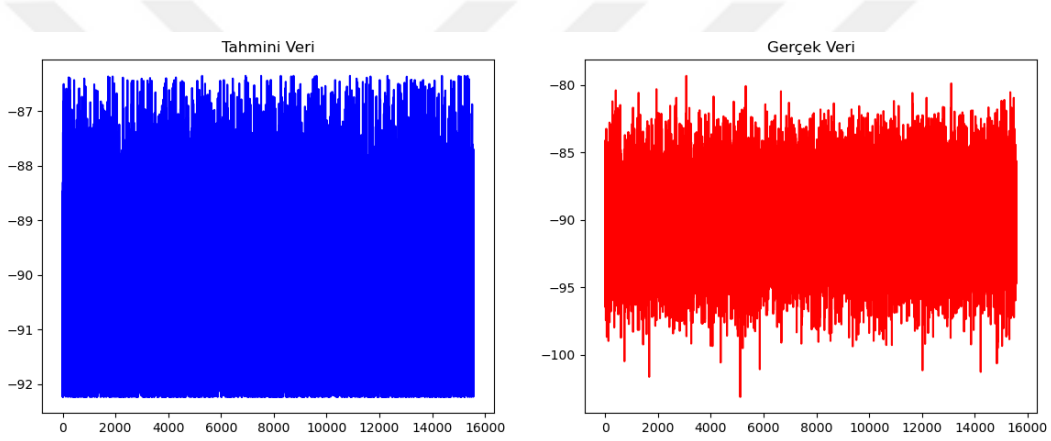
Şekil 3.18 Rassal orman algoritması 838 MHz'de karar ağacı sayısı

Belirlenen optimum karar ağacı sayısı ile Rassal Orman algoritması ile 800 MHz frekansındaki kanalın gelecek durumu ile ilgili kestirimler yapılmıştır.

Rassal Orman algoritması test ve doğrulama kestirimleri şekil 3.19 ve 3.20’de gösterilmiştir.



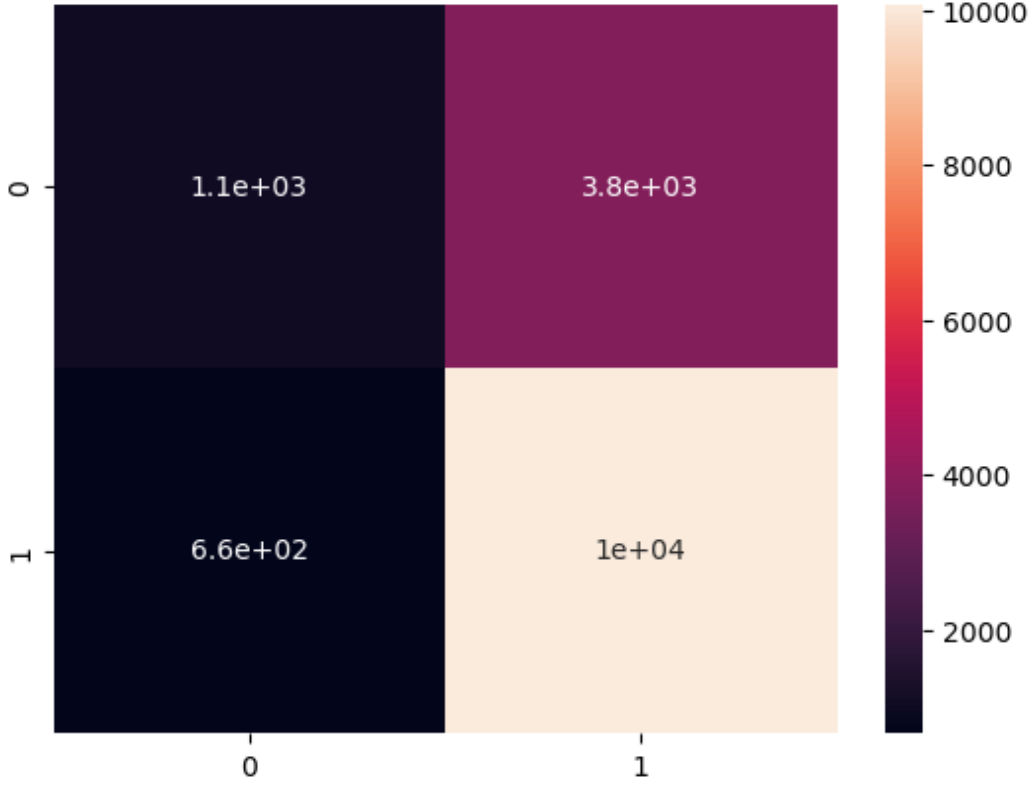
Şekil 3.19 800 MHz’de rassal orman algoritması test tahminleri



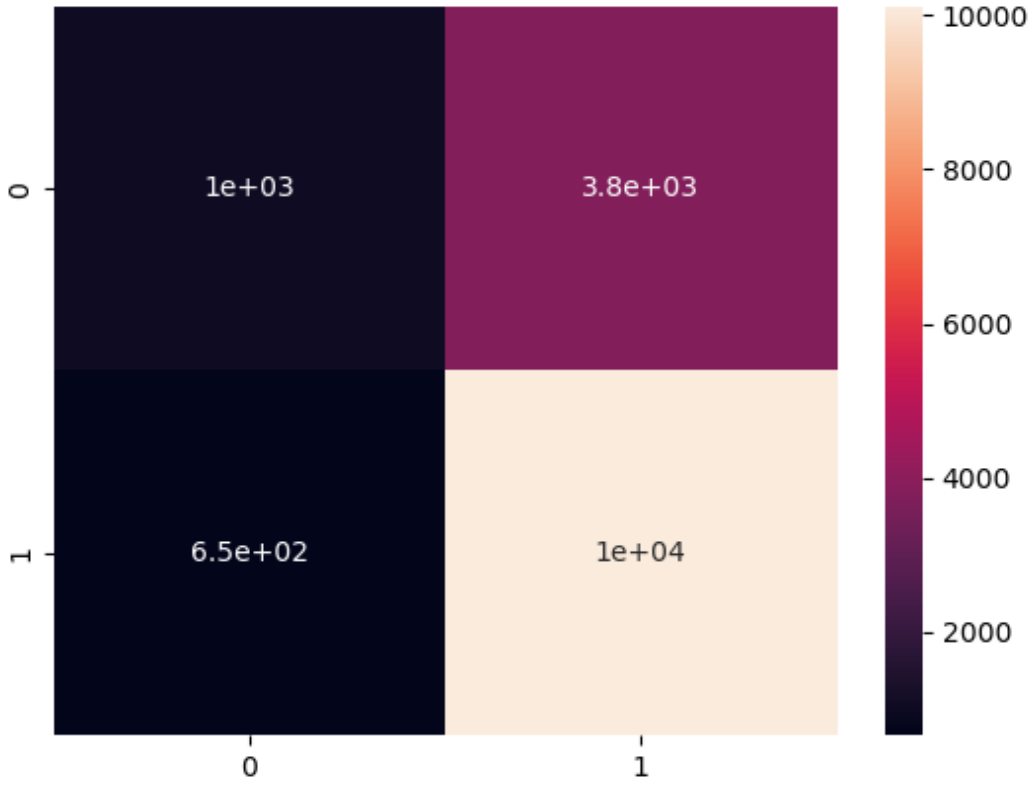
Şekil 3.20 800 MHz’de rassal orman algoritması doğrulama tahminleri

Kestirilen değerler eşik değerinden geçirilerek, kanalın mevcut durumunu (0 veya 1) belirlenir, burada 0 spektrumun boş olduğunu, 1 ise dolu olduğunu gösterir. Şekil 3.21’de gösterildiği üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 4801’dir. 1055 adet boş olan durum boş olarak kestirilmiştir, 3746 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 10747’dir. 10147 adet dolu olan durum dolu olarak kestirilmiştir, 660 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 1715 durumun 1055 durumu boş 660 durumu ise dolu olarak kestirilmiştir. Algoritmanın dolu olarak kestirdiği 13800’e yakın durumun 10000’e yakın durumu dolu 3800’e yakın durumu ise boş olarak kestirilmiştir.

Şekil 3.22’de gösterildiği üzere doğrulama veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı yaklaşık olarak 4800’dir. 1000 adet boş olan durum boş olarak kestirilmiştir, 3800 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı yaklaşık olarak 10750’dir.



Şekil 3.21 800 MHz'de rassal orman algoritması test verisi hata matrisi

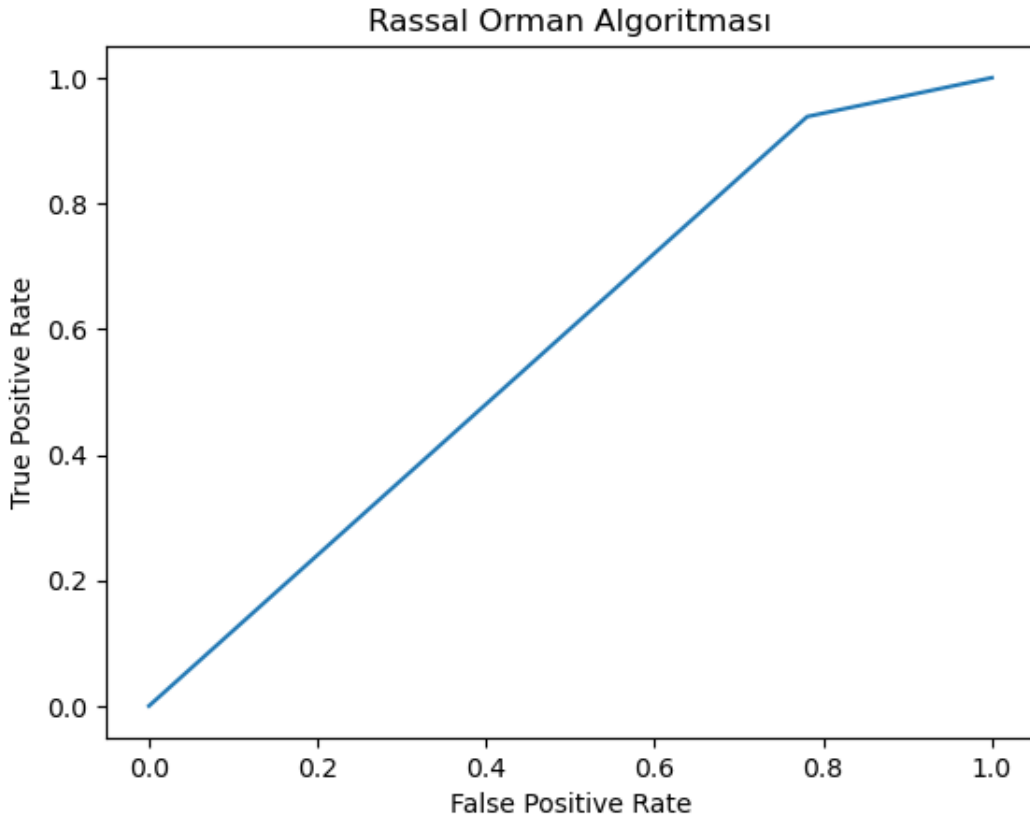


Şekil 3.22 800 MHz'de rassal orman algoritması doğrulama verisi hata matrisi

10000 adet dolu olan durum dolu olarak kestirilmiş, 650 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği yaklaşık 1650 durumun 1000 durumu boş 650 durumu ise dolu olarak kestirilmiştir. Algoritmanın dolu olarak kestirdiği yaklaşık 13800 durumun 10000 durumu dolu 3800 durumu ise boş olarak kestirilmiştir.

