

T.C.
İSTANBUL AYDIN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



**MAKİNE ÖĞRENME TEKNİKLERİNE DAYALI OLARAK
DENİZ YÜZEY SICAKLIĞI DEĞİŞİMİNİN MODELLENMESİ**

YÜKSEK LİSANS TEZİ

Cemal ERDEM

**Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Programı**

NİSAN, 2023

T.C.
İSTANBUL AYDIN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



**MAKİNE ÖĞRENME TEKNİKLERİNE DAYALI OLARAK
DENİZ YÜZEY SICAKLIĞI DEĞİŞİMİNİN MODELLENMESİ**

YÜKSEK LİSANS TEZİ

**Cemal ERDEM
(Y2013.010005)**

**Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Programı**

Tez Danışmanı: Prof. Dr. Zafer ASLAN

NİSAN, 2023

TEZ SINAV TUTANAĐI

ONUR SÖZÜ

Yüksek Lisans Tezi olarak sunduđum “MAKİNE ÖĐRENME TEKNİKLERİNE DAYALI OLARAK DENİZ YÜZEY SICAKLIKLARI DEĐİŐİM MODELLENMESİ” adlı alıŐmanın, tezin proje safhasından sonuçlanmasına kadarki bütün süreçlerde bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurulmaksızın yazıldığını ve yararlandığım eserlerin Kaynaka’da gösterilenlerden oluştuđunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve onurumla beyan ederim. (18/04/2023)

Cemal ERDEM

ÖNSÖZ

Bu tez çalışmasında, küresel ısınmanın deniz yüzey sıcaklıkları üzerindeki etkisini incelemek ve ömüzündeki yıllarda deniz yüzey sıcaklıklarının tahminlenmesi amaçlanmıştır.

Projemi hazırladığım tüm süreçte destekleriyle beni ayakta tutan ve süreçteki devamlılığımı kolaylaştıran başta ailem olmak üzere değerli arkadaşlarıma şükranlarımı sunuyorum.

Tez çalışmamı gerçekleştirdiğim süre içerisinde bilgi birikimini ve desteğini esirgemeyen değerli hocam Prof. Dr. Zafer Aslan'a sonsuz teşekkür, eder saygılarımı sunarım.

Nisan, 2023

Cemal ERDEM

MAKİNE ÖĞRENME TEKNİKLERİNE DAYALI OLARAK DENİZ YÜZEY SICAKLIĞI DEĞİŞİMİNİN MODELLENMESİ

ÖZET

Deniz yüzey sıcaklığı (DYS), okyanus ve atmosferik sistemlerin dinamiklerini anlamak ve gelecekteki iklim trendlerini tahmin etmek için kritik bir parametredir. Bu çalışmada, Avrupa Uzay Ajansı İklim Değişikliği Girişimi'nden, özellikle Deniz Yüzey Sıcaklığı İklim Değişikliği Girişimi (DYS IDG) projesinden elde edilen veriler Zonguldak ve Bartın illerindeki DYS'nın ARIMA (otoregresif hareketli ortalamalar) ve LSTM yöntemlerini kullanarak modellenmektedir. Veri seti (boylam 31.25°E ve enlem 40.95°N) 1981'den 2022'ye kadar olan 40 yıllık günlük verilere dayalı olarak hesaplanan aylık değerleri kapsamakta ve DYS trendleri ile mevsimsel varyasyonların bir değerlendirmesini içermektedir.

Bu çalışmada, deniz yüzey sıcaklıklarının ARIMA ve LSTM modelleri kullanılarak tahmini yapılmıştır. Her iki model de önümüzdeki 5-10 yıl içinde deniz yüzey sıcaklıklarının artış eğiliminde olacağını göstermektedir. ARIMA modeli ile yapılan çalışmada gözlem ve tahmin değerleri arasında ortalama kare hatası (MSE) değeri 0,07 olarak hesaplanmıştır. LSTM modeli ise daha düşük bir hata payı ile (0.005 MSE) tahmin yapılabilmektedir. Bu sonuçlar, LSTM modelinin ARIMA modeline göre daha yüksek doğrulukta sonuçlar verdiğini göstermektedir.

Zonguldak ve Bartın illerindeki değişen DYS trendleri, balıkçılık ve turizm de dahil olmak üzere birçok endüstri vb. sektörler için önemli sonuçları oluşturmaktadır. Bu çalışmanın sonuçlarının bu alanlardaki karar verme süreçlerine ve iklim değişikliği uyum ve stratejileriyle ilgili politika kararlarına yardımcı olması beklenmektedir. Bulgularımız ayrıca, DYS verilerinin alındığı DYS IDG projesi ve ARIMA ve LSTM yöntemlerinin DYS verilerini modelleme etkinliği ve potansiyel sınırlamaları hakkında da bilgi sağlamaktadır.

Anahtar Kelimeler: Deniz Yüzey Sıcaklığı, Makine Öğrenmesi, ARIMA, LSTM, Performans

SEA SURFACE TEMPERATURE CHANGES BASED ON MACHINE LEARNING TECHNIQUES

ABSTRACT

Sea Surface Temperature (SST) is a critical parameter in understanding the dynamics of oceanic and atmospheric systems and predicting future climate trends. In this study, we use data obtained from the European Space Agency Climate Change Initiative, specifically from the Sea Surface Temperature Climate Change Initiative (SST CCI) project, to model SST in the Zonguldak and Bartın provinces using the Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) methods. The dataset covers 40 years from 1981 to 2022 (longitude 31.25°E and latitude 40.95°N) and includes monthly values calculated based on 40 years of daily data from an assessment of SST trends and seasonal variations.

In this study, sea surface temperatures were estimated by using ARIMA and LSTM models. Both models show that sea surface temperatures will be on an increasing trend in the next 5-10 years. In the study with the ARIMA model, the mean square error (MSE) value was calculated as 0.07. The LSTM model, on the other hand, has a lower margin of error of 0.005 MSE. These results show that the LSTM model gives more accurate results than the ARIMA model.

The changing SST trends in the Zonguldak and Bartın provinces have important implications for several industries and sectors, including fisheries and tourism. The results of this thesis are expected to assist decision-making processes in these areas and policy decisions regarding with climate change adaptation and strategies. Our findings also provide information about the SST CCI project from which SST data was obtained, and the potential limitations and effectiveness of ARIMA and LSTM methods for modeling SST data.

Keywords: Sea Surface Temperature, Machine Learning, ARIMA, LSTM, Performance

İÇİNDEKİLER

ONUR SÖZÜ	iv
ÖNSÖZ.....	iv
ÖZET.....	v
ABSTRACT	vi
İÇİNDEKİLER	vii
ŞEKİLLER LİSTESİ.....	ix
ÇİZELGELER LİSTESİ.....	x
KISALTMALAR LİSTESİ.....	xi
I. GİRİŞ.....	1
II. LİTERATÜR TARAMASI.....	3
III. YÖNTEM	6
A. Arıma	6
1. AIC (The Akaike Information Criteria – Akaike Ölçütü)	7
3. Kullanım Alanları	7
B. LSTM (Long Short Term Memory / Uzun Kısa Süreli Bellek)	8
1. LSTM Kullanım Alanları	9
2. Parametreler	10
3. LSTM Aktivasyon Fonksiyonları	11
IV. VERİ SETİ VE YÖNTEM.....	16
A. Avhrr / 3 MetOp uydusu.....	20
V. UYGULAMA.....	21
A. Araştırmanın Amacı.....	22
B. Bulgular	22

1. ARIMA Analiz Sonuçları	22
2. LSTM Analiz Sonuçları	28
3. Eğitim Performans Optimizasyonu ve Ölçümleri.....	30
VI. SONUÇ VE ÖNERİLER	32
VII. KAYNAKLAR.....	33
EKLER.....	36
ÖZGEÇMİŞ.....	39

ŞEKİLLER LİSTESİ

Şekil 1 Hindistan ve Hint Okyanusu çevresinde çalışma alanları.....	5
Şekil 2 Klasik RNN Yinelenen Modül.....	9
Şekil 3 Sigmoid Fonksiyonu ve Türevi Grafiği (Kızrak, 2019).....	12
Şekil 4 Hiperbolik Tanjant Fonksiyon ve Türevi (Kızrak, 2019).....	13
Şekil 5 ReLU Aktivasyon Fonksiyonu ve Türevi (Kızrak, 2019).....	14
Şekil 6 Tam Çözünürlüklü (L2P) Deniz Yüzey Sıcaklıklarının Özellikleri.....	18
Şekil 7 İklim Veri Kayıtlarının Üretim Aşaması.....	19
Şekil 8 LSTM Model Akış Şeması.....	22
Şekil 9 Yıllara Göre Zonguldak-Bartın Sıcaklık.....	23
Şekil 10 ARIMA Modeli Eğitim- Test / Zonguldak-Bartın.....	25
Şekil 11 ARIMA Modeli Eğitim / Test Dinamik Veriler Zonguldak-Bartın.....	26
Şekil 12 ARIMA Model ile Gelecek 5 Yıl Tahmini-Zonguldak-Bartın.....	27
Şekil 13 ARIMA Model ile Gelecek 10 Yıl Tahmini-Zonguldak-Bartın.....	27
Şekil 14 LSTM-Zonguldak/Bartın Loss/Epoch.....	29
Şekil 15 LSTM-Zonguldak/Bartın 5 Yıllık Tahmin Grafik.....	29

ÇİZELGELER LİSTESİ

Çizelge 1 Veri İşleme Sonrası Format (Örnek).....	16
Çizelge 2 Veri Seti Anahtar Kelimeler	17
Çizelge 3 AIC Değerleri Zonguldak-Bartın.....	24
Çizelge 4 Model: "RNN-LSTM model parametreleri. Model: "model""	28

KISALTMALAR LİSTESİ

ADAM	: Adaptive Moment Estimation (Uyarlanabilir Moment Tahmini)
AIC	: Akaike Information Criteria (Akaike Bilgi Kriterleri)
ARIMA	: Auto-Regressive Integrated Moving Average (Oto-regresif Entegre Hareketli Ortalama)
ATSR	: Along-Track Scanning Radiometers (İz Boyunca Taramalı Radyometreler)
AVHRR	: Advanced Very High Resolution Radiometers (Gelişmiş Çok Yüksek Çözünürlüklü Radyometreler)
DYS	: Deniz Yüzey Sıcaklığı
ESA	: European Space Agency
IDG	: Climate Change Initiative (İklim Değişikliği Girişimi)
IFOV	: Instantaneous Field of View (Anlık Görüş Alanı)
K	: Kelvin
LSTM	: Long-Short Tem Memory (Uzun-Kısa Süreli Hafıza Ağı)
MSE	: Mean Square Error (Ortalama Kare Hatası)
NLP	: Natural Language Processing (Doğal Dil İşleme)
OISST	: Optimum Interpolation Sea Surface Temperature (Optimum Deniz Yüzey Sıcaklığı Interpolasyonu)
RELU	: Rectifier Linear Unit (Doğrultulu Lineer Birim)
RMSE	: Root Mean Square Error (Kök Ortalama Kare Hatası)
RNN	: Recurrent Neural Network (Tekrarlayan Sinir Ağı)
SST	: Sea Surface Temperature
YSA	: Yapay Sinir Ağı

I. GİRİŞ

Deniz Yüzey Sıcaklığı (DYS), okyanus ve atmosfer sistemlerinin dinamiklerini anlamak için temel bir parametredir ve gelecekteki iklim eğilimlerini tahmin etmede kritik bir rol oynamaktadır (Manabe, 1988). DYS, okyanusun üst tabakasının sıcaklığının bir ölçüsüdür ve güneş radyasyonu, rüzgâr ve akımlar dahil olmak üzere çeşitli faktörlerden etkilenir. DYS 'deki değişiklikler, okyanus dolaşımı, deniz ekosistemleri ve hava koşulları dahil olmak üzere çeşitli doğal süreçler üzerinde geniş kapsamlı etkilere sahip olabilir.