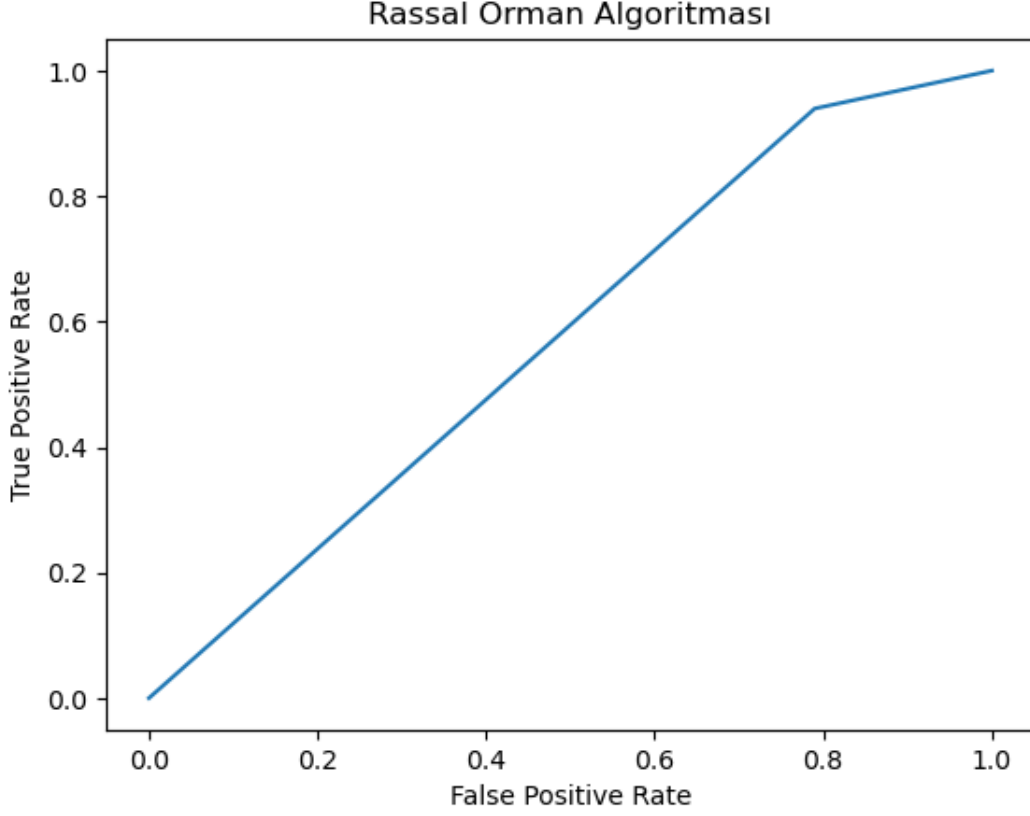
Algoritmanın başarı oranı test veri setinde %72, doğrulama veri setinde %71'dür. Kanalın boş ve dolu olma durumu dengeli bir dağılıma sahip olmadığı için sadece başarıma bakılması eksik olacaktır. Bu sebepten ayrıca ROC eğrisi de incelenmelidir.

Şekil 3.23'de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.57 olarak hesaplanmıştır. Şekil 3.24'de gösterildiği gibi doğrulama verisinin ayırt edicilik değeri olan AUC değeri 0.58 olarak hesaplanmıştır.



Şekil 3.23 800 MHz'de rassal orman algoritması test verisi ROC eğrisi

800 MHz frekanslı kanalın Test verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %61, duyarlılığı %22, f1 score değeri ise %32, kanalın dolu olma durumunun kesinliği %73, duyarlılığı %94, f1 score değeri ise %82 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %22'si kestirilmiş, boş olarak tahmin edilen durumların %61'i gerçekten boş olma durumudur.



Şekil 3.24 800 MHz’de rassal orman algoritması doğrulama verisi ROC eğrisi

Kanalın gerçekte dolu olan durumlarının %94’ü kestirilmiş, dolu olarak tahmin edilen durumların %73’ü gerçekten dolu olma durumudur.

Tablo 3.6 800 MHz’de rassal orman algoritma test verisi raporu

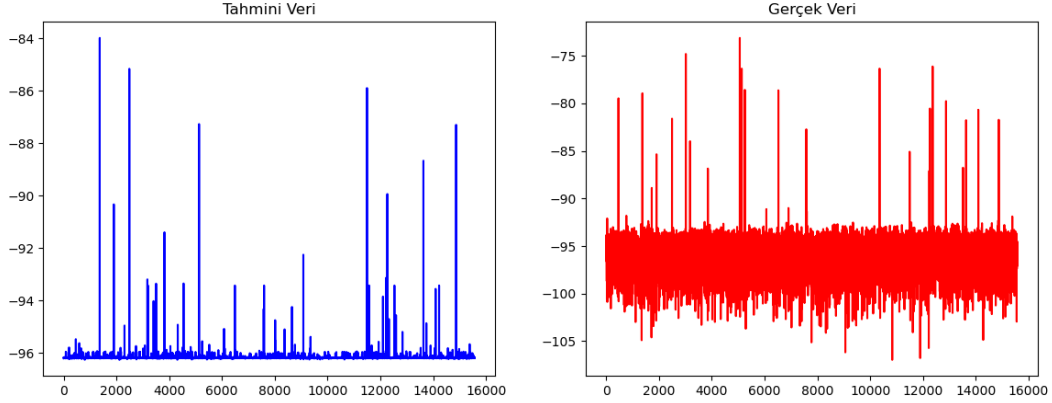
Kanalın Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.61	0.22	0.32	4801
Dolu	0.73	0.94	0.82	10747

800 MHz frekanslı kanalın doğrulama verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %61, duyarlılığı %21, f1 score değeri ise %31, kanalın dolu olma durumunun kesinliği %73, duyarlılığı %94, f1 score değeri ise %82 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %21’i kestirilmiş, boş olarak tahmin edilen durumların %61’i gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %94’ü kestirilmiş, dolu olarak tahmin edilen durumların %73’ü gerçekten dolu olma durumudur.

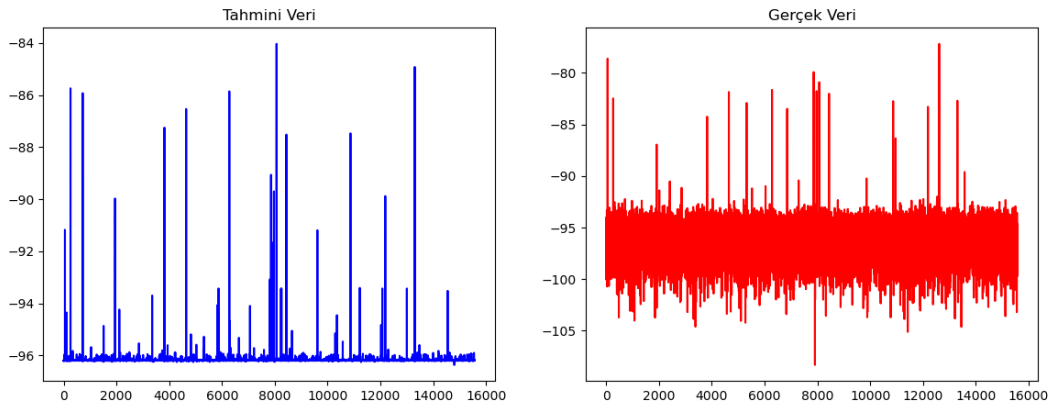
Tablo 3.7 800 MHz’de rassal orman algoritma doğrulama verisi raporu

Kanalın Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.61	0.21	0.31	4801
Dolu	0.73	0.94	0.82	10747

Belirlenen optimum karar ağacı sayısı ile Rassel Orman algoritması ile 800 MHz frekansındaki kanalın gelecek durumu ile ilgili kestirimler yapılmıştır. Rassel Orman algoritması test ve doğrulama kestirimleri şekil 3.25 ve 3.26'de gösterilmiştir.



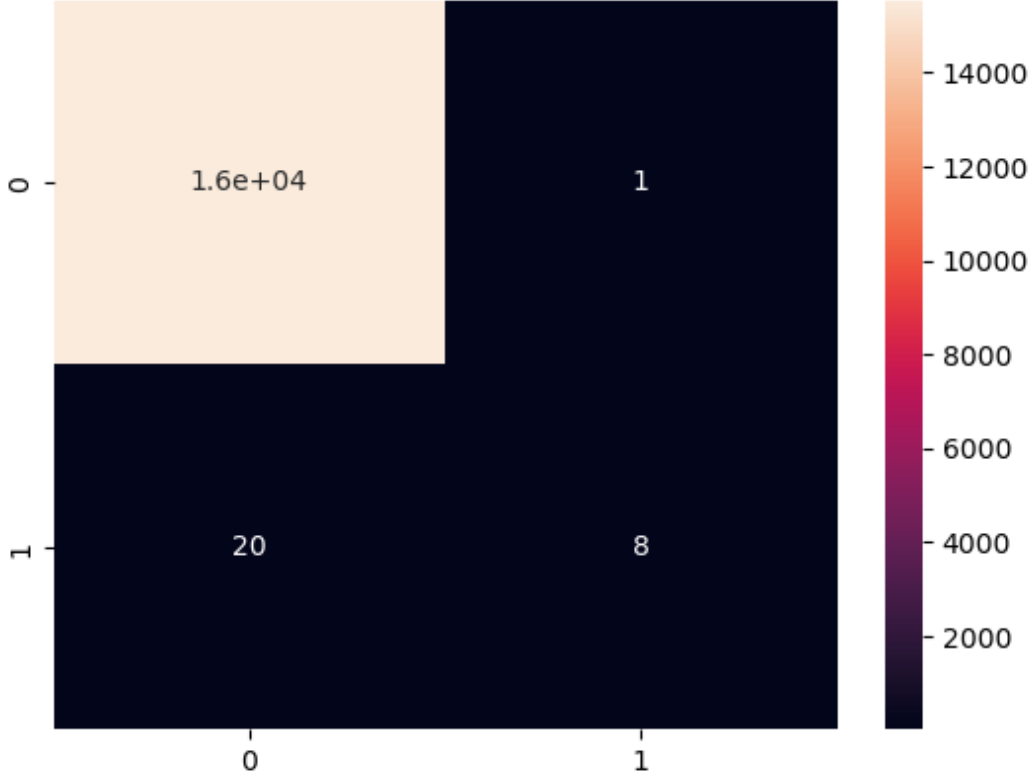
Şekil 3.25 838 MHz'de rassel orman algoritması test tahminleri



Şekil 3.26 838 MHz'de rassel orman algoritması doğrulama verisi tahminleri

Şekil 3.27'de gösterildiği üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 16001'dir. 16000 adet boş olan durum boş olarak kestirilmiş, 1 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 28'dir. 8 adet dolu olan durum dolu olarak kestirilmiş, 20 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 16020 durumun 16000 durumu boş 20 durumu ise doludur. Algoritmanın dolu olarak kestirdiği 9 durumun 8 durumu dolu 1 durumu ise boş olarak kestirilmiştir.

Şekil 3.28'de gösterildiği üzere doğrulama veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 16004'dir. 16000 adet boş olan durum boş olarak kestirilmiş, 4 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 26'dir. 12 adet dolu olan durum dolu olarak kestirilmiş, 14 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 16014 durumun 16000 durumu boş 14 durumu ise doludur.



Şekil 3.27 838 MHz’de rassal orman algoritması test verisi hata matrisi

Algoritmanın dolu olarak kestirdiği 16 durumun 12 durumu dolu 4 durumu ise boştur.

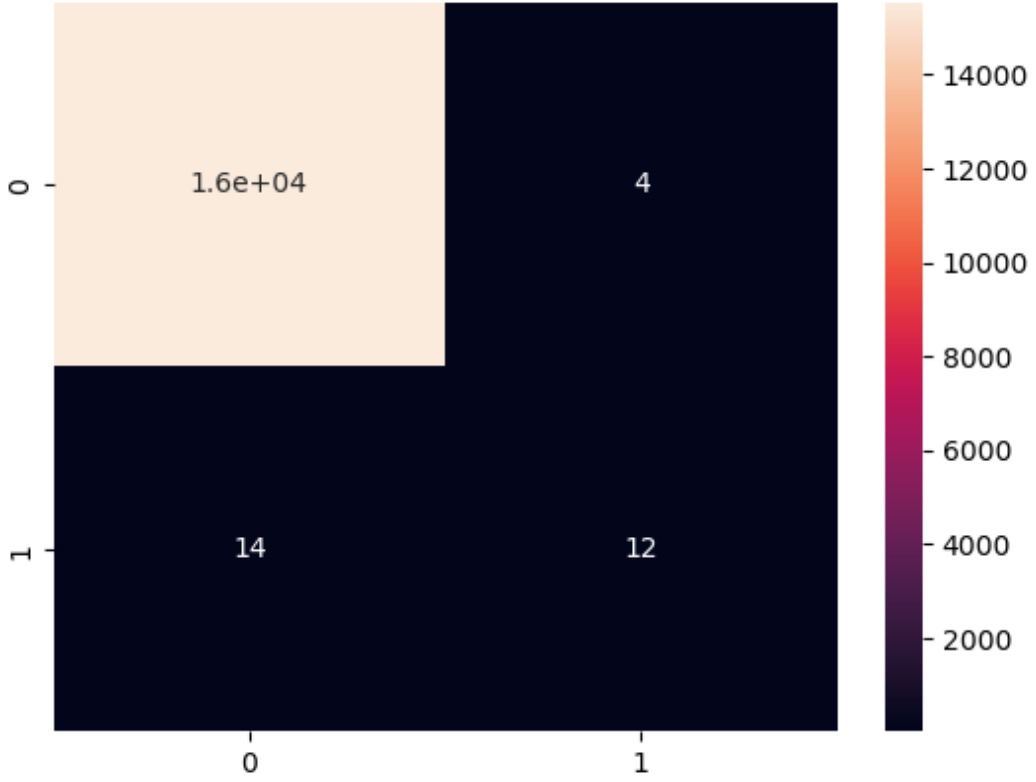
Şekil 3.29’de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.73 olarak hesaplanmıştır.

Şekil 3.30’de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.64 olarak hesaplanmıştır.

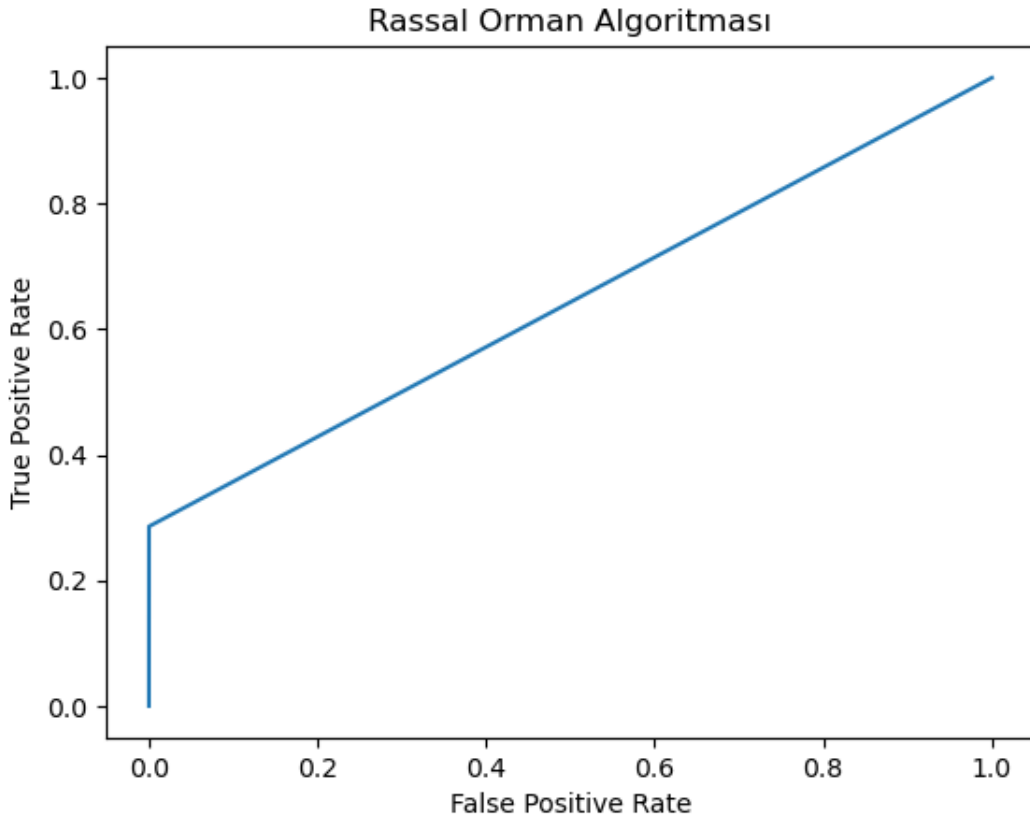
838 MHz frekanslı kanalın Test verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %100, duyarlılığı %100, f1 score değeri ise %100, kanalın dolu olma durumunun kesinliği %89, duyarlılığı %29, f1 score değeri ise %43 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %100’ü kestirilmiş, boş olarak tahmin edilen durumların %100’ü gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %29’u kestirilmiş, dolu olarak tahmin edilen durumların %89’si gerçekten dolu olma durumudur.

Tablo 3.8 838 MHz’de rassal orman algoritma test verisi raporu

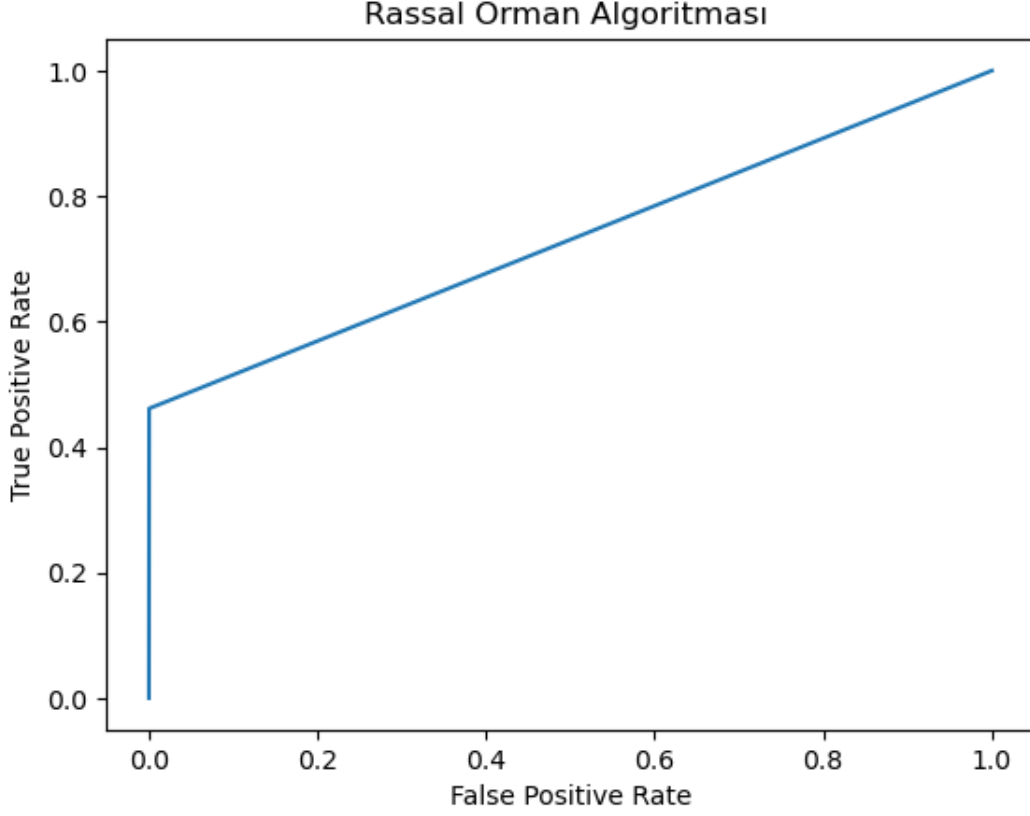
Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.99	0.99	0.99	15547
Dolu	0.89	0.29	0.43	28



Şekil 3.28 838 MHz’de rassal orman algoritması doğrulama verisi hata matrisi



Şekil 3.29 838 MHz’de rassal orman algoritması test verisi ROC eğrisi



Şekil 3.30 838 MHz’de rassal orman algoritması doğrulama verisi ROC eğrisi

838 MHz frekanslı kanalın doğrulama verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %100, duyarlılığı %100, f1 score değeri ise %100, kanalın dolu olma durumunun kesinliği %75, duyarlılığı %46, f1 score değeri ise %57 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %100’ü kestirilmiş, boş olarak tahmin edilen durumların %100’ü gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %46’sı kestirilmiş, dolu olarak tahmin edilen durumların %75’i gerçekten dolu olma durumudur.