Deniz yüzeyi sıcaklık anomalileri, hava durumu modelleri ve şiddetli hava olayları üzerinde önemli etkilere sahip olabilir. Araştırmalar, SST anomalilerinin öngörülebilirliğinin, geleneksel oşinografi modelleme ve makine öğrenimi tekniklerinin bir kombinasyonu kullanılarak iyileştirilebileceğini göstermiştir. Ek olarak, DYS anomalilerinin veya okyanus özelliklerinin yıldan yıla kalıcılığının, küresel yüzey ısınmasına yanıt olarak üst okyanus karışık katman derinliğinin sığlaşması nedeniyle önümüzdeki 20-30 yılda küresel olarak azalması beklenmektedir. (Hui Shi, 2022)

Ayrıca DYS, balıkçılık, turizm ve tarım gibi çeşitli insan faaliyetleri ve endüstrileriyle yakından bağlantılıdır. DYS 'deki değişiklikler, balık popülasyonlarının dağılımını ve bolluğunu etkileyerek gelirleri ve gıda güvenliği için onlara bağımlı olanların geçim kaynaklarını etkileyebilecektir (Mohamed, Ibrahim & Nagy, 2022). Kıyı bölgelerinde bir başka önemli sektör olan turizm, plajların ve diğer kıyı cazibe merkezlerinin kalitesini etkileyebileceği için DYS'deki değişiklikler ve gelecek dönemde tahmini önem taşımaktadır. Tarım, DYS 'deki değişikliklerden de etkilenebilir, çünkü deniz ve okyanuslar önemli bir nem kaynağı olup öngörülebilir yağış modellerine dayalı mahsullerin büyümesini ve verimini etkileyebilir.

DYS'nın hem doğal süreçler hem de insan faaliyetleri için önemi göz önüne alındığında, DYS eğilimlerinin doğru modellenmesi ve tahmini bilim adamları, kanun koyucular ve paydaşlar için çok önemlidir. Bu çalışma, Türkiye'nin Karadeniz bölgesinde bulunan ve son birkaç on yılda DYS'de değişiklikler gözlenen Zonguldak

ve Bartın civarı inceleme bölgeleri olarak seçilmiştir. Avrupa Uzay Ajansı İklim Değişikliği Girişimi'nden (European Space Agency Climate Change Initiative) elde edilen verileri analiz ederek, bu bölgedeki DYS eğilimlerinin değişmesinin anlaşılmasına katkıda bulunulması amaçlanmıştır. Ayrıca stratejileri ile ilgili politika kararları için girdi verisi oluşturulması hedeflenmiştir. (Hoegh-Guldberg, 2014)

II. LİTERATÜR TARAMASI

Eldeki verilerden yararlanılarak, literatürde deniz yüzey sıcaklığı tahmini için farklı yöntemler ve modeller önerilmiştir. Bu yöntemler arasında istatistiksel modeller, yapay sinir ağı modelleri ve makine öğrenme yöntemleri yer almaktadır. Örneğin, ARIMA modeli deniz yüzey sıcaklığı tahmininde sıkça kullanılan bir istatistiksel yöntemdir. Benzer şekilde, LSTM modeli gibi yapay sinir ağı modelleri de deniz yüzey sıcaklığı tahmininde etkili olabilmektedir. Ayrıca, son araştırmalarda makine öğrenimi yöntemleri de deniz yüzey sıcaklığı tahmini için önerilmekte ve başarılı sonuçlar elde edilmektedir. Ancak, her bir yöntemin performansı ve uygulama alanı, veri özelliklerine, model parametrelerine ve kullanım amacına bağlı olarak değişebilmektedir. Bu nedenle, bu literatür taraması bölümünde, farklı yöntemler ve modellerin literatürdeki performansı ve uygulama alanları incelenerek, tezin amaçları doğrultusunda uygun yöntemlerin seçilmesine çalışılmıştır.

(Mohamed, Ibrahim & Nagy, 2022), yüksek çözünürlüklü deniz yüzey sıcaklığı uydu verileri kullanılarak 1982'den 2020'ye kadar Karadeniz'deki sıcaklık trendleri, değişkenlik ve deniz sıcaklık dalgalarının mekânsal ve zamansal değişkenliğini, trendlerini ve örüntülerini incelemiştir. El Nino-Güney Salınımı (ENSO) ile sıcaklık dalgalarının sıklığı arasındaki ilişki araştırılmıştır. Sonuçlar, Karadeniz'de 1982-2020 döneminde yükselen bir sıcaklık artış hızı varlığını ve özellikle son 20 yılda (2001-2020) yüksek sıcak dalga sıklığı eğilimi olduğunu göstermektedir. Karadeniz'in batısından doğusuna doğru lineer bir sıcaklık artış eğilimi de gözlenmiştir. Yıllık ortalama deniz yüzeyi sıcaklığı ile yıllık sıcak dalga sıklığı arasında yüksek bir korelasyon ($R = 0.90$) gözlemlenmiştir.

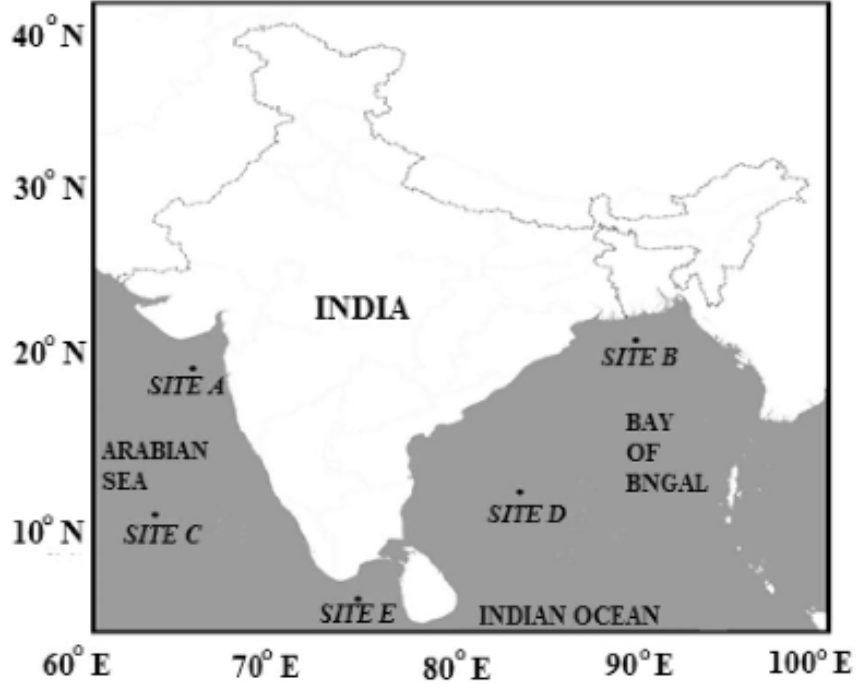
EL Nino (sıcak su) ve La Nina (soğuk su) Pasifik Okyanusu'ndaki akıntı özellikleri olup, hava koşullarını küresel ölçekte etkilemektedir. El Nino ve La Nina rüzgarları 2-7 yılda bir gözlenir. El Nino olayında ticaret rüzgarları zayıflar, Güney Amerika'dan Asya'ya doğru olan sıcak deniz suyu Ekvator boyunca batıya doğru hareket eder (Noaa, 2017). El Nino ve La Nina'nın hava durumu, orman yangınları, ekosistemler ve ekonomiler üzerinde önemli küresel etkileri olmaktadır.

El Nino ve sıcak dalga olayı örtüşmektedir. Bu çalışma, gelecekte deniz sıcaklık dalgalarının sıklığının küresel ısınmaya bağlı olarak artabileceğini ve bu nedenle Karadeniz'deki ana sürücülerinin incelenmesi gerektiğini belirtmektedir.

(Jia X, 2022) Deniz yüzey sıcaklığı tahmininde uzun kısa vadeli hafızalı ağlar (LSTM) kullanarak, bir haftalık ve bir aylık tahmin yapmak için ilk kez tekrarlayan sinir ağı yöntemini kullandıklarını belirtmektedirler. Makalede, deniz yüzey sıcaklığı tahmini problemini bir zaman serisi regresyon problemi olarak formüle ederek, LSTM ağlarını kullanarak bir model önerilmektedir. Önerilen ağ mimarisi iki tür katmandan oluşmaktadır: LSTM katmanı ve tam bağlantılı yoğun katman. LSTM katmanı, zaman serisi ilişkileri modellemek için kullanılırken, tam bağlantılı katman, LSTM katmanının çıktısını nihai tahmin ve haritalamak amacıyla kullanılır. Yapılan deneylerle optimal mimari ayarları incelenerek, Çin kıyı denizleri için tahmin doğruluğu rapor edilmiştir. Ayrıca, önerilen yöntemin online güncellenebilir özelliklerini de göstermektedirler.

(Aydınlı, 2022) 2022 yılında Karadeniz'de yapılan bir çalışmada, LSTM yöntemi kullanılarak gerçekleştirilen bir çalışmanın sonuçları incelenmiştir. Bu çalışmada, deniz ortamı ve deniz türleri hakkında kapsamlı bilgi edinmek için yüksek zaman çözünürlüklü uzaktan algılama görüntüleri kullanılmıştır. Çalışmanın amacı, Karadeniz'in orta bölgesinde günlük deniz yüzey sıcaklığı (DYS) değerini tahmin etmek için bir tahmin modeli oluşturmaktır. Bu amaçla, 2008-2021 yılları arasındaki zaman serisi uydu verileri kullanılmıştır. Çalışmanın sonuçlarına göre, LSTM yöntemi kullanılarak elde edilen tahmin modeli, DHS tahmininde önemli bir potansiyele sahiptir. Doğrulama için kullanılan Adam stokastik optimizasyonu ile her bir konum için ortalama kare hata (MSE) değerleri hesaplanmıştır. Sonuçlar, Ordu, Samsun ve Sinop şehirleri için sırasıyla 0.914, 0.815 ve 0.802 MSE değerlerini göstermektedir. (Zhang Q. W., 2017) 2017 yılında Çin Kıyı denizlerinde yapılan çalışma da deniz yüzey sıcaklığını (DYS) tahmin etmek için uzun kısa vadeli hafıza (LSTM) yöntemi kullanılmıştır. Analiz için kullanılan veri seti, 1981/09'dan 2016/11'e kadar olan dönem kapsayan NOAA yüksek çözünürlüklü DHS verileridir. Özellikle, Bohai Denizi'nde yapılan doğrulama sonuçlarına odaklanılmıştır. Elde edilen sonuçlara göre, LSTM modeli Bohai Denizi'nde DHS tahmininde etkili bir performans sergilemiştir. RMSE (Ortalama Kare Hata) değeri 1.1363 olarak hesaplanmıştır.

(Sarkar, Janardhan, & Roy, 2020) yapılan çalışma da LSTM derin öğrenme sinir ağıları kullanılarak deniz yüzey sıcaklıklarının tahmini yapılmıştır. Çalışma Hindistan ve Hint Okyanusu çevresinde beş farklı konumu içermektedir. (SITE A: 19°N-68°E, SITE B: 20°N-90°E, SITE C: 12°N-64°E, SITE D: 13°N-84°E, SITE E: 8°N-75°E)



Şekil 1 Hindistan ve Hint Okyanusu çevresinde çalışma alanları

Günlük tahminler için veri seti 36 ay (2015-2018), haftalık tahminler için 38 yıl (1981-2018) ve aylık analizler için 149 yıl (1870-2018) sürecini kapsamaktadır. Veri setleri NOAA tarafından sağlanmış olup özellikle OISST sürüm 2 veri seti kullanılmıştır. Tahminlerde derin öğrenme sinir ağı yöntemi olan LSTM kullanılmıştır. Yapılan tahminlerde RMSE (Root Mean Square Error) değerleri 0.37'den küçük bulunmuştur.

Sonuç olarak, literatürde ARIMA ve LSTM modelleri gibi istatistiksel ve yapay sinir ağı modellerinin deniz yüzey sıcaklığı tahmini için yaygın olarak kullanıldığı görülmüştür. Ancak, her bir modelin performansı ve uygulama alanı, veri özelliklerine, model parametrelerine ve kullanım amacına bağlı olarak değişebilir.

III. YÖNTEM

A. Arima

Otoregresif Hareketli Ortalama anlamına gelen ARIMA, geçmiş gözlemlere dayalı olarak gelecekteki değerleri tahmin etmek için kullanılan popüler bir zaman serisi analiz tekniğidir. Ekonomik, finansal ve meteorolojik veriler dahil olmak üzere çok çeşitli zaman serisi verilerini analiz etmek ve modellemek için esnek ve güçlü bir araçtır (Zhang G. P., 2003).