Tablo 3.9 838 MHz’de rassal orman algoritma doğrulama verisi raporu

Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.99	0.99	0.99	15548
Dolu	0.75	0.46	0.57	26

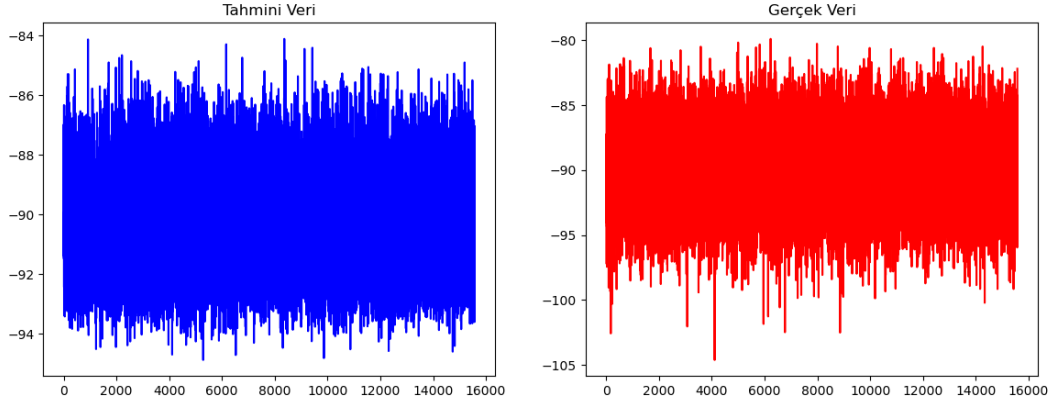
3.4 Uzun Kısa Süreli Bellek (LSTM)

Uzun Kısa Süreli Bellek (LSTM) algoritmasında veri seti düzenlenmesi yapılmıştır. Tahmin edilen $t + 1$. durum kendinden önceki 5 saniye (durum) ile ilişkilendirilmiştir. $t - 4$. andan $t - 4$. ana kadar olan durumlar ile ilişkilendirilmiştir.

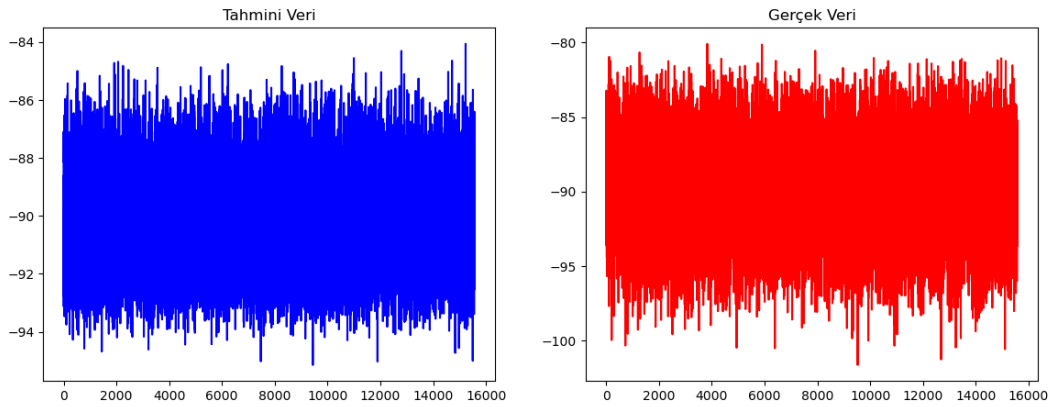
Veri seti eğitim, test ve doğrulama seti olarak üçe ayrılmıştır. Veri setinin %60’nın

eđitim seti, %20'si test seti ve %20'si de dođrulama seti olarak ayrılmıřtır.

Uzun Kısa Süreli Bellek (LSTM) 800 MHz frekansındaki kanalın test ve dođrulama kestirimleri řekil 3.31 ve 3.32'de gösterilmiřtir.



Şekil 3.31 800 MHz'de LSTM algoritması test tahminleri

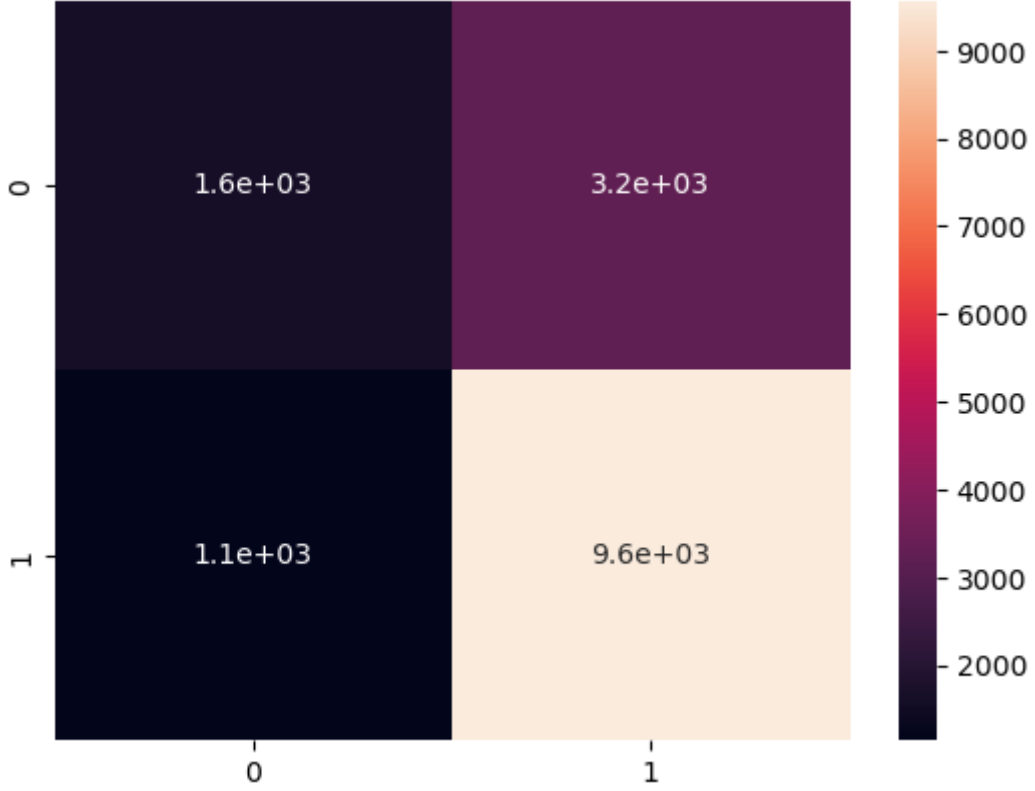


Şekil 3.32 800 MHz'de LSTM algoritması dođrulama tahminleri

Kestirilen deđerler eřik deđerinden geđirilerek, kanalın mevcut durumunu (0 veya 1) belirlenir, burada 0 spektrumun boř olduđunu, 1 ise dolu olduđunu gösterir.

Şekil 3.33'de gösterildiđi üzere test veri setinde belirlenen kanal durumlarının kanalın boř olduđu durum sayısı 4863'dir. Yaklařık olarak 1600 adet boř olan durum boř olarak kestirilmiř, 3200 adeti ise dolu olarak kestirilmiřtir. Kanalın dolu olduđu durum sayısı 10713'dir. Yaklařık olarak 9600 adet dolu olan durum dolu olarak kestirilmiř, 1100 adeti ise boř olarak kestirilmiřtir. Algoritmanın boř olarak kestirdiđi 2700 durumun 1600 durumu boř 1100 durumu ise doludur. Algoritmanın dolu olarak kestirdiđi 12800 durumun 9600 durumu dolu 3200 durumu ise bořtur.

Şekil 3.34'de gösterildiđi üzere dođrulama veri setinde belirlenen kanal durumlarının kanalın boř olduđu durum sayısı 4712'dir. Yaklařık olarak 1600 adet boř olan durum boř olarak kestirilmiř, 3200 adeti ise dolu olarak kestirilmiřtir.



Şekil 3.33 800 MHz’de LSTM algoritması test verisi hata matrisi

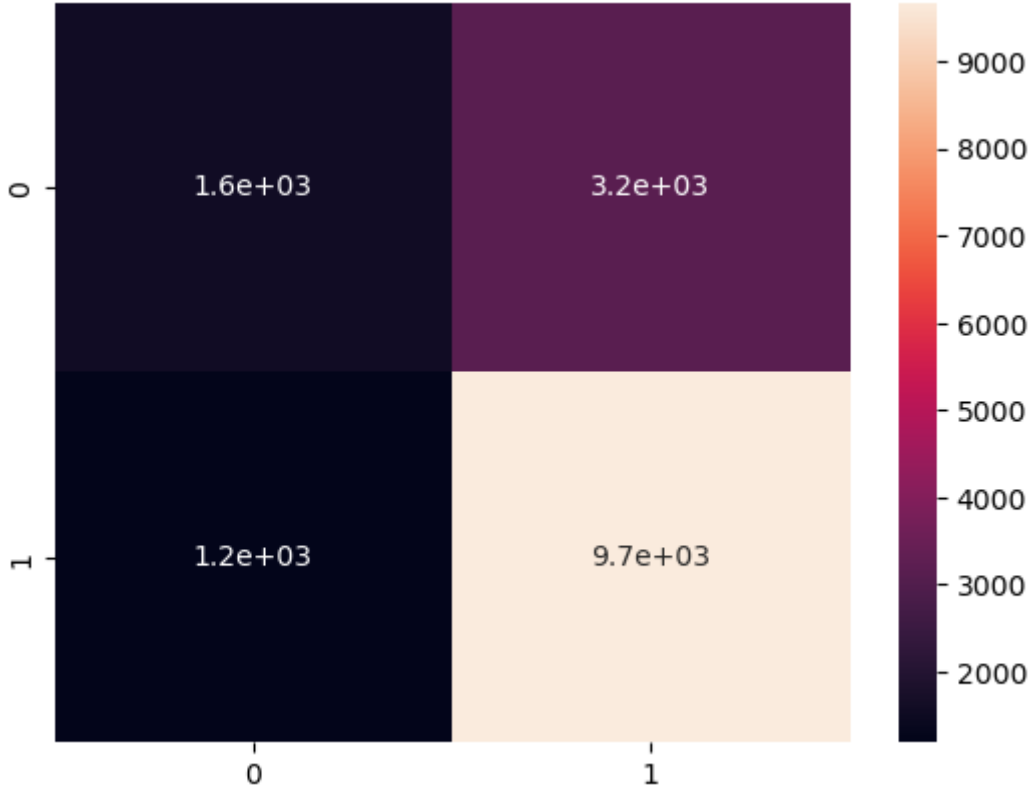
Kanalın dolu olduğu durum sayısı 10863’dir. Yaklaşık olarak 9700 adet dolu olan durum dolu olarak kestirilmiş, 1200 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 2800 durumun 1600 durumu boş 1200 durumu ise doludur. Algoritmanın dolu olarak kestirdiği 12900 durumun 9700 durumu dolu 3200 durumu ise boştur.