ARIMA modeli üç ana bileşenden oluşur: otoregresif (AR) bileşeni, fark (I) bileşeni ve hareketli ortalama (MA) bileşeni. Otoregresif bileşen, gelecekteki değerleri tahmin etmek için geçmiş gözlemleri kullanırken, hareketli ortalama bileşeni, tahmini iyileştirmek için geçmiş tahminlerden gelen hataları kullanır. Fark bileşeni, verilerdeki herhangi bir trend veya mevsimsel etkiyi ortadan kaldırmak için kullanılır.

ARIMA, bir zaman serisi tahmin yöntemi olarak birçok avantaja sahiptir. İlk olarak, doğrusal olmayan ve durağan olmayan veriler de dahil olmak üzere geniş bir veri deseni yelpazesini ele alabilen esnek bir modele sahiptir. İkinci olarak, verilerden model parametrelerini tahmin edebilen bir parametrik modele sahiptir ve bu sayede veri kullanımını açısından daha verimlidir. Üçüncü olarak, karar verme süreci için önemli olan tahmin belirsizliğinin istatistiksel ölçüsünü sağlar. Son olarak, tahmin modeli yeni verilerle kolayca güncellenebilir ve böylece sürekli izleme ve uyarılama sağlanabilir.

ARIMA'daki her bileşen, standart notasyonlu bir parametre olarak işlev görür. ARIMA modelleri için standart notasyon, p , d ve q ile ARIMA olacaktır; burada kullanılan ARIMA modelinin türünü belirtmek için parametrelerin yerine tamsayı değerler geçmektedir. Parametreler şu şekilde tanımlanabilir:

- p , modelin otomatik gerileyen kısmıdır. Geçmiş değerlerin etkisini modelimize dahil etmemizi sağlar. Sezgisel olarak bu, son 3 gündür hava sıcaksa yarın da sıcak olacağını söylemeye benzer.

- d , modelin bütünleşmiş parçasıdır. Bu, zaman serisine uygulanacak fark miktarını (yani mevcut değerden çıkarılacak geçmiş zaman noktalarının sayısı) içeren modeldeki terimleri içerir. Sezgisel olarak bu, son üç gündeki sıcaklık farkı çok küçükse, yarın da aynı sıcaklık olacağını söylemek mümkündür.
- q , modelin hareketli ortalama kısmıdır. Bu, modelimizin hatasını geçmişte önceki zaman noktalarında gözlemlenen hata değerlerinin doğrusal bir kombinasyonu olarak belirlenmesine olanak tanır.

1. AIC (The Akaike Information Criteria – Akaike Ölçütü)

ARIMA modellerinde AIC (Akaike Information Criterion) Akaike ölçütü bir modelin kalitesinin değerlendirilmesinde kullanılan bir ölçüttür. AIC, modelin performansını hem uyum hem de aşırı uyum açısından değerlendirir. AIC, düşük hatayla uyumlu bir modeli tercih ederken aşırı uyumdan kaçınır. Bu nedenle, AIC değeri düşük olan bir model, daha iyi bir model olarak kabul edilir.

AIC değeri, modelin karmaşıklığı ile log-likelihood değerinin bir fonksiyonudur. AIC değerinin hesaplanması için, öncelikle modelin log-likelihood değeri hesaplanır. Log-likelihood, modelin uyum kalitesinin bir ölçüsüdür ve logaritmik bir olasılık değeridir. Modelin karmaşıklığı arttıkça, log-likelihood değeri azalır. AIC, log-likelihood değerine, modelin karmaşıklığını (örneğin, p , d ve q değerleri gibi) azaltmak için eklenen bir terimle hesaplanır. Bu nedenle, daha karmaşık bir modelin AIC değeri daha yüksek olacaktır.

AIC, özellikle regresyon analizi gibi büyük veri setleri üzerinde çalışan araştırmacılar tarafından sıklıkla kullanılır. Ancak, yalnızca iki model arasındaki karşılaştırmalar için kullanılması önerilir (Cavanaugh, 2019).

3. Kullanım Alanları

ARIMA modelleri, zaman serileri analizinde kullanılan ve istatistiksel yöntemler arasında en yaygın olanıdır. Bu model, geçmiş verilere dayanarak gelecekteki değerleri tahmin etmek için kullanılır.

- Ekonomi: ARIMA modelleri, ekonomik verilerin analizinde yaygın olarak kullanılır. Örneğin, hisse senedi fiyatları, faiz oranları, döviz kurları, enflasyon oranları gibi finansal verilerin tahmini için ARIMA modelleri kullanılır.

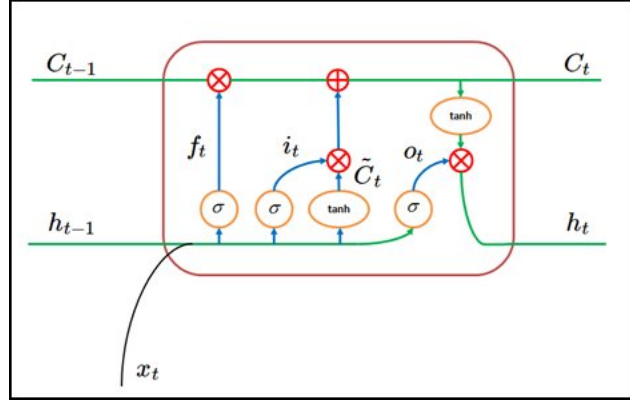
- Meteoroloji: ARIMA modelleri, meteorolojik verilerin analizinde de kullanılır. Özellikle, hava sıcaklığı, yağış miktarı ve rüzgâr şiddeti gibi parametrelerin tahmini için sıklıkla kullanılır. Bu tür tahminler, tarım, su kaynakları yönetimi ve enerji sektörü gibi birçok sektörde kullanılır.
- Enerji: ARIMA modelleri, enerji talebi tahmini ve enerji fiyatları gibi konuların analizinde de kullanılır. Özellikle, elektrik tüketimi tahmininde çok etkilidir.
- Endüstriyel Üretim: ARIMA modelleri, endüstriyel üretim tahmininde de kullanılır. Özellikle, malzeme stokları, üretim hacmi ve sipariş miktarı gibi verilerin analizinde kullanılır.
- Sağlık: ARIMA modelleri, sağlık verilerinin analizinde de kullanılır. Örneğin, hastane yatışları, bulaşıcı hastalık oranları, ölüm oranları gibi parametrelerin tahmininde kullanılır.
- İletişim: ARIMA modelleri, telefon trafiği, internet trafiği, veri transfer hızı gibi konuların analizinde de kullanılır. Bu tür tahminler, iletişim sektöründe kapasite planlaması ve ağ optimizasyonu gibi konuların yönetiminde etkili bir araçtır.

ARIMA modelleri, veri setlerinin zaman serilerine ayrılması, trend, döngü ve rasgele etkilerin belirlenmesi ve gelecekteki verilerin tahmin edilmesi için kullanılan bir dizi matematiksel teknik ve yöntemlerdir. Bu modeller, verilerin düzenlenmesi ve analiz edilmesi için çok etkili bir araçtır ve birçok sektörde kullanılır.

B. LSTM (Long Short Term Memory / Uzun Kısa Süreli Bellek)

LSTM (uzun kısa-süreli bellek) ağları, zaman serileri veya metin gibi yapısal verilerin anlaşılması ve gelecekteki tahminlerin yapılması için kullanılan "tekrarlayan sinir ağı" (RNN) mimarisinin bir türüdür (Han, 2023). Hochreiter ve Schmidhuber (1997) tarafından tanıtılmış ve daha sonra birçok araştırmacı tarafından geliştirilmiştir. LSTM, özel bir RNN türüdür ve uzun süreli bağımlılıkları öğrenebilme özelliğiyle geleneksel RNN modellerinin zayıf yönlerini ele alarak, zaman serileri tahmini, dil modelleri ve doğal dil işleme gibi alanlarda başarılı sonuçlar verir. (Hochreiter, 1997).

Zaman serileri tahmini, ekonomik trendlerdeki öngörülemez değişiklikler ve eksik bilgi nedeniyle zor bir süreçtir ve bazı yöntemlerin uygulamalarda kısıtlamaları olduğu için doğruluğun değerlendirilmesi gereklidir. LSTM'ler, uzun vadeli bağımlılık sorununu önlemek için açıkça tasarlanmıştır. Tüm tekrarlayan sinir ağları (RNN), birbirini tekrarlayan sinir ağı modüllerinin bir zincir şeklinde oluşturduğu bir yapıya sahiptir. Standart RNN'lerde bu tekrarlayan modül, tek bir tanh katmanı gibi basit bir yapıya sahiptir.



Şekil 2 Klasik RNN Yinelenen Modül

LSTM'ler de zincir şeklinde bir yapıya sahip olmakla birlikte, tekrarlayan modüller farklı bir yapıya sahiptir. Bu ağlar, birbirine bağlı hücrelerden (LSTM blokları) oluşur. Her LSTM bloğu, sırasıyla giriş kapısı (input gate), çıkış kapısı (output gate) ve unutma kapısı (forget gate) adında üç tip kapı içerir. Bu kapılar, hücre belleğindeki bilgilerin yazılması, okunması ve unutulması işlemlerini gerçekleştirir. Her kapı, sigmoid aktivasyon fonksiyonunu kullanır ve bu sayede hücrenin belleğindeki bilgiyi belirli bir süre boyunca tutar. Bu özellik, LSTM'lerin uzun süreli bağımlılıkları öğrenebilmesini sağlar ve zaman serileri tahmini, dil modelleri ve doğal dil işleme gibi alanlarda başarılı sonuçlar verir. (Siame-Namini, 2018)

1. LSTM Kullanım Alanları

LSTM modelleri, yüksek boyutlu ve zamansal serilerin analizi için son derece etkili bir araçtır. Geleneksel sinir ağı modellerine kıyasla daha yüksek performans gösterir ve özellikle uzun vadeli bağımlılıkları modelleme konusunda oldukça etkilidirler. Bu nedenle, birçok sektörde geniş bir uygulama yelpazesine sahip olmalarıyla birlikte, büyük başarılar elde etmişlerdir. LSTM ağları, konuşma işleme,

dil çevirisi ve duyarlılık analizi gibi çeşitli uygulamalarda başarıyla kullanılmıştır, (Tai, 2015) (Li, 2021) (Pal, 2018). Başlıca kullanım alanları şu şekildedir:

- Doğal Dil İşleme (NLP): LSTM modelleri, dil modellerinin oluşturulması, kelime gömme (word embedding) ve duygusal analiz gibi NLP uygulamalarında yaygın olarak kullanılır.
- Ses Tanıma: LSTM modelleri, ses tanıma sistemlerinde sıklıkla kullanılır. Özellikle, konuşma tanıma, konuşma sentezi ve arka plan gürültüsü azaltma gibi uygulamalar için kullanılır.
- Görüntü İşleme: LSTM modelleri, görüntü sınıflandırma, nesne tanıma, yüz tanıma gibi uygulamalarda kullanılır. Özellikle, resim dizilerinin analizi için kullanılırlar.
- Robotik: LSTM modelleri, robotik sistemlerde kullanılır. Özellikle, hareket kontrolü, hareket planlaması ve nesne algılama gibi uygulamalar için kullanılır.
- Tıp: LSTM modelleri, tıp uygulamalarında kullanılır. Özellikle, hastalık teşhisi, ilaç keşfi ve hastalık gelişme tahmini gibi uygulamalarda kullanılır.
- Finansal Tahmin: LSTM modelleri, finansal piyasa tahminlerinde kullanılır. Özellikle, hisse senedi fiyatı tahmini, döviz kuru tahmini ve trend tahmini gibi uygulamalarda kullanılır.

2. Parametreler

LSTM ağları, performanslarını artırmak için optimize edilebilecek birkaç parametreye sahiptir (Greff, 2017) (Alamri, 2023). LSTM ağlarının en önemli parametreleri, bellek hücrelerinin sayısı, gizli birimlerin sayısı, öğrenme hızı, toplu işlem boyutu ve epoch sayısıdır (Greff, 2017) (Alamri, 2023). Bellek hücrelerinin sayısı, ağın bellek kapasitesini belirlerken, gizli birimlerin sayısı modelin karmaşıklığını kontrol eder (Greff, 2017). Öğrenme hızı, optimizasyon algoritmasının adım boyutunu kontrol ederken, toplu işlem boyutu optimizasyon algoritmasının her iterasyonunda kullanılan örnek sayısını belirler (Alamri, 2023). Epoch sayısı, eğitim sırasında tüm eğitim veri kümesinin ağdan geçirilme sayısını belirler (Greff, 2017). Bu parametrelerin yanı sıra, LSTM ağları da bilgi akışını kontrol eden birkaç kapıya sahiptir, bunlar arasında giriş kapısı, unutma kapısı ve çıkış kapısı bulunur (Okut,

2021) (Khan, 2023) (Y. Fang, 2019). Bu kapılar, bilgi akışını kontrol etmek ve ağın performansını artırmak için ayarlanabilir.

3. LSTM Aktivasyon Fonksiyonları

LSTM ağları, kapılarında kullanılan dört farklı aktivasyon fonksiyonuna sahiptir. Bu fonksiyonlar, sigmoid, hiperbolik tanjant (tanh), doğrultulmuş lineer birim (ReLU) ve lineer aktivasyon fonksiyonudur. Her bir aktivasyon fonksiyonunun farklı bir amaca hizmet ettiği bilinmektedir. Sigmoid fonksiyonu, bellek hücresine bilginin akışını kontrol etmek için kullanılırken, hiperbolik tanjant fonksiyonu bellek hücresi durumunu güncellemek için kullanılır. ReLU fonksiyonu, doğrusal olmayan özelliklerin modelleme sürecini artırmak için kullanılırken, lineer aktivasyon fonksiyonu doğrusal işlemler için kullanılır. LSTM ağının her bir kapısında doğru aktivasyon fonksiyonlarının kullanımı, ağı daha etkili ve doğru sonuçlar üretmesini sağlar.

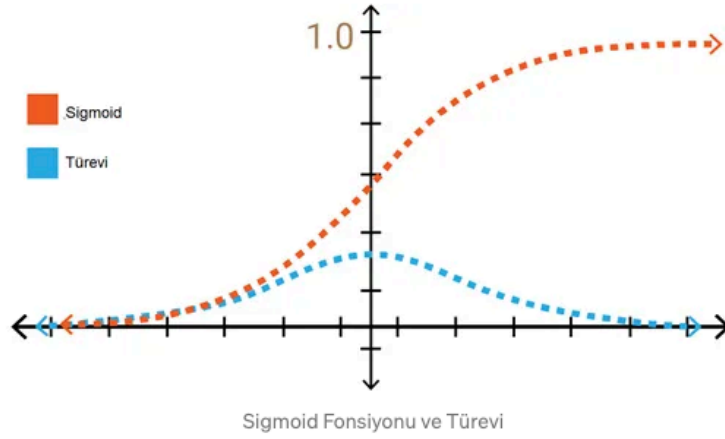
a. Sigmoid aktivasyon fonksiyonu

LSTM ağlarında, bellek hücresine akışı kontrol etmek için Sigmoid fonksiyonu kullanılır. Bu fonksiyon, herhangi bir giriş değerini 0 ile 1 arasında bir değere eşler ve bu değer, girişin ilgili olma olasılığı olarak yorumlanabilir. LSTM kapılarının Giriş (i) ve Unutma (f) kapıları, sigmoid aktivasyon fonksiyonunu kullanır. Bu sayede, ağı öğrenme sürecinde bellek hücresindeki bilginin akışı kontrol edilebilir ve ağ, girdiler arasında uygun bir ilişki kurabilir. (Kizrak, 2019)

$$\sigma(x) = \frac{1}{e^{-x} + 1}$$

"e", doğal logaritma tabanını temsil eder ve fonksiyonun denkleminde sabit bir değerdir.

"x", sigmoid fonksiyonunun girdisidir ve herhangi bir sayı olabilir.



Şekil 3 Sigmoid Fonsiyonu ve Türevi Grafiği (Kızrak, 2019)

b. Hiperbolik tanjant aktivasyon fonksiyonu

Hiperbolik tanjant (tanh) fonksiyonu, yapay sinir ağlarında özellikle bellek hücresi durumunu güncellemek için sıklıkla kullanılır. Bu fonksiyon, herhangi bir giriş değerini -1 ile 1 arasında bir değere eşitler. Bu özelliği sayesinde, ağın çıktılarının sınırlanması ve aşırı uyumaya (overfitting) karşı korunması sağlanır.

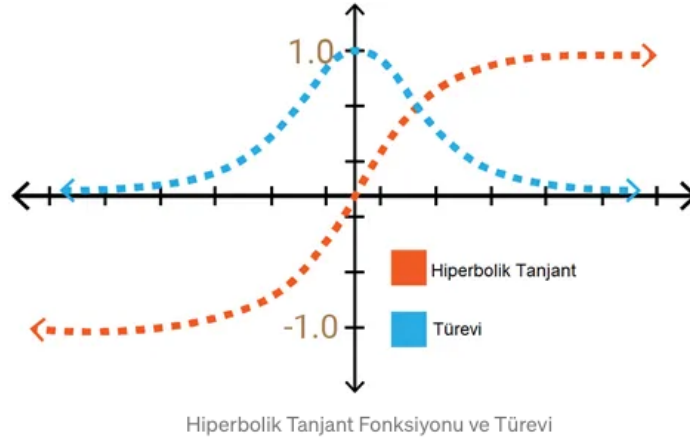
LSTM ağlarında çıkış kapısı (output gate), tanh aktivasyon fonksiyonunu kullanarak çıktı üretir. Bu sayede, ağın daha doğru sonuçlar vermesi ve daha iyi performans göstermesi sağlanır. Tanh fonksiyonu, sigmoid fonksiyonuna benzer ancak sigmoid fonksiyonundan farklı olarak simetrik bir yapıya sahiptir ve bu özelliği sayesinde daha iyi sonuçlar elde edilir.

Hiperbolik tanjant fonksiyonu, birçok yapay sinir ağı algoritmasında kullanılan etkili bir aktivasyon fonksiyonudur. Ayrıca, ağın geriye doğru yayılımı (backpropagation) sırasında da türevlenebilir olması, ağın eğitim sürecinde başarılı olmasını sağlar. (Kızrak, 2019)

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Burada "e" doğal logaritma tabanını, "x" ise fonksiyonun girdisini temsil eder. Fonksiyon, girdinin pozitif veya negatif olmasına bağlı olarak -1 ve 1 arasında bir çıkış değeri üretir.

Hiperbolik tanjant fonksiyonu, sigmoid fonksiyonuna benzer bir şekilde, sıklıkla yapay sinir ağlarında kullanılır ve girdinin ağırlıklı toplamını bir sınırlandırma işlemiyle normalize eder.



Şekil 4 Hiperbolik Tanjant Fonksiyon ve Türevi (Kizrak, 2019)

c. Düzeltilmiş doğrusal birim aktivasyon fonksiyonu (ReLU)

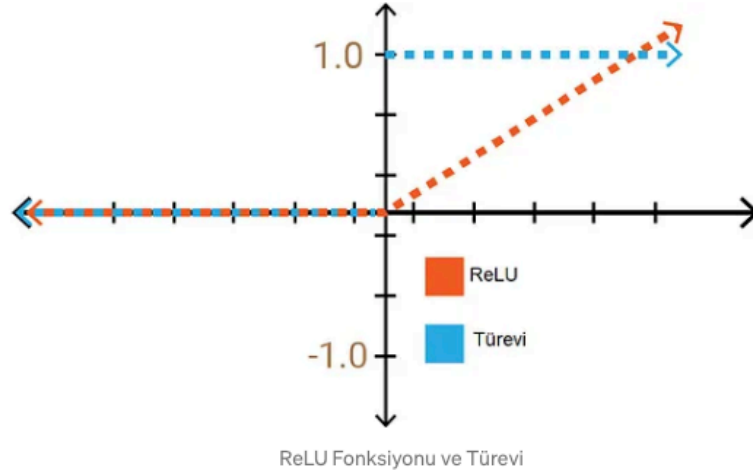
Düzeltilmiş Doğrusal Birim Aktivasyon Fonksiyonu, yapay sinir ağlarında doğrusallığı bozmak ve non-lineerliği sağlamak için kullanılır. Bu basit bir fonksiyondur ve giriş değeri pozitifse çıkışı giriş değeri olarak, negatifse ise sıfır olarak verir.

ReLU, LSTM ağlarının bazı varyasyonları ve Kapılı Yinelemeli Birim (GRU) ağlarında sıklıkla kullanılan bir aktivasyon fonksiyonudur. ReLu, özellikle büyük boyutlu veri kümeleriyle çalışan ağlar için tercih edilir, çünkü hesaplamaları hızlıdır ve ağın performansını artırır.

GRU ağlarında ReLu, kapıların (gates) kontrolünde kullanılır ve bu sayede ağın daha doğru sonuçlar vermesine yardımcı olur. LSTM ağlarında ise ReLu, hücre durumunu güncellemek için kullanılır. Düzeltilmiş Doğrusal Birim Aktivasyon Fonksiyonu ve ReLu gibi doğru aktivasyon fonksiyonlarının seçimi, yapay sinir ağlarının daha iyi performans göstermesini ve daha doğru sonuçlar vermesini sağlar. (Kizrak, 2019)

$$f(x) = \max(0, x)$$

Burada "x", fonksiyonun girdisini temsil eder ve "max" işlemi, girdi değerinin sıfırdan büyük olup olmadığını kontrol eder. Eğer girdi sıfırdan büyükse, fonksiyon girdiyi çıktı olarak verir. Eğer girdi sıfır veya daha küçük ise, fonksiyon sıfır çıktısı verir.



Şekil 5 ReLU Aktivasyon Fonksiyonu ve Türevi (Kızrak, 2019)

ReLU fonksiyonunun tek bir parametresi vardır, o da girdi değeridir. ReLU, sıfırın altındaki değerleri kestiği için aktivasyon fonksiyonu olarak tercih edilir. Bu sayede, yapay sinir ağlarının daha hızlı ve etkili bir şekilde öğrenmesine yardımcı olur. Ayrıca, sigmoid ve hiperbolik tanjant gibi diğer aktivasyon fonksiyonlarına göre hesaplama maliyeti daha düşüktür.

Aktivasyon fonksiyonları, yapay sinir ağlarının temel bir bileşenidir ve doğru bir şekilde seçilmesi, ağın performansını büyük ölçüde etkiler. (M. H. Essai Ali, 2022). Aktivasyon fonksiyonları, girdi verilerinin ağırlıklı toplamını alarak, çıktı değerini belirler ve bu değer, bir sonraki katmana aktarılır. Bu nedenle, aktivasyon fonksiyonları, ağın doğru çalışması için çok önemlidir.

Eldeki verileri işleyerek birçok farklı uygulama alanında kullanılan yapay sinir ağları, birçok farklı katman ve her katmanda birçok farklı nöron ile oluşturulurlar. Her nöronun çıktısı, aktivasyon fonksiyonu adı verilen bir matematiksel işleme tabi tutulur. Bu fonksiyonlar, ağın öğrenme sürecini belirleyen önemli bir faktördür.

Aktivasyon fonksiyonları, doğrusal ve non-lineer fonksiyonlar olarak iki ana gruba ayrılır. Doğrusal fonksiyonlar ağın yapısını basitleştirerek öğrenme sürecini hızlandırabilir ancak ağın daha kompleks örüntüleri öğrenmesini sınırlayarak

performansını kısıtlayabilir. Non-linear fonksiyonlar ise ađın öğrenme kapasitesini artırarak daha kompleks örüntüleri öğrenmesine izin verir ve daha yüksek bir performans sağlar.

Aktivasyon fonksiyonlarının türevlenebilir olması, geriye doğru yayılım (backpropagation) sırasında kullanıldığı için oldukça önemlidir. Bu sayede ađın eğitim süreci başarılı bir şekilde gerçekleştirilebilir. Ayrıca, aktivasyon fonksiyonları, ađın çıktılarını belirli bir aralıkta tutarak performansını artırabilir ve aşırı uyumaya (overfitting) karşı koruma sağlayabilir.

Aktivasyon fonksiyonlarının hesaplamalarının hızlı olması, büyük boyutlu veri kümeleri ile çalışan ađlar için de oldukça önemlidir. Bu sayede ađın performansı artırılabilir ve eğitim süreci hızlandırılabilir. Doğru aktivasyon fonksiyonlarının seçimi, yapay sinir ađlarının daha iyi performans göstermesini ve daha doğru sonuçlar vermesini sağlar.