Algoritmanın başarımlı oranı test veri setinde %72, doğrulama veri setinde %72’dir. Kanalın boş ve dolu olma durumu dengeli bir dağılıma sahip olmadığı için sadece başarımlı bakılması eksik olacaktır. Bu sebepten ayrıca ROC eğrisi de incelenmelidir.

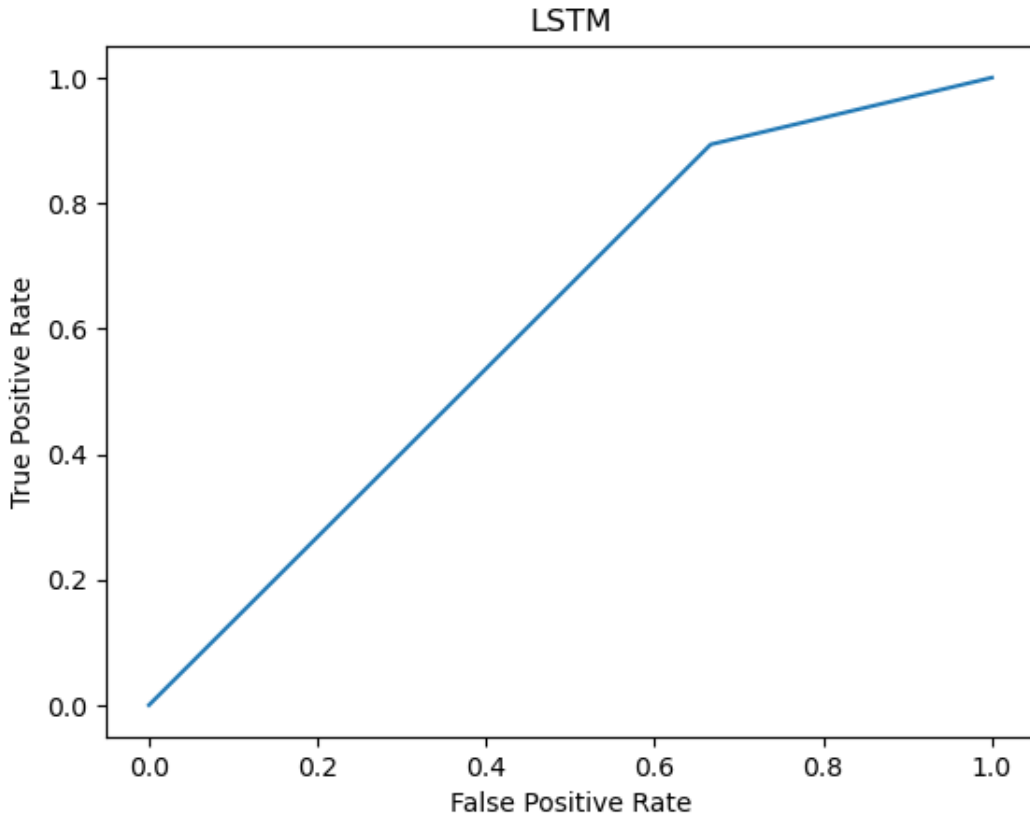
Şekil 3.35’de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.61 olarak hesaplanmıştır.

Şekil 3.36’de gösterildiği gibi doğrulama verisinin ayırt edicilik değeri olan AUC değeri 0.61 olarak hesaplanmıştır.

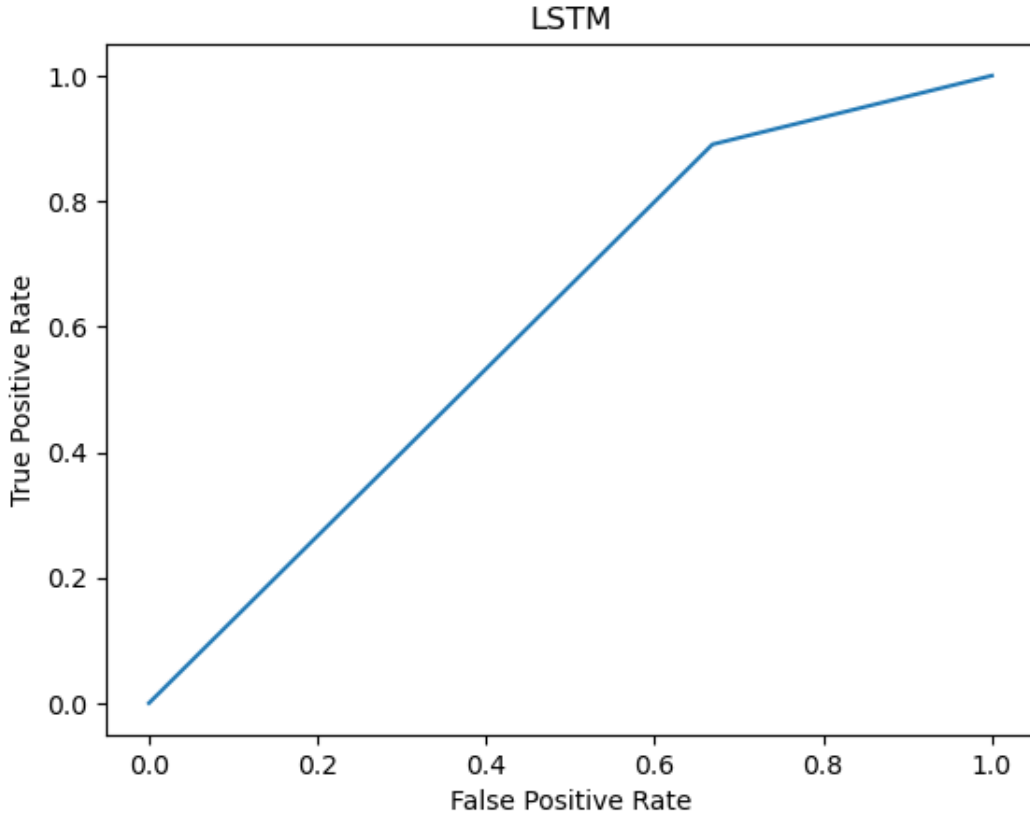
Şekil 3.37’de LSTM algoritmasının model kaybının grafiği gösterilmiştir. Belli bir epoch değerinden sonra kayıp değeri sabitleşmiştir. Bu yüzden epoch sayısını arttırmak algoritmanın doğruluğunu arttırmamaktadır.



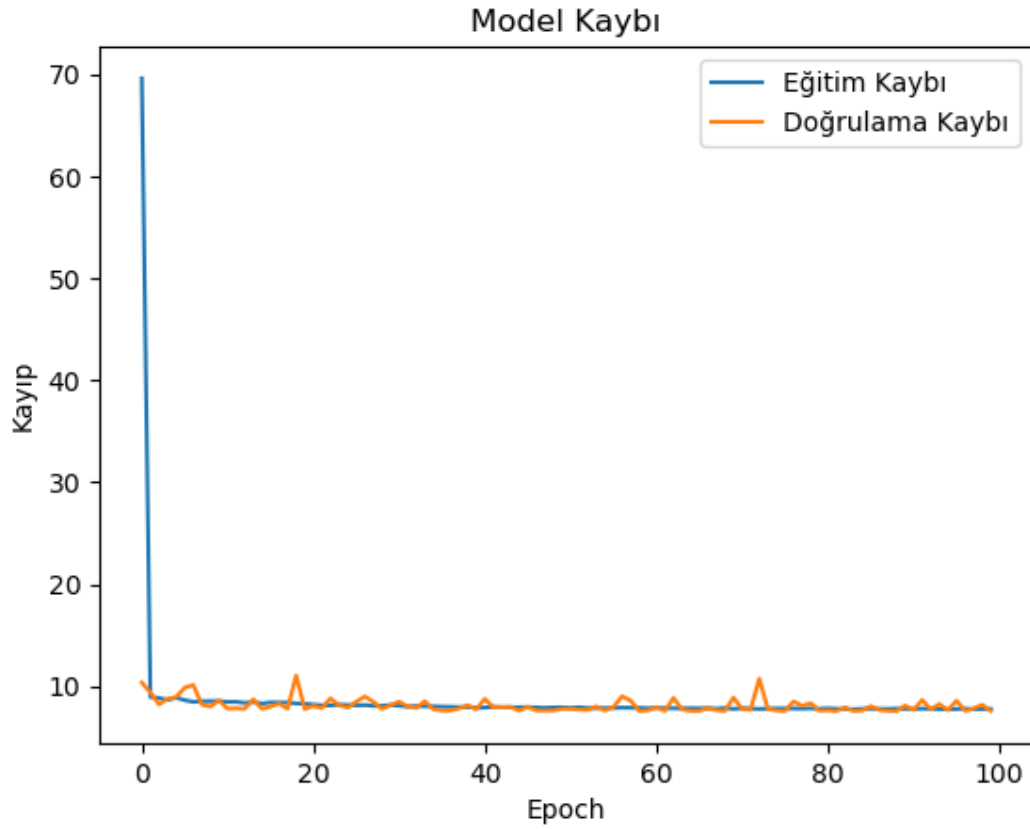
Şekil 3.34 800 MHz’de LSTM algoritması doğrulama verisi hata matrisi



Şekil 3.35 800 MHz’de LSTM algoritması test verisi ROC eğrisi

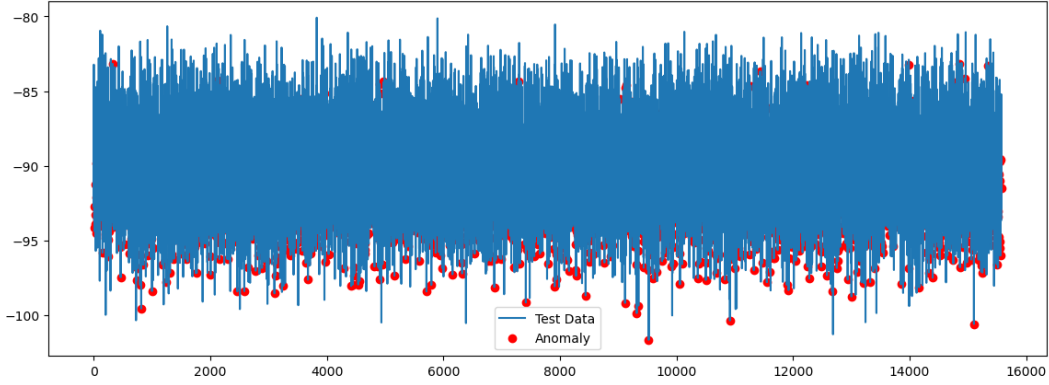


Şekil 3.36 800 MHz'de LSTM algoritması doğrulama verisi ROC eğrisi



Şekil 3.37 800 MHz'de LSTM algoritması kayıp grafiği

Şekil 3.38’de LSTM modelinin kanalın boş olduğu durumları boş olarak tahminlediği noktaları göstermektedir.