IV. VERİ SETİ VE YÖNTEM

Bu çalışmada kullanılan veriler, Avrupa Uzay Ajansı İklim Değişikliği Girişimi'nden (ESA CCI) elde edilmiştir (Merchant C. E., 2019) (Merchant C. J., 2020). Çalışma Zonguldak ve Bartın illerimizi kapsayan 40.95°E enlem ve 31.25°N boylam koordinatlarında yer alan Karadeniz kıyılarını kapsamaktadır. Veri kümesi, Kelvin cinsinden günlük ortalama deniz yüzeyi sıcaklığı (DYS) değerlerinin yanı sıra santigrat derece cinsinden ortalama sıcaklığı, sıcaklık belirsizliğini ve deniz buzu ile kaplı okyanusun oranını içerir. Veri seti 1982'den 2022'ye kadar olan dönemi kapsamaktadır.

Aylık ortalama DYS hesaplanarak hazırlanan verilere dayalı olarak, aylık ve yıllık ortalamalar incelenmiştir. Tablo 1'de, veri ön işleme sonrası formatıyla birlikte örneklenmiştir.

Çizelge 1 Veri İşleme Sonrası Format (Örnek)

	Tarih	Sıcaklık (°K)
1	198501	300
2	198502	312
3	312	298

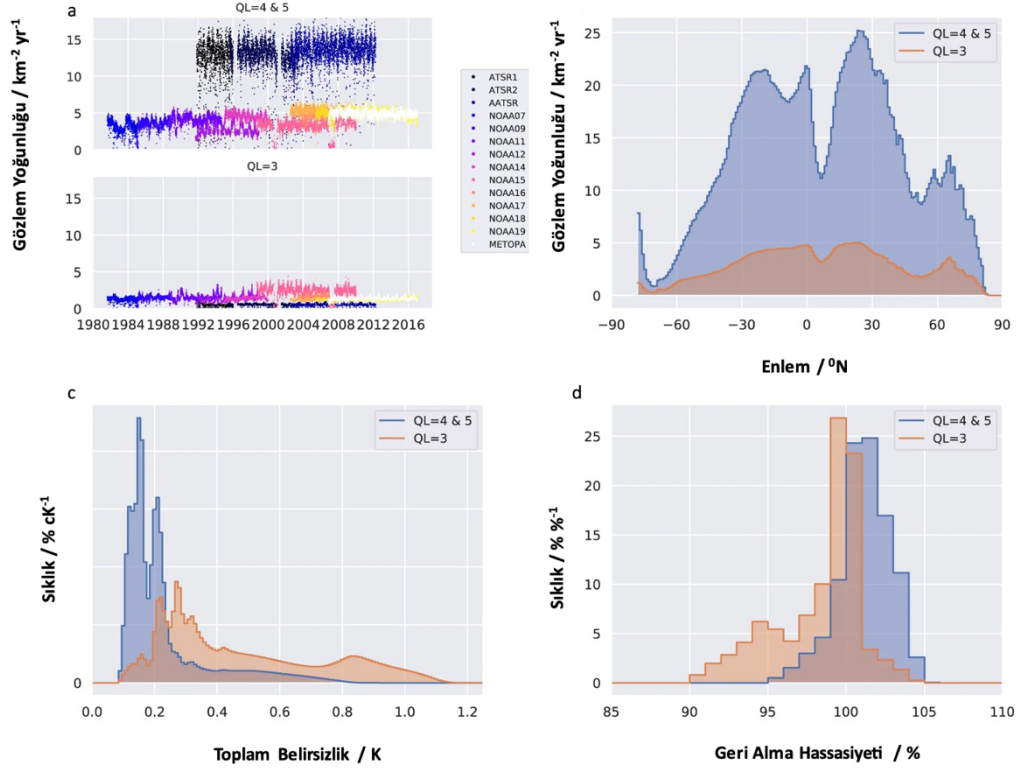
1981–2016 yıllarını kapsayan küresel deniz yüzeyi sıcaklığının (DYS) bir iklim veri kaydı (İVK), termal kızılötesi parlaklığın 4×10^{12} ölçümlerinden geliştirilmiştir. Piksel DYS tahminleriyle temsil edilen alan 1 km² ile 45 km² arasındadır. Kaliteli gözlemlerin ortalama yoğunluğu $13 \text{ km}^{-2} \text{ yıl}^{-1}$ dir. DYS belirsizliği veri başına değerlendirilir, piksel DYS'leri için medyan belirsizlik 0,18 K'dir. Deniz şamandıra ölçümlerine göre uzun dönem değişim DYS'lerde yüksek güvenle sıfırın $0,003 \text{ K yıl}^{-1}$ içindedir, (20 yıllık rekor). Veriler doğal çözünürlükte sağlanır, 0,05° enlem-boylam çözünürlüğünde gridlerde (bireysel sensörler) ve günlük

0,05°lik gridler gözönüne alınır. SST'leri, günlük döngüye göre filtrelenmiştir, farklı derinliklerde DYS anomalileri Saptanır. Veri setinin hedef uygulamaları şunları içerir: İklim ve okyanus modeli değerlendirmesi; denizdeki değişim ve değişkenliğin miktarının belirlenmesi (deniz sıcaklık dalgaları dahil); iklim ve okyanus-atmosfer süreçleri ve okyanus ekolojisi, oşinografi ve jeofizikteki özel uygulamalar vb. Veriler MetOp uyduları kullanılarak elde edilmiştir.

Çizelge 2 Veri Seti Anahtar Kelimeler

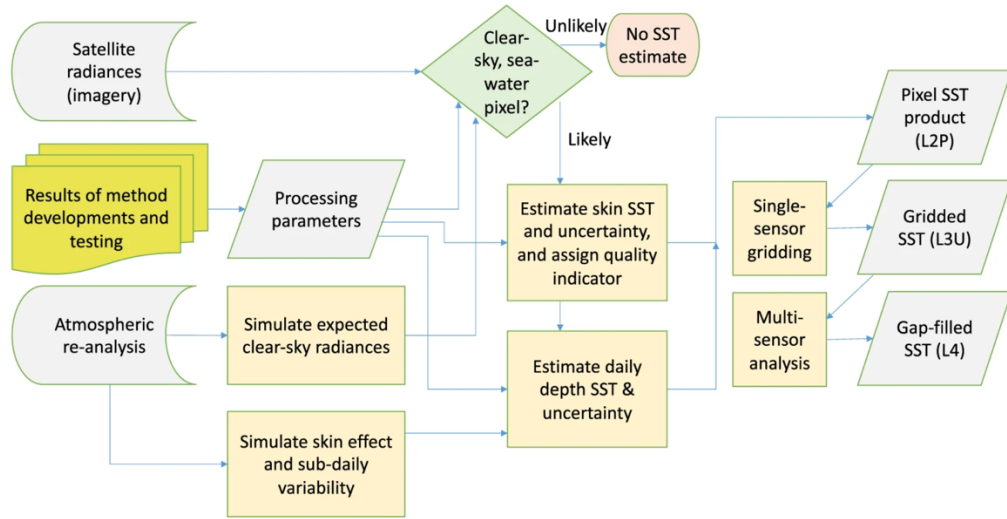
	Anahtar Kelimeler	Açıklama
1	Ölçümler	Su Sıcaklığı
2	Teknoloji Tipi	Uydu Görüntüleme
3	Faktör Tipleri	Zaman - Coğrafi Konum
4	Örnek Karakteristik - Çevre	Deniz
5	Örnek Karakteristik - Konum	Dünya (Gezegen)

Veriler, Dünya yörüngesindeki uydulardan elde edilen 4×10^{12} termal kızılötesi (TIR) ölçümlerinden küresel deniz yüzey sıcaklığı (DYS) verilerini içeren bir iklim veri kaydı sunmaktadır. 11 AVHRR ve 3 ATSR sensörü kullanılarak elde edilen veriler, bulutsuz, deniz buzu içermeyen okyanuslardan geçerli DYS değerleri sağlar. Her verinin geçerliliği ve belirsizliğine göre bir değerlendirme sağlanır ve iklim uygulamaları için önerilen kalite seviyesi 4 veya 5'tir. Elde edilen verilerin yoğunluğu zamana ve mekâna göre değişir Şekil 6.



Şekil 6 Tam Çözünürlüklü (L2P) Deniz Yüze Sıcaklıklarının Özellikleri

- Okyanusun birim alanı başına DYS gözlemlerinin sayısı (“gözlem yoğunluğu”), zaman içinde her bir sensör için göz önüne alınır. Her nokta, tek bir günün gözlemlerini temsil eder. Üst panel, iklim uygulamaları için önerilen kalite seviyeleri için verileri gösterir. (QL 4 ve 5), alt panel ise QL 3’ü gösterir.
- Tüm sensörlerin birleştirilmiş ve zaman serisi boyunca ortalaması alınmış birim alana göre enlem DYS yoğunluğu elde edilmektedir.
- Tüm sensörlerin zaman serisi boyunca birleştirilmiş DYS gözlemlerinin belirsizliği histogramı oluşturulur.
- Tüm sensörlerin birleştirilmiş zaman serisi boyunca değerlendirilmiş DYS gözlemlerinin hassasiyeti histogramı oluşturulur.



Şekil 7 İklim Veri Kayıtlarının Üretim Aşamaları

Şekil 7 farklı ürün seviyelerinin ilişkileri de dahil olmak üzere, iklim veri kayıtlarının üretim mantığına genel bir bakış sağlar. Deniz yüzeyi sıcaklığı (DYS) için iklim veri kayıtlarının üretimi şöyle gerçekleştirilir:

- Kalibre edilmiş, konumlandırılmış uydu verileri temel kaynak verilerini oluşturur.
- Uydudan gönderilen elektromanyetik dalgaların kalibrasyon tutarlılığı için birçok işleme parametresi belirlenir.
- Geçerli DYS'lerin hangi piksellerde elde edilebileceği olasılıklara dayalı (Bayesçi) piksel sınıflandırmasıyla belirlenir.
- Cilt DYS'si ayarlanarak, termal ışınların doğrudan ilişkili olduğu SST'ye (yerinde ölçümlerle karşılaştırılabilir bir nominal derinlikte) tahmin edilir.
- Geçmiş atmosferik koşulların tahminleri, geri alma ve yüzeyden derinliğe ayarlama adımlarında kullanılır.
- Düşük seviyeli DYS iklim veri kayıtları, giriş görüntüsü tam piksel çözünürlüğündedir, ancak bulutlar nedeniyle boşluklar içerir. L2P DYS'ler, sensör L3U ürünleri elde etmek için 0,05°'lik bir gridler üzerinde ortalaması alınır.
- L3U SST'ler, belirsizlik tahminleriyle birleştirilerek, birçok sensörde enterpolasyonlu, boşluklarla dolu bir L4 analizi oluşturulur.

A. Avhrr / 3 MetOp uydusu

(Avhrr MetOp, 2022) Gelişmiş Çok Yüksek Çözünürlüklü Radyometre (AVHRR/3), Ulusal Okyanus ve Atmosfer İdaresi (NOAA) tarafından MetOp-A, B ve C'de uçmak üzere sağlanan Amerikan cihazlarının tamamlayıcılarından biridir.

AVHRR/3, Dünya yüzeyini 0,58- 12,5 mikron aralığında altı spektral bantta tarar. Kara, su ve bulutların gündüz ve gece görüntülenmesini sağlar, deniz yüzey sıcaklığını, buz, kar ve bitki örtüsünü ölçer.

AVHRR/3, elektromanyetik spektrumun görünür ve kızılötesi (IR) bölümlerindeki enerjiyi saptayan altı kanallı bir görüntüleme radyometresidir. Cihaz, karadan, denizden, bulutlardan ve araya giren atmosferden yansıyan güneş enerjisini (görünür ve yakın kızılötesi) ve yayılan termal enerjiyi ölçer. Cihaz, en düşük seviyede 1,1 km (0,69 mi) nominal uzay çözünürlüğü sağlayan 1,3 mili radyan anlık görüş alanına (IFOV) sahiptir. Sürekli dönen eliptik bir tarama aynası, Dünya'yı nadirden $\pm 55,4^\circ$ 'den tarayarak çapraz tarama sağlar. Ayna, sürekli kapsama alanı sağlamak için saniyede altı devirde tarama yapar.