Şekil 3.38 800 MHz’de LSTM tahminleri

800 MHz frekanslı kanalın Test verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %59, duyarlılığı %33, f1 score değeri ise %43, kanalın dolu olma durumunun kesinliği %75, duyarlılığı %89, f1 score değeri ise %81 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %33’ü kestirilmiş, boş olarak tahmin edilen durumların %59’u gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %89’u kestirilmiş, dolu olarak tahmin edilen durumların %75’i gerçekten dolu olma durumudur.

Tablo 3.10 800 MHz’de LSTM algoritma test verisi raporu

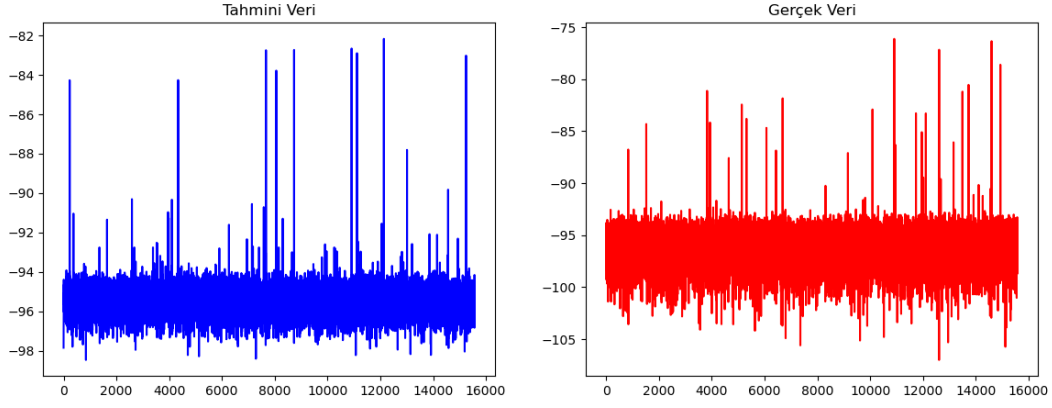
Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.59	0.33	0.43	4863
Dolu	0.75	0.89	0.81	10713

800 MHz frekanslı kanalın doğrulama verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %57, duyarlılığı %33, f1 score değeri ise %42, kanalın dolu olma durumunun kesinliği %75, duyarlılığı %89, f1 score değeri ise %82 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %33’ü kestirilmiş, boş olarak tahmin edilen durumların %57’si gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %89’u kestirilmiş, dolu olarak tahmin edilen durumların %75’i gerçekten dolu olma durumudur.

Tablo 3.11 800 MHz’de LSTM algoritma doğrulama verisi raporu

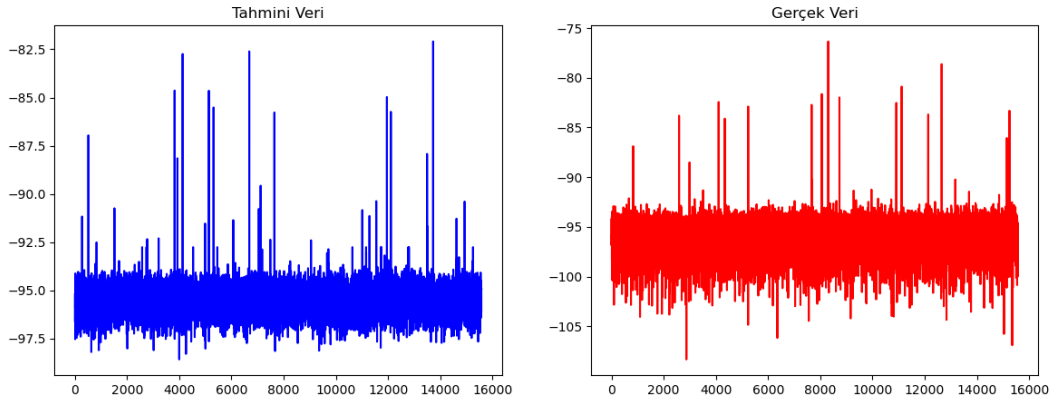
Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.57	0.33	0.42	4712
Dolu	0.75	0.89	0.82	10863

Uzun Kısa Süreli Bellek (LSTM) 800 MHz frekansındaki kanalın test ve doğrulama kestirimleri şekil 3.40 ve 3.39’de gösterilmiştir.



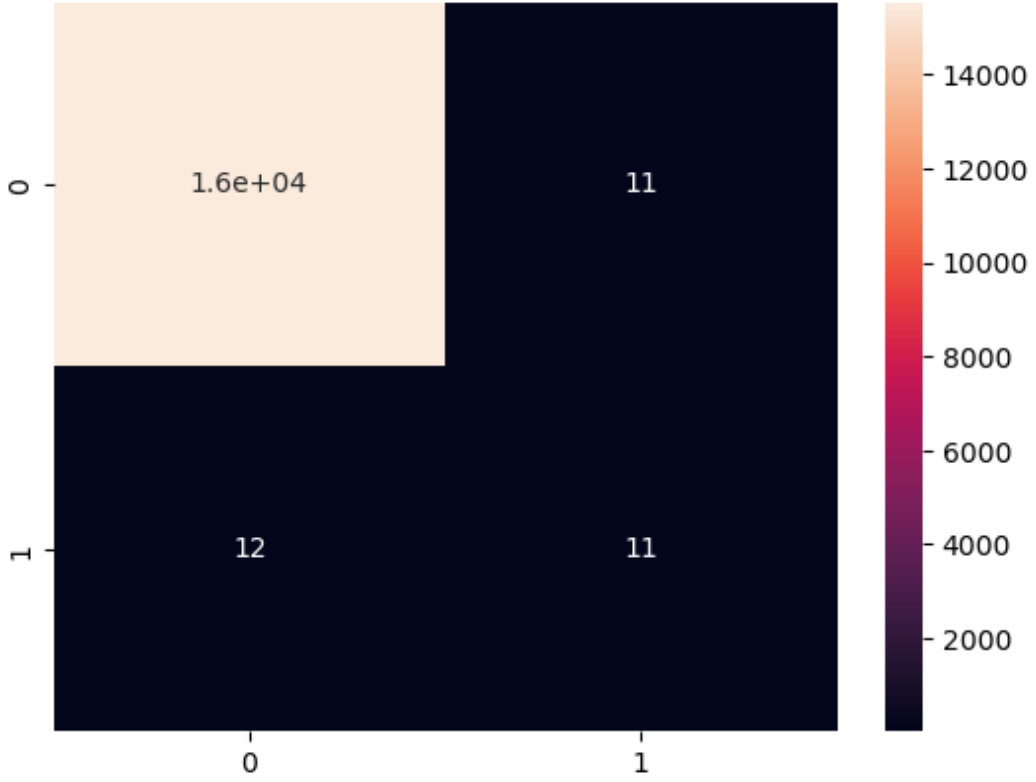
Şekil 3.39 838 MHz'de LSTM algoritması doğrulama verisi tahminleri

Şekil 3.41'de gösterildiği üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 15553'dir. 15542 adet boş olan durum boş olarak kestirilmiş, 11 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 23'dür. 11 adet dolu olan durum dolu olarak kestirilmiş, 12 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 15554 durumun 15542 durumu boş 12 durumu ise doludur. Algoritmanın dolu olarak kestirdiği 22 durumun 11 durumu dolu 11 durumu ise boştur.

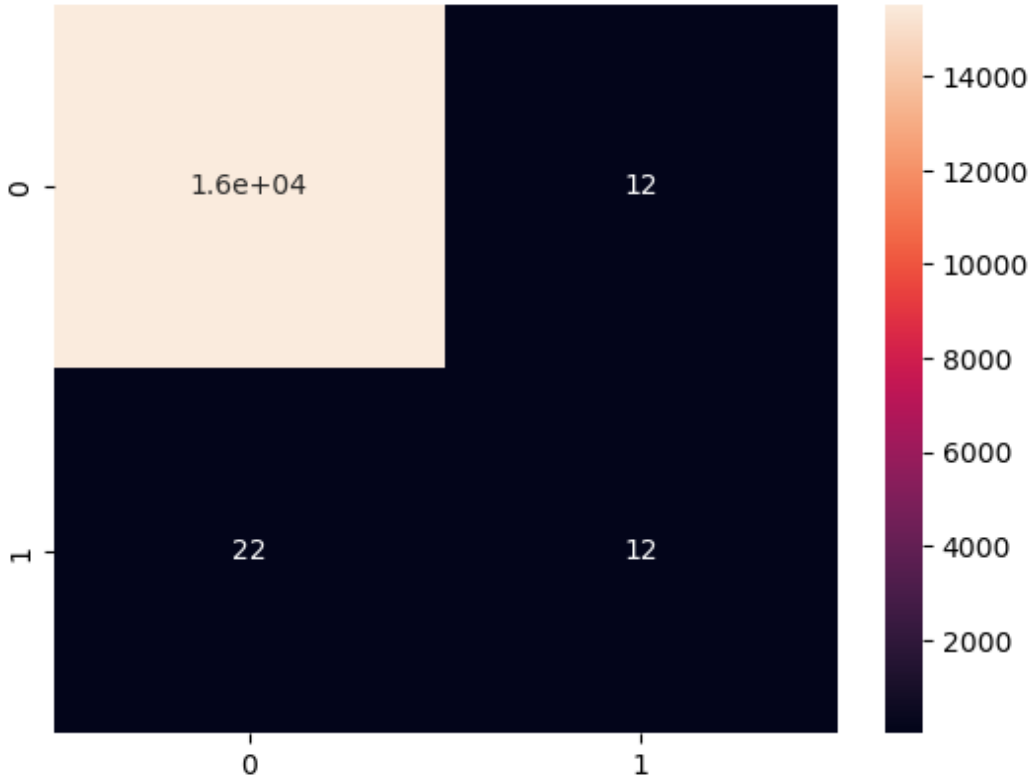


Şekil 3.40 838 MHz'de LSTM algoritması test tahminleri

Şekil 3.42'de gösterildiği üzere test veri setinde belirlenen kanal durumlarının kanalın boş olduğu durum sayısı 15541'dir. 15529 adet boş olan durum boş olarak kestirilmiş, 12 adeti ise dolu olarak kestirilmiştir. Kanalın dolu olduğu durum sayısı 34'dür. 12 adet dolu olan durum dolu olarak kestirilmiş, 22 adeti ise boş olarak kestirilmiştir. Algoritmanın boş olarak kestirdiği 15551 durumun 15529 durumu boş 22 durumu ise doludur. Algoritmanın dolu olarak kestirdiği 24 durumun 12 durumu dolu 12 durumu ise boştur. Algoritmanın başarımları test veri setinde %100, doğrulama veri setinde %100'dür. Kanalın boş ve dolu olma durumu dengeli bir dağılıma sahip olmadığı için sadece başarımları bakılması eksik olacaktır.