Cihaz, düşük ışık enerjisi tespiti sağlayan güneşte görünür kanallarda spektral ve kazanç iyileştirmeleri sağlar. Kanal 3A, 1,6 mikronda kar, buz ve bulut ayrımı sağlar. Kanal 3A, beş sürekli veri kanalı sağlamak için 3B olarak adlandırılan 3,7 mikronluk kanalla zaman paylaşımli olacaktır. Aletin optik boşluğuna ve detektörlerine güneş ışığı çarpmasını azaltmak için harici bir güneş kalkanı ve dahili bir bölme eklenmiştir. MetOp uydusunun Dünya'ya en yakın yörünge yükseliği 530 km'dir (Europe's first meteorological satellite in polar orbit ends its run, 2021).

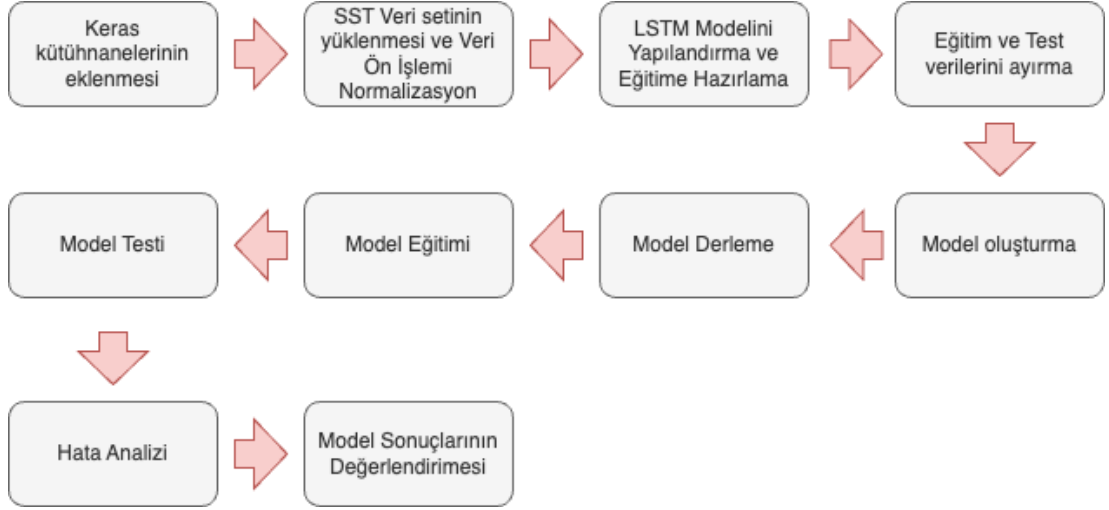
V. UYGULAMA

LSTM modeli, doğal dil işleme, konuşma tanıma, çeviri gibi alanlarda son yıllarda başarılı sonuçlar elde eden bir derin öğrenme modelidir. LSTM modeli, önceki zaman adımlarında aldığı bilgiyi hatırlama yeteneği sayesinde, zaman içindeki bağımlılıkları tanımlama konusunda oldukça başarılıdır.

LSTM modelinin uygulanması, derin öğrenme kütüphaneleri aracılığıyla kolayca gerçekleştirilebilir. Bu kapsamda, Keras kütüphanesi LSTM modelinin uygulanması için popüler bir tercih haline gelmiştir. Keras kütüphanesi, basit bir kullanım arayüzü ve yüksek seviye fonksiyonları ile LSTM modelinin yapılandırılmasını ve eğitilmesini kolaylaştırır.

LSTM modelinin eğitimi ve test edilmesi, veri setinin boyutuna ve özelliklerine göre değişkenlik gösterir. Bu nedenle, LSTM modelinin doğru şekilde eğitilmesi ve test edilmesi için özenle seçilmiş veri setleri kullanılmalıdır. Ayrıca, LSTM modelinin hiper parametrelerinin de doğru bir şekilde ayarlanması, modelin performansını büyük ölçüde etkileyebilir.

Bu nedenle, araştırma çalışmalarında LSTM modelinin doğru bir şekilde yapılandırılması ve eğitilmesi için literatürdeki uygun yöntemlerin incelenmesi gerekmektedir. LSTM modelinin avantajları ve dezavantajları ayrıca göz önünde bulundurulmalıdır. LSTM modeli, doğal dil işleme ve konuşma tanıma alanlarında oldukça başarılı sonuçlar vermesine rağmen, modelin karmaşıklığı nedeniyle diğer basit modellerle karşılaştırıldığında daha yavaş çalışabilir. Bu nedenle, LSTM modeli uygulanmadan önce veri setinin özellikleri ve problemin gereklilikleri dikkate alınmalıdır. Bir LSTM modelinin akış şeması Şekil 8’de gibidir.



Şekil 8 LSTM Model Akış Şeması

A. Araştırmanın Amacı

Bu çalışma, deniz yüzeyi sıcaklıklarının değişimlerini modellenmesi amacıyla makine öğrenme teknikleri kullanılarak yapılmıştır. Araştırma, iklim değişikliğinin etkilerini anlamak ve makine öğrenimi tabanlı sıcaklık tahmin modelinin doğruluğunu artırmanın mümkün olup olmadığını test etmektir.

B. Bulgular

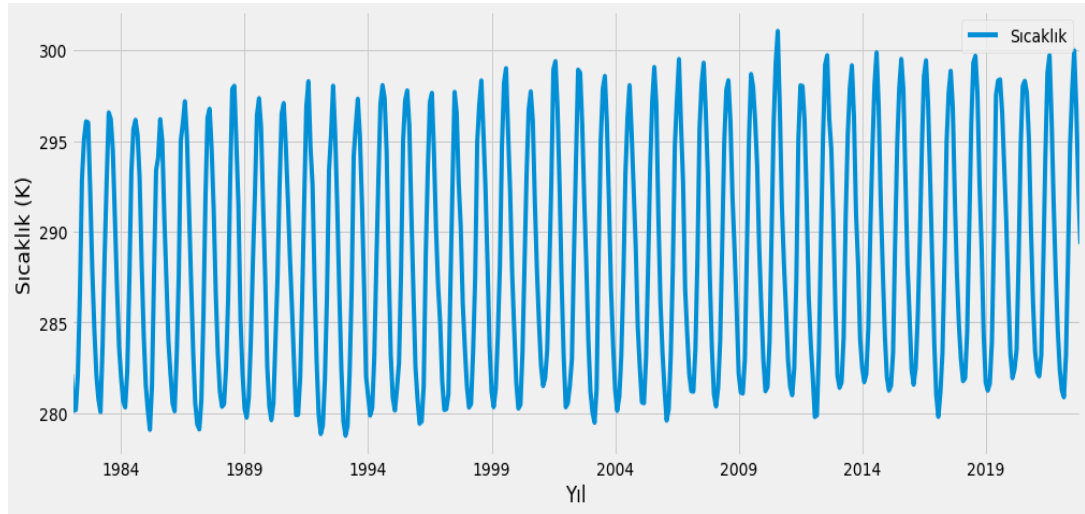
Bu çalışma içerisinde kullanılan yöntemlere ait bulgu sonuçları bu bölümde yer almaktadır.

1. ARIMA Analiz Sonuçları

Elde edilen verilerin analizi sonucunda, deniz yüzey sıcaklıklarının artış trendinde olduğu gözlemlenmiştir. Grafik üzerinde yapılan incelemelerde, en düşük sıcaklığın 1992-02-26 tarihinde 278.21 Kelvin derece, en yüksek sıcaklığın ise 2010-08-14 tarihinde 302.47 Kelvin derece olarak kaydedildiği görülmüştür. Ayrıca, son yıllara bakıldığında sıcaklıkların artış trendinin daha belirgin hale geldiği tespit edilmiştir.

Bu sonuçlar, iklim değişikliği ve küresel ısınma gibi çevresel faktörlerin deniz yüzey sıcaklıklarına olan etkisini ortaya koyabileceği gibi çevresel faktörler ve insan aktiviteleri nedeniyle oluşabilir. Örneğin, fosil yakıtların kullanımı, sera gazı emisyonlarına neden olarak atmosferdeki sıcaklığı artırır ve sonuçta deniz yüzey

sıcaklıklarını da etkiler. Bu nedenle nüfusun ve sanayinin artmasının deniz yüzeyi sıcaklıklarına olan etkisini araştırmak için etkili bir neden olabilir. Bu veri setindeki sıcaklık trendleri, deniz yüzey sıcaklıklarının arttığını göstermektedir.



Şekil 9 Yıllara Göre Zonguldak-Bartın Sıcaklık

ARIMA modellerinde AIC (Akaike Information Criterion) Akaike ölçütü bir modelin kalitesinin değerlendirilmesinde kullanılan bir ölçüttür. AIC, modelin performansını hem uyum hem de aşırı uyum açısından değerlendirir. AIC, düşük hatayla uyumlu bir modeli tercih ederken aşırı uyumdan kaçınır. Bu nedenle, AIC değeri düşük olan bir model, daha iyi bir model olarak kabul edilir.

AIC değeri, modelin karmaşıklığı ile log-likelihood değerinin bir fonksiyonudur. AIC değerinin hesaplanması için, öncelikle modelin log-likelihood değeri hesaplanır. Log-likelihood, modelin uyum kalitesinin bir ölçüsüdür ve logaritmik bir olasılık değeridir. Modelin karmaşıklığı arttıkça, log-likelihood değeri azalır. AIC, log-likelihood değerine, modelin karmaşıklığının (örneğin, p, d ve q değerleri gibi) giderilmesi için eklenen bir terimle hesaplanır. Bu nedenle, daha karmaşık bir modelin AIC değeri daha yüksek olacaktır.

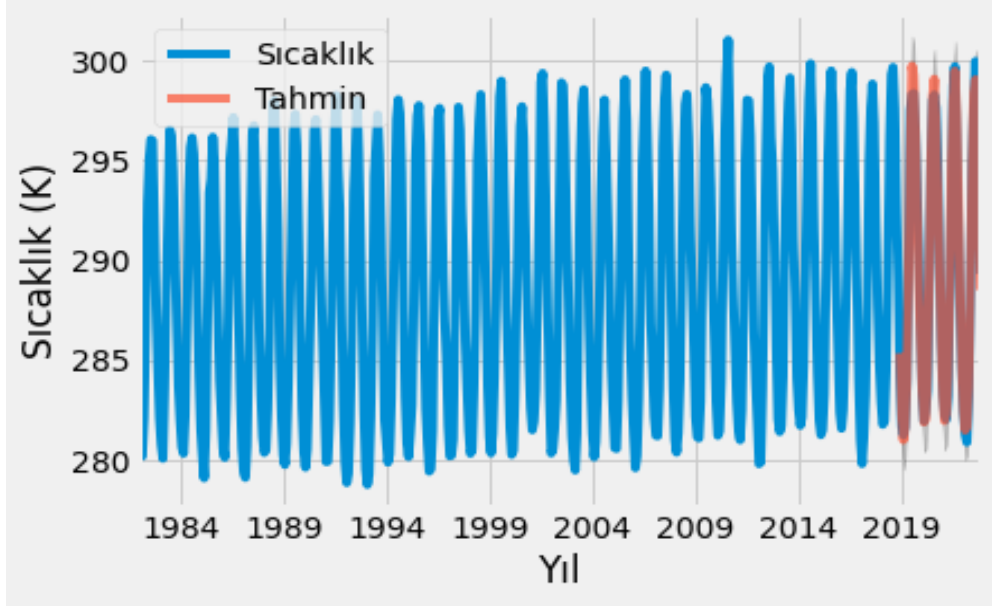
AIC, özellikle regresyon analizi gibi büyük veri setleri üzerinde çalışan araştırmacılar tarafından sıklıkla kullanılır. Ancak, yalnızca iki model arasındaki karşılaştırmalar için kullanılması önerilir.

ARIMA modelin de order (düzenleme) (p, q, d) ve seasonal order (mevsimsel düzen) için hesaplanan AIC skorları hesaplanmıştır. En düşük AIC skoru, ARIMA (1, 0, 1) x (0, 1, 1, 12) uygulama içerisinde kullanılmıştır.

Çizelge 3 AIC Değerleri Zonguldak-Bartın

	ARIMA Parametreleri	AIC Skor
1	ARIMA (0, 0, 0) x (0, 0, 0, 12)	6944.87
2	ARIMA (0, 0, 0) x (0, 1, 0, 12)	1511.11
3	ARIMA (0, 0, 0) x (1, 0, 1, 12)	1271.88
4	ARIMA (0, 0, 1) x (0, 0, 0, 12)	6262.79
5	ARIMA (0, 0, 1) x (0, 1, 1, 12)	1133.97
6	ARIMA (0, 0, 1) x (1, 0, 1, 12)	1134.24
7	ARIMA (0, 1, 0) x (0, 0, 0, 12)	2613.1
8
9	ARIMA (1, 1, 1) x (0, 1, 1, 12)	1088.62
10	ARIMA (1, 1, 1) x (1, 0, 0, 12)	1360.91
11	ARIMA (1, 1, 1) x (1, 1, 1, 12)	1098.14

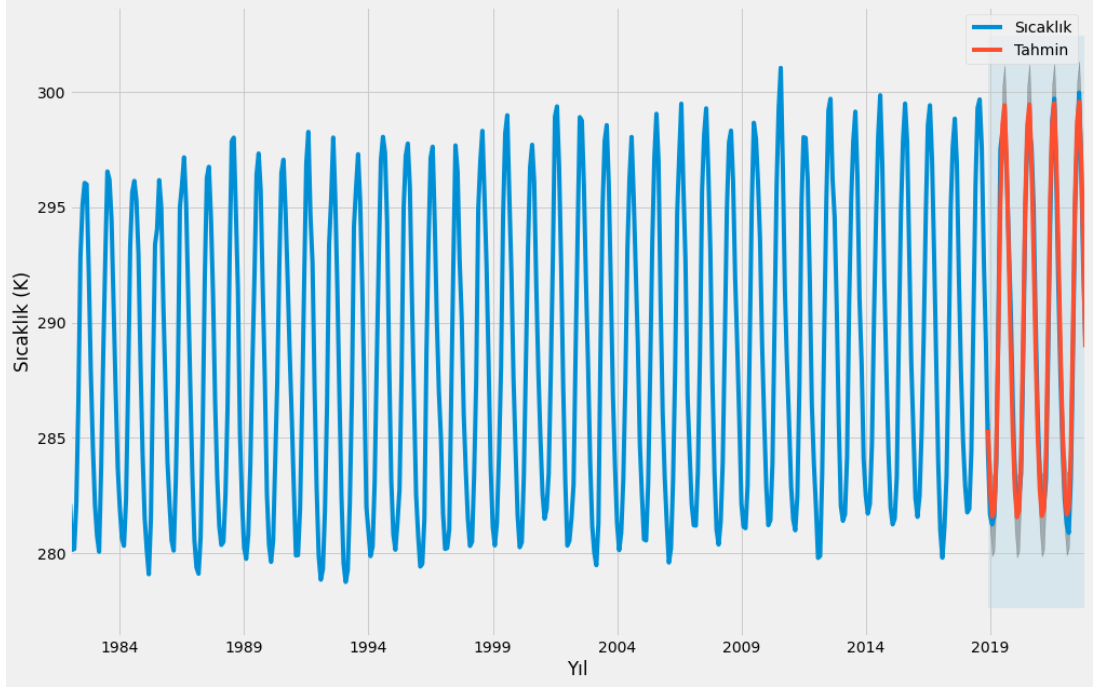
Tez çalışmasının ana amacı, modelin tahmin hatalarının ilintisiz olmasını ve ortalamanın sıfır ortalama ile normal dağılmasını sağlamaktır. Eğer mevsimsel ARIMA modeli bu özellikleri karşılamıyorsa, bu modelin daha da iyileştirilmesi gerektiği anlamına gelir.



Şekil 10 ARIMA Modeli Eğitim- Test / Zonguldak-Bartın

Eldeki veri setini kullanarak ARIMA eğitim modeli uygulandığında, 1982-2018 yılları arasındaki verileri eğitim amaçlı kullanırken, 2018 ve 2022 yılları arasındaki verileri test etmek için ayrılmıştır. Bu süreçte, modelin performansını ölçmek için ortalama kare hatası (MSE) oranını hesaplanmıştır. Sonuç olarak 0.0713 değerini elde edilmiştir. Şekilden de anlaşılacağı üzere, modelin başarılı olduğu gözlemlenmektedir. Bu sonuçlar, modelin kullanılabilirliğini ve gelecekteki deniz yüzey sıcaklığı tahminlerinde doğru sonuçlar sağlayabileceğini göstermektedir.

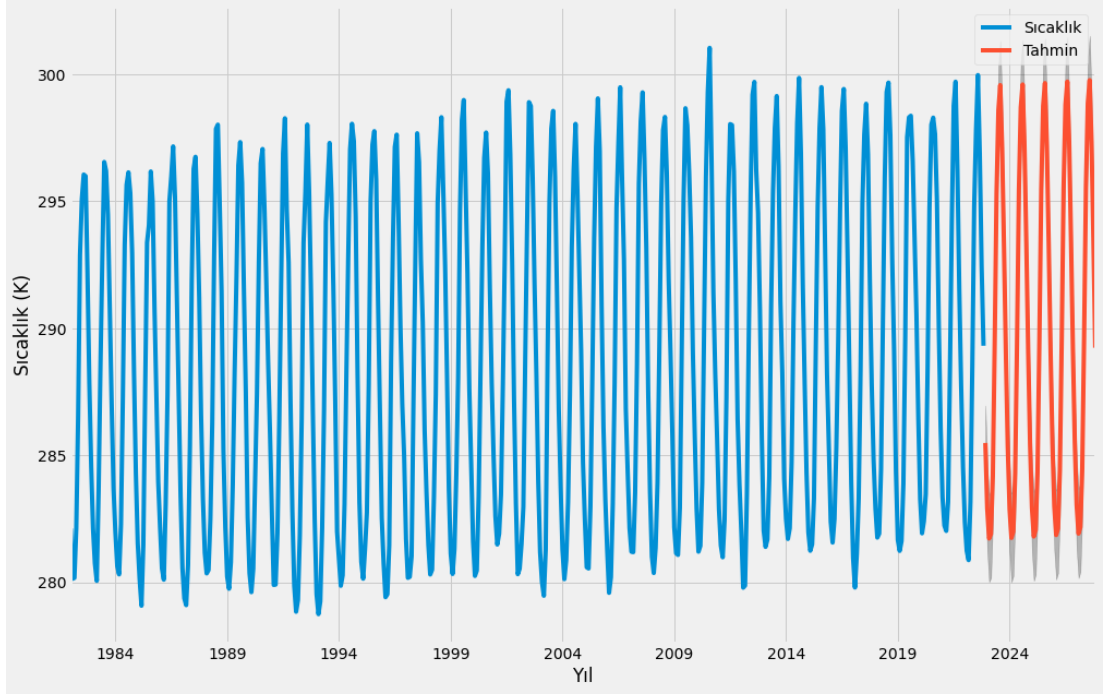
Ancak, bu modelden tamamen emin olmak için dinamik tahminler veri seti olarak kullanılabilir. Bu durumda, zaman serisinden belirli bir noktaya kadar olan bilgileri kullanıp, bundan sonrası için önceki tahmin edilen zaman noktalarından alınan değerler kullanılarak yeni tahminler oluşturulur.



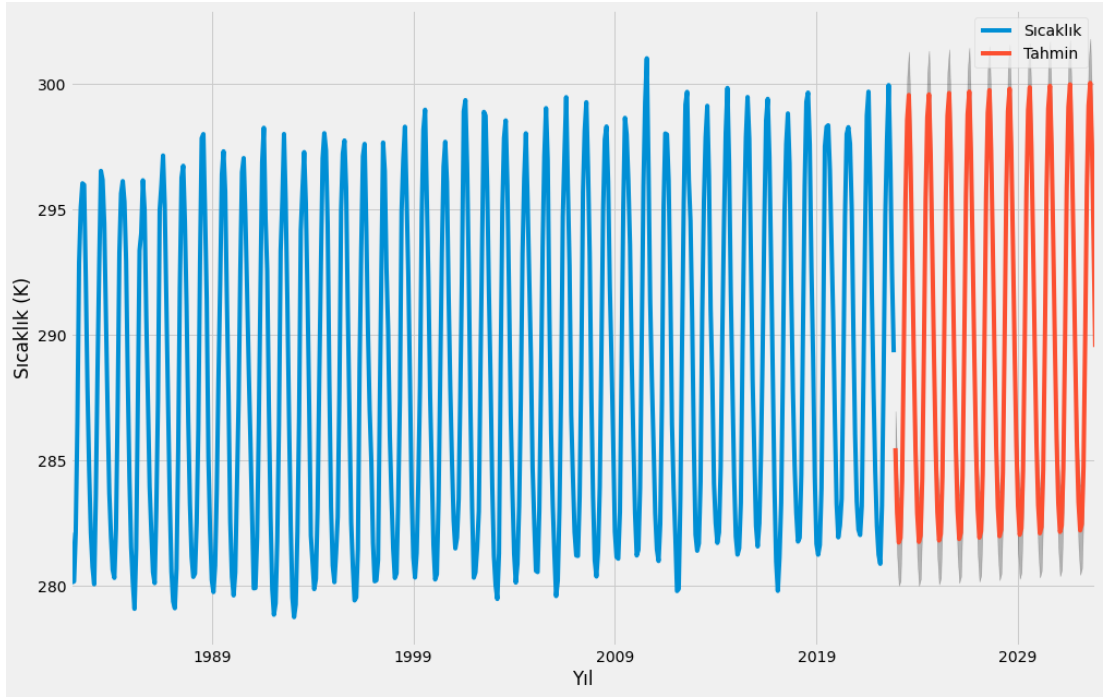
Şekil 11 ARIMA Modeli Eğitim / Test Dinamik Veriler Zonguldak-Bartın

Dinamik tahminlerden elde edilen tahmin edilen yeni değerler, 0.16'lık bir MSE verir. Bu bir önceki MSE oranından 2 kat daha yüksek olmasına rağmen yine mühendislik uygulamaları için bu modelin geçerli olduğunu doğrulamaktadır.

Elde edilen geçerli ARIMA modeli ile 2022 yılından başlamak üzere gelecek 5 ve 10 yıllık dönem için tahmin edilen deniz yüzeyi sıcaklıkları Şekil 12 ve Şekil 13'de ki gibidir.



Şekil 12 ARIMA Model ile Gelecek 5 Yıl Tahmini-Zonguldak-Bartın



Şekil 13 ARIMA Model ile Gelecek 10 Yıl Tahmini-Zonguldak-Bartın

Zonguldak ili için elde edilen model ile gelecek 60 ay (5 yıl- Grafik 1 ve 120 ay (10 yıl) – Grafik 5 için deniz yüzey sıcaklığı tahmini yapılmıştır. Bu sonuçlara göre deniz yüzey sıcaklık ortalaması yükseliş trendini koruduğu gözlemlenmektedir.

2. LSTM Analiz Sonuçları

Tüm model RNN-LSTM özellikleri kullanılarak eğitilmiş ve değerlendirilmiştir. Eğitim ve doğrulama verileri için özel bir yöntem belirlenmiş ve veri miktarı saptanmıştır. Eğitim veri seti için son 60 ayın verileri hariç tüm veriler kullanılmıştır. Test veri seti ise sadece son 60 ayın verileri göz önüne alınmıştır. Toplam veri boyu günlük veri seti üzerinde yapılan işlemler sonucu 491 olarak belirlendi.

Tahmin modeline göre, bir sonraki ayın tahmin edilmesi için son 12 ayın verisi kullanılmış ve hiperparametreler bu düzen üzerine belirlenmiştir. Seçilen modellerin öğrenme oranı 0.001, toplam veri boyutu 491 ve epoch sayısı 100 ve batch size 32'dir. RNN-LSTM modeli aşağıdaki Çizelge 4'te detaylar belirtilmiştir.