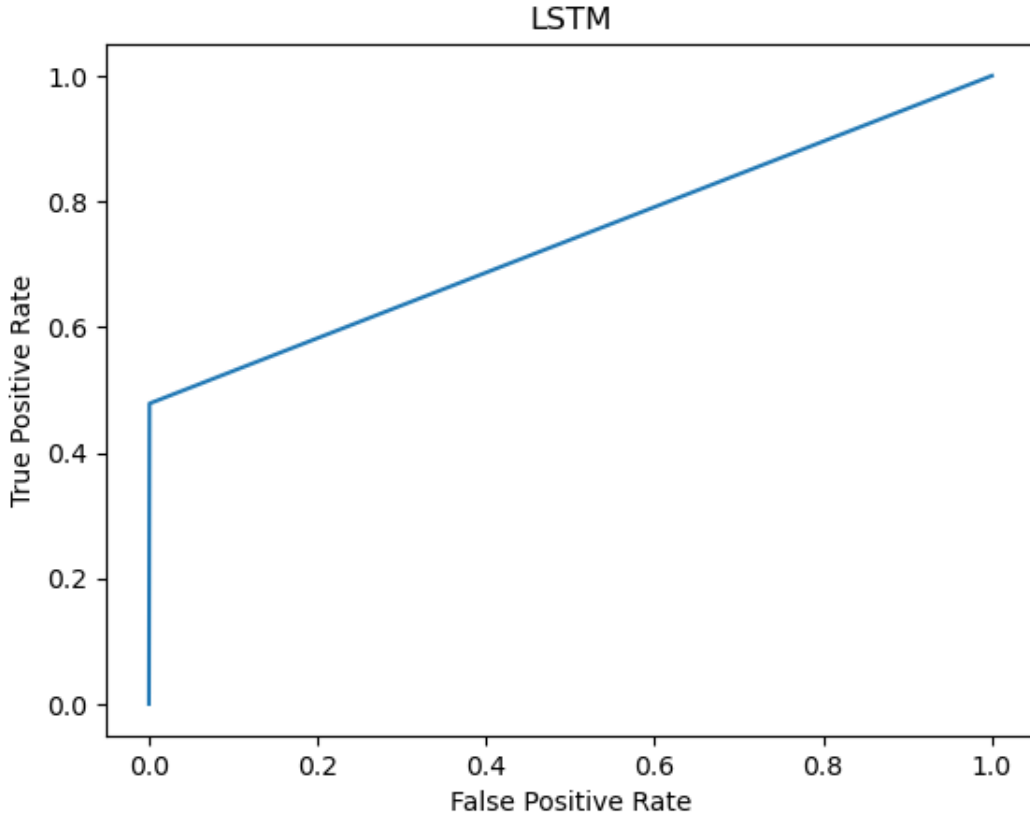


Şekil 3.41 838 MHz’de LSTM algoritması test verisi hata matrisi



Şekil 3.42 838 MHz’de LSTM algoritması doğrulama verisi hata matrisi

Şekil 3.43’de gösterildiği gibi test verisinin ayırt edicilik değeri olan AUC değeri 0.67 olarak hesaplanmıştır.



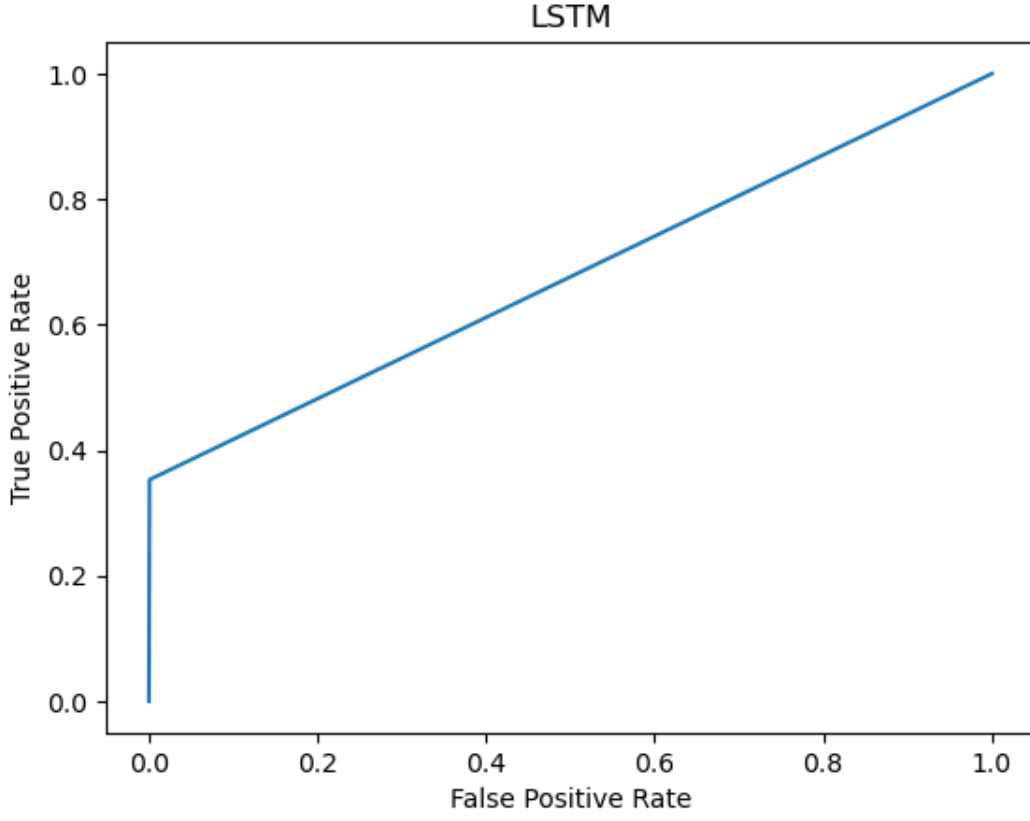
Şekil 3.43 838 MHz’de LSTM algoritması test verisi ROC eğrisi

Şekil 3.44’de gösterildiği gibi doğrulama verisinin ayırt edicilik değeri olan AUC değeri 0.74 olarak hesaplanmıştır.

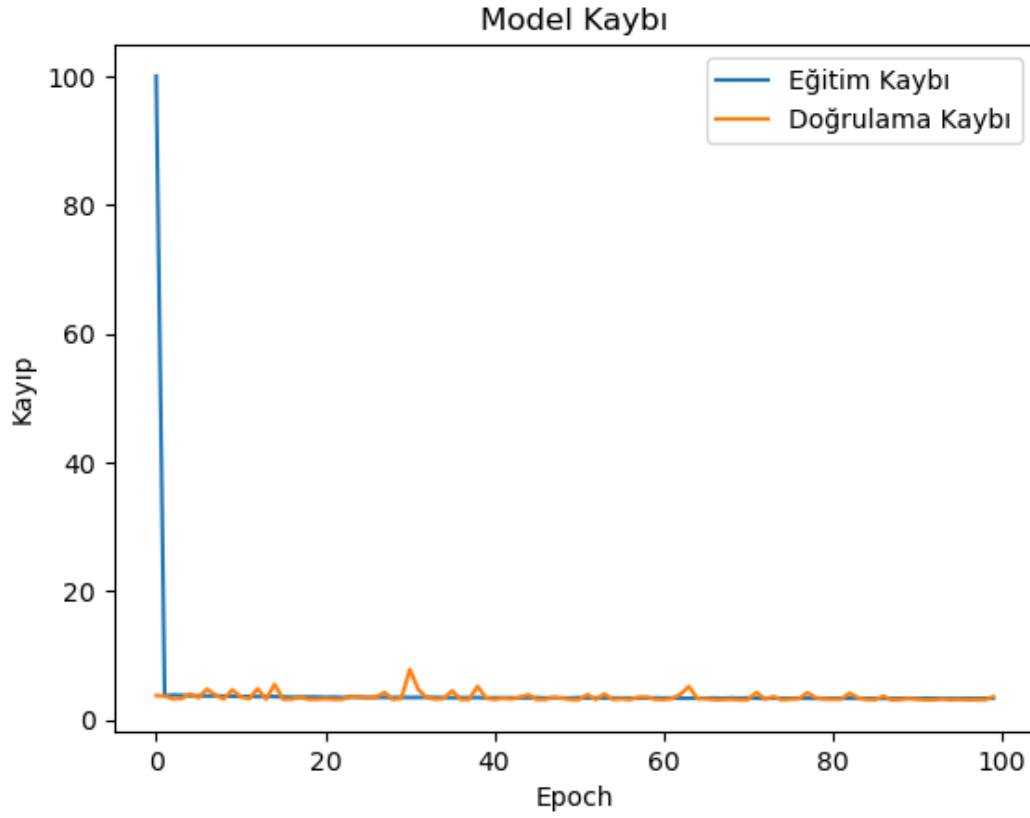
Şekil 3.45’de LSTM algoritmasının model kaybının grafiği gösterilmiştir. Belli bir epoch değerinden sonra kayıp değeri sabitleşmiştir. Bu yüzden epoch sayısını arttırmak algoritmanın doğruluğunu arttırmamaktadır.

838 MHz frekanslı kanalın Test verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %99, duyarlılığı %99, f1 score değeri ise %99, kanalın dolu olma durumunun kesinliği %50, duyarlılığı %48, f1 score değeri ise %49 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %99’ü kestirilmiş, boş olarak tahmin edilen durumların %99’ü gerçekte boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %48’i kestirilmiş, dolu olarak tahmin edilen durumların %50’si gerçekte dolu olma durumudur.

Şekil 3.46’de LSTM modelinin kanalın dolu olduğu durumları dolu olarak tahminlediği noktaları göstermektedir.



Şekil 3.44 838 MHz’de LSTM algoritması doğrulama verisi ROC eğrisi

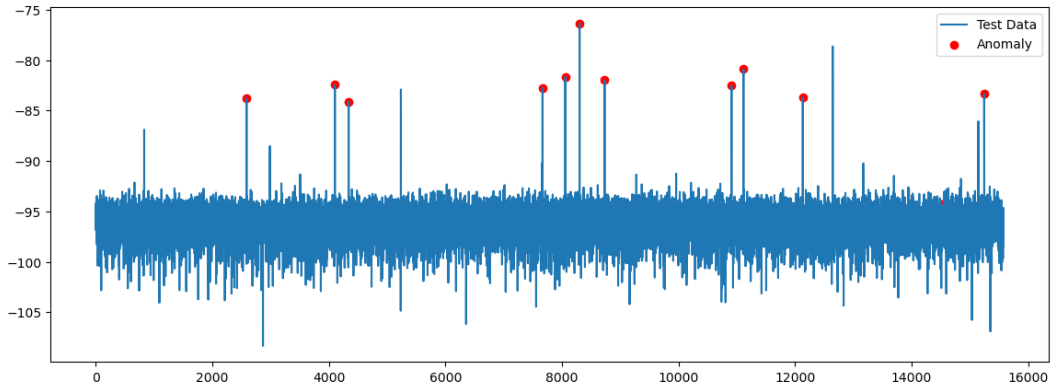


Şekil 3.45 838 MHz’de LSTM algoritması kayıp grafiği

Tablo 3.12 838 MHz'de LSTM algoritma test verisi raporu

Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.99	0.99	0.99	15553
Dolu	0.50	0.48	0.49	23

838 MHz frekanslı kanalın Test verisinin sonuçları incelendiğinde, kanalın boş olma durumunun kesinliği %99, duyarlılığı %99, f1 score değeri ise %99, kanalın dolu olma durumunun kesinliği %50, duyarlılığı %48, f1 score değeri ise %49 olarak hesaplanmıştır. Kanalın gerçekte boş olan durumlarının %99'ü kestirilmiş, boş olarak tahmin edilen durumların %99'ü gerçekten boş olma durumudur. Kanalın gerçekte dolu olan durumlarının %45'i kestirilmiş, dolu olarak tahmin edilen durumların %50'si gerçekten dolu olma durumudur.



Şekil 3.46 838 MHz'de LSTM tahminleri

Tablo 3.13 838 MHz'de LSTM algoritma doğrulama verisi raporu

Kanal Durumu	Kesinlik	Duyarlılık	f1 score	Tahmin Sayısı
Boş	0.99	0.99	0.99	15541
Dolu	0.50	0.45	0.41	34

4

SONUÇ

Bu çalışmada, spektrum kullanımının kestirimi üzerine makine öğrenmesi yöntemleri kullanılarak kanalın dolu ya da boş olma durumunun sınıflandırılması incelenmiştir. Günümüzde popüler olan makine öğrenmesi ve derin öğrenme algoritmaları, regresyon ve sınıflandırma problemlerinde oldukça başarılı performans göstermektedir.

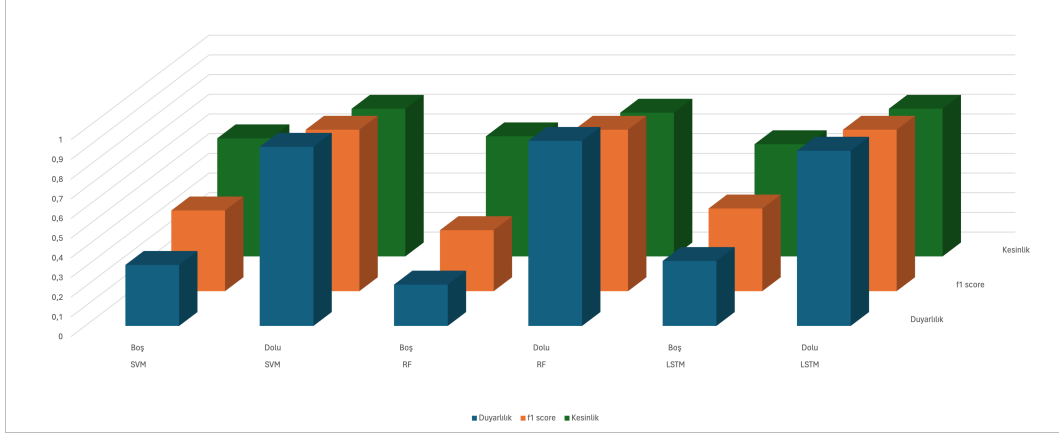
Çalışmada, SVM, RF ve LSTM algoritmaları karşılaştırılmıştır. 800 MHz frekansındaki yoğun kullanılan kanalda SVM ve LSTM algoritmaları aynı başarı metriklerini gösterse de, 838 MHz frekansındaki az yoğun kullanılan kanalda ise LSTM algoritmasının ayırt edicilik özelliği daha yüksek çıkmıştır. LSTM algoritmasının sonuçlarının SVM ve RF algoritmalarına kıyasla daha iyi olmasının nedeni, veri setinin zaman serisi olması ve problemin kendinden önceki durumlar ile korelasyonunun olmasıdır.

Bu çalışma, spektrum kullanımının etkin bir şekilde yönetilmesi ve optimizasyonu için makine öğrenmesi yöntemlerinin potansiyelini göstermektedir. Özellikle zaman serisi verileriyle çalışan problemler için LSTM algoritmasının üstün performansı, bu tür problemlerde tercih edilmesini önermektedir.

Tablo 4.1 ve Şekil 4.1'de görüldüğü üzere üç algortmada iyi sonuçlar vermemiştir. SVM ve LSTM algoritmaları aynı başarı metriklerini gösterse de kötü sonuç vermiştir. Doğruluk oranları %70'i geçmesine rağmen sınıflandırma özellikleri %61'i geçemeyerek düşük çıkmıştır.

Tablo 4.1 800 MHz'de algoritmaların doğrulama veri seti için doğruluk ve AUC değerleri

Algoritmalar	Doğruluk	AUC Değeri
SVM	%73	%61
RF	%71	%57
LSTM	%72	%61

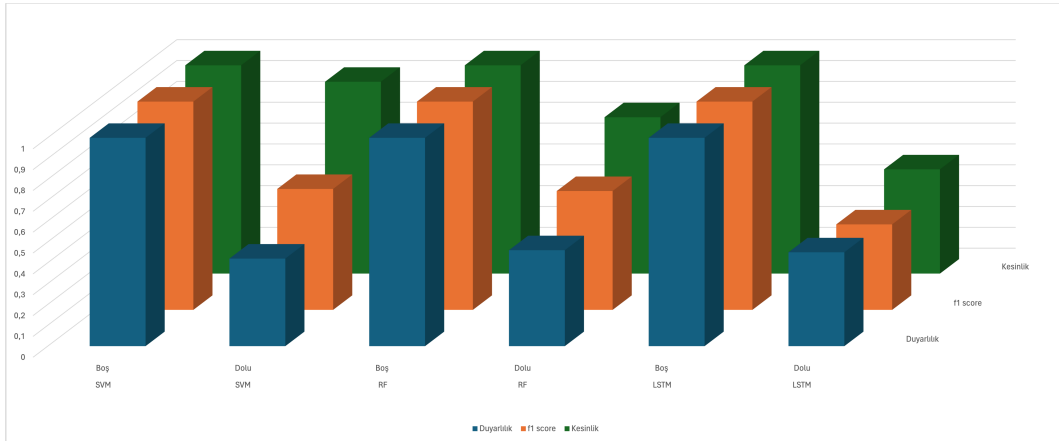


Şekil 4.1 800 MHz'de algoritmaların başarımları parametreleri özetini

Tablo 4.2 838 MHz'de algoritmaların doğrulama veri seti için doğruluk ve AUC değerleri

Algoritmalar	Doğruluk	AUC Değeri
SVM	%99	%71
RF	%99	%64
LSTM	%99	%74

Tablo 4.2'de ve Şekil 4.2 görüldüğü üzere LSTM algoritması diğer algoritmalara göre daha başarılıdır. Doğruluk değerlerinin yüksek çıkması veri setinin sınıflarının dengesiz dağılımından kaynaklanmaktadır.



Şekil 4.2 838 MHz'de algoritmaların başarımları parametreleri özetini

Başarımları oranı yüksek LSTM mimarisi için ise farklı veri setleri kullanarak ve veri miktarını değiştirerek başarımların artırılması planlanmaktadır. Ayrıca algoritmalar spektrumda bulunan tüm frekanslara uygulanarak algoritmaların performansını yükseltmek yapılabilir.

- [1] I. A. Akbar, W. H. Tranter, “Dynamic spectrum allocation in cognitive radio using hidden markov models: Poisson distributed case,” in *Proceedings 2007 IEEE SoutheastCon*, IEEE, 2007, pp. 196–201.
- [2] B. Bolat, M. O. Kelek, “Spectrum usage estimating and predicting,” in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2020, pp. 1–4.
- [3] T. Fujii, K. Umebayashi, “Smart spectrum for future wireless world,” *IEICE Transactions on Communications*, vol. 100, no. 9, pp. 1661–1673, 2017.
- [4] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, S. Mohanty, “A survey on spectrum management in cognitive radio networks,” *IEEE Communications magazine*, vol. 46, no. 4, pp. 40–48, 2008.
- [5] T. Black, B. Kerans, A. Kerans, “Implementation of hidden markov model spectrum prediction algorithm,” in *2012 international symposium on communications and information technologies (ISCIT)*, IEEE, 2012, pp. 280–283.
- [6] A. Ghosh, S. Kasera, J. Van Der Merwe, “Spectrum usage analysis and prediction using long short-term memory networks,” in *Proceedings of the 24th International Conference on Distributed Computing and Networking*, 2023, pp. 270–279.
- [7] A. Shenfield, Z. Khan, H. Ahmadi, “Deep learning meets cognitive radio: Predicting future steps,” in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, IEEE, 2020, pp. 1–5.
- [8] N. Radhakrishnan, S. Kandeepan, X. Yu, G. Baldini, “Performance analysis of long short-term memory-based markovian spectrum prediction,” *IEEE Access*, vol. 9, pp. 149 582–149 595, 2021.
- [9] H. Byun, S.-W. Lee, “A survey on pattern recognition applications of support vector machines,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 17, no. 03, pp. 459–486, 2003.
- [10] M.-W. Huang, C.-W. Chen, W.-C. Lin, S.-W. Ke, C.-F. Tsai, “Svm and svm ensembles in breast cancer prediction,” *PloS one*, vol. 12, no. 1, e0161501, 2017.
- [11] V. N. Vapnik, V. Vapnik, *et al.*, “Statistical learning theory,” 1998.
- [12] L. Breiman, *Random forests mach learn 45 (1): 5–32*, 2001.
- [13] J. D. Malley, J. Kruppa, A. Dasgupta, K. G. Malley, A. Ziegler, “Probability machines,” *Methods of information in medicine*, vol. 51, no. 01, pp. 74–81, 2012.

- [14] K. K. Nicodemus, J. D. Malley, “Predictor correlation impacts machine learning algorithms: Implications for genomic studies,” *Bioinformatics*, vol. 25, no. 15, pp. 1884–1890, 2009.
- [15] F. Degenhardt, S. Seifert, S. Szymczak, “Evaluation of variable selection methods for random forests and omics data sets,” *Briefings in bioinformatics*, vol. 20, no. 2, pp. 492–503, 2019.
- [16] F. A. Gers, J. Schmidhuber, F. Cummins, “Learning to forget: Continual prediction with lstm,” *Neural computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
- [17] Y. Yu, X. Si, C. Hu, J. Zhang, “A review of recurrent neural networks: Lstm cells and network architectures,” *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.
- [18] S. Hochreiter, J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [19] A. Sherstinsky, “Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network,” *Physica D: Nonlinear Phenomena*, vol. 404, p. 132 306, 2020.

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirisi

1. M.S.U. Türel, M.O. Kelek, "Estimation of Spectrum Usage with Machine Learning Algorithms", VII. International Halich Congress on Multidisciplinary Scientific Research, 2024, pp 860-870.