Çizelge 4 Model: "RNN-LSTM model parametreleri. Model: "model""

	Layer (type)	Output Shape	Param #
1	lstm_1 (LSTM)	(None, 64)	16896
2	dense_1 (DENSE)	(None, 1)	65

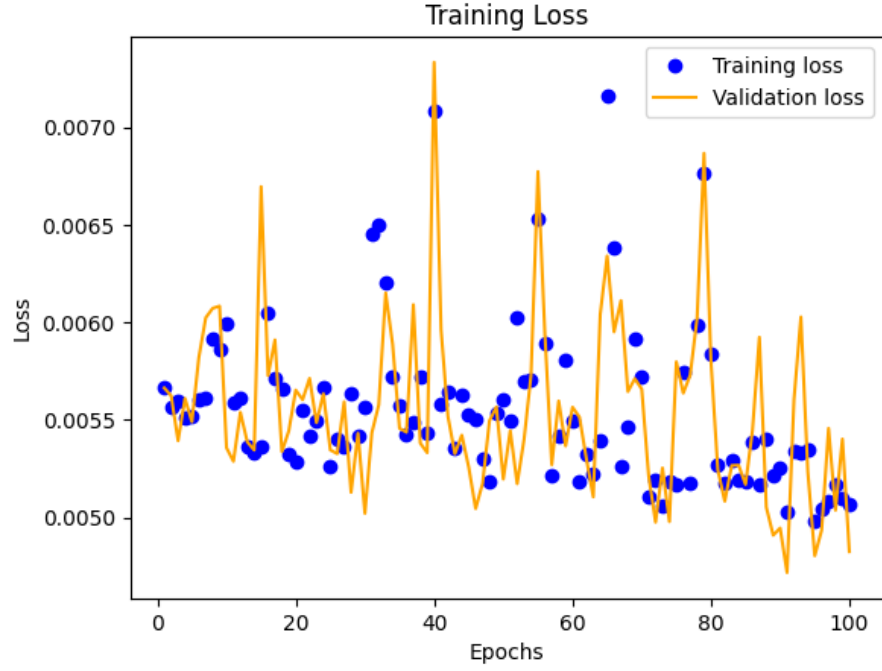
Total params: 16,961

Trainable params: 16,961

Non-trainable params: 0

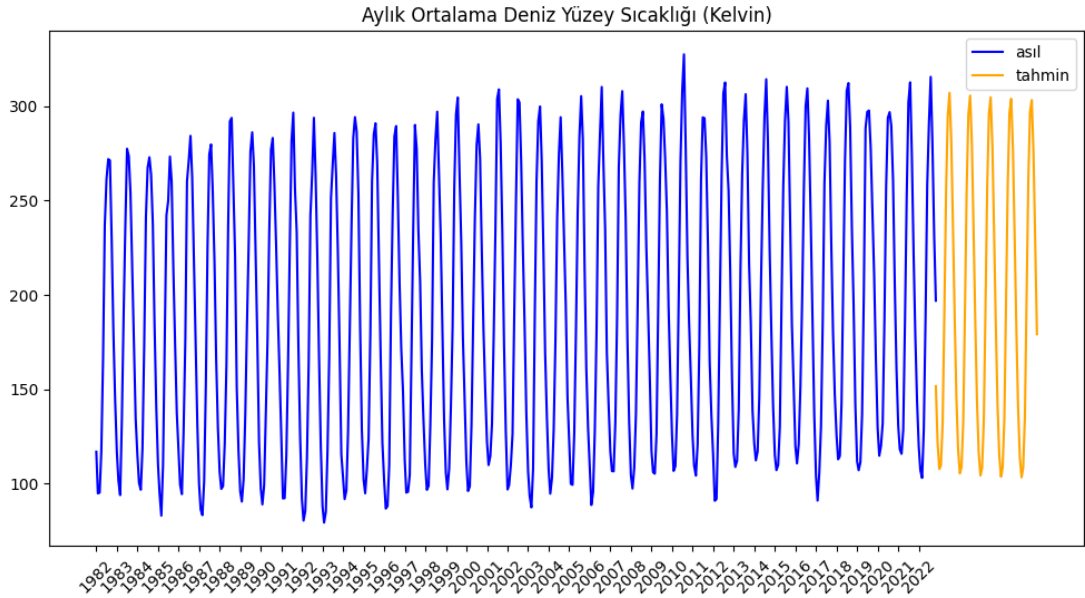
Kullanılan optimizasyon fonksiyonu ADAM, loss fonksiyonu ise MSE olarak belirlenmiştir. MSE sonucu 0.005 gibi düşük bir oran olarak gözlemlenmiştir.

Error! Reference source not found.'de görüleceği üzere, eğitim kaybı değerleri değişken bir görünüme sahip olmasına rağmen öğrenim eğrisi sabit değerlere doğru azalarak değişim göstermektedir.



Şekil 14 LSTM-Zonguldak/Bartın Loss/Epoch

Zonguldak-Bartın bölgesindeki deniz yüzey sıcaklıklarının 2022 ve öncesindeki verilerine bakıldığında azalma trendinde olduğu gözlemlenmektedir. LSTM modeli tarafından yapılan 5 yıllık tahminler ise bu azalma trendinin devam edeceğini doğrulamaktadır **Error! Reference source not found.**



Şekil 15 LSTM-Zonguldak/Bartın 5 Yıllık Tahmin Grafik

3. Eğitim Performans Optimizasyonu ve Ölçümleri

Makine öğrenme algoritmalarında, eğitim sürecinin hızlandırılması, modelin başarı performansında kritik bir rol oynamaktadır. Yüksek performans gösteren modeller, sınırlı miktarda eğitim verisiyle daha hızlı bir şekilde eğitilebilir ve daha üstün sonuçlar üretebilir.

Eğitim performansının artırılması, makine öğrenme modelinin başarısını olumlu yönde etkileyen önemli bir faktördür. Daha hızlı ve etkili bir eğitim süreci, modelin daha iyi bir şekilde öğrenmesini ve veriler arasındaki ilişkileri daha doğru bir şekilde yakalamasını sağlar. Bu da daha keskin tahminler ve daha yüksek doğruluk oranlarıyla sonuçlanır.

Normalizasyon, genellikle makine öğrenmesi için veri hazırlığı sürecinde uygulanan bir tekniktir. Normalizasyonun amacı, veri setindeki sayısal sütunların değerlerini ortak bir ölçüğe dönüştürmektir. LSTM gibi derin öğrenme modellerinde normalizasyon, daha stabil bir eğitim süreci sağlar ve aşırı öğrenmeyi azaltabilir.

Makine öğrenimde en popüler normalizasyon yöntemi olan Min-Max normalizasyonu, veri setindeki değerleri belirli bir aralığa dönüştürmek için kullanılan bir normalizasyon yöntemidir. Bu yöntemde, veri setindeki her bir değer, minimum ve maksimum değerler arasında yeni bir değer aralığına yeniden ölçeklendirilir. Min-Max normalizasyonu aşağıdaki formülle uygulanır.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Burada, x , orijinal veri setindeki bir değeri temsil ederken, x' ise normalleştirilmiş değerini ifade eder. $\min(x)$ veri setindeki en küçük değeri, $\max(x)$ ise veri setindeki en büyük değeri temsil eder.

Min-Max normalizasyonu, veri setindeki değerleri 0 ile 1 arasında bir aralığa dönüştürür. Bu, veri setinin ölçeklendirilmesini ve karşılaştırmaların daha anlamlı hale gelmesini sağlar. Aynı zamanda, veri setindeki anormalliklerin etkisini minimize ederek modelin daha istikrarlı bir şekilde öğrenmesine yardımcı olur.

Yaptığımız normalizasyon testleri x86 mimarisine sahip Intel işlemci, 16 GB RAM ve 2100 MB/S okuma 2700 MB/S yazma kapasiteli bir SSD diske sahip bir cihaz

üzerinde test edilmiştir. Test sonucuna göre 14943 veri noktasının normalizasyonu için 0.0018 saniyede normalizasyon işlemi tamamlanmıştır.

VI. SONUÇ VE ÖNERİLER

Bu çalışmada, Deniz Yüzey Sıcaklığının (DYS) ARIMA yöntemi kullanılarak Zonguldak ve Bartın illerindeki sıcaklık trendleri modellenmiş ARIMA ve LSTM yöntemleri ile incelenmiş ve gelecek tahmini yapılmıştır.

Çalışma 1982-2022'ye kadar olan dönemi kapsamaktadır. ARIMA model 0.07 gibi düşük bir RMSE oranına sahip olmasına karşın LSTM modeli 0.005 gibi daha düşük bir RMSE oranına sahiptir. Bu durumda LSTM modelinin ARIMA modeline göre daha doğru sonuçlar verdiği gözlemlenmiştir.

Analiz sonuçlarına göre, çalışma dönemi boyunca Zonguldak ve Bartın bölgelerinde DYS'da yavaş bir artış gözlemlenmiştir. (Aydınlı, 2022) Karadeniz bölgesinde yapılan başka bir çalışmada deniz yüzey sıcaklığı tahminleri için LSTM modelinin önemli bir potansiyele sahip olduğu belirtilmiştir. Doğrulama için kullanılan RMSE değerleri, Ordu, Samsun ve Sinop için sırasıyla 0.914, 0.815 ve 0.802 RMSE olarak hesaplanmıştır.

Çin Kıyı denizlerinde yapılan (Zhang Q. W., 2017) çalışması da LSTM modelinin DYS tahmininde etkili bir performans sergilediğini göstermektedir. Bohai Denizi için hesaplanan RMSE değeri 1.1363 olarak bulunmuştur.

(Sarkar, Janardhan, & Roy, 2020) çalışmasında da LSTM modelinin deniz yüzey sıcaklıklarının tahmininde başarılı olduğu görülmektedir. Yapılan tahminlerde RMSE değerleri 0.37'den küçük bulunmuştur.

Sonuç olarak, LSTM modeli DYS tahmininde etkili bir araç olarak kullanılabilir. Bu çalışma, deniz yüzey sıcaklığının gelecekteki eğilimlerini belirlemek ve ilgili bölgelerdeki iklim değişikliklerini anlamak için önemli bir adımdır. Bir sonraki çalışmada, gelecekteki çalışmalarda daha geniş bir coğrafi kapsamı ve daha uzun süreli veri setleri kullanarak daha kapsamlı analizler yapılabilir.

VII. KAYNAKLAR

KİTAPLAR

HOEGH-GULDBERG, O. R. (2014). **The Ocean**. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

TAI, K. S. (2015). **Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks**. Beijing, China: Association for Computational Linguistics.

OKUT, H. (2021). **Deep Learning for Subtyping and Prediction of Diseases: Long-Short Term Memory**. IntechOpen.

MAKALELER

MANABE, S. A. (1988). "Two stable equilibria of a coupled ocean-atmosphere model", **Journal of Climate**, 841-866.

HUI SHI, F.-F. J.-R. (2022). "Global decline in ocean memory over the 21st century", **Science Advances**, N/A.

MOHAMED, B., IBRAHIM, O., & NAGY, H. (2022). "Sea Surface Temperature Variability and Marine Heatwaves in the Black Sea", **Remote Sensing**, 2383.

JIA X, J. Q. (2022). "Prediction of Sea Surface Temperature in the East China Sea Based on LSTM Neural Network", **Remote Sensing**, 3300.

HAKAN OKTAY AYDINLI, A. E.-A. (2022). "Sea surface temperature prediction model for the Black Sea by employing time-series satellite data: a machine learning approach", **Applied Geomatics**, 669–678.

ZHANG, Q. W. (2017). "Prediction of Sea Surface Temperature Using Long Short-Term Memory", **IEEE Geoscience and Remote Sensing Letters**, 1745-1749.

ZHANG, G. P. (2003). "Time series forecasting using a hybrid ARIMA and neural network model", **Neurocomputing**, 159-175.

- CAVANAUGH, J. E. (2019). "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements", **Wiley Interdisciplinary Reviews: Computational Statistics**, e1460.
- HAN, C. &. (2023). "Challenge and Opportunity: Deep Learning-Based Stock Price Prediction by Using Bi-Directional LSTM Model", **Frontiers in Business, Economics and Management**, 51–54.
- HOCHREITER, S. &. (1997). "Long Short-term Memory", **Neural Computation**, 1735-80.
- Li, M. (2021). "Star Business Value Prediction based on Sentiment Analysis", **BCP Business & Management**, 229–232.
- PAL, S. G. (2018). "Sentiment Analysis in the Light of LSTM Recurrent Neural Networks", **International Journal of Synthetic Emotions (IJSE)**, 33-39.
- GREFF, K. S. (2017). "LSTM: A Search Space Odyssey", **IEEE Transactions on Neural Networks and Learning Systems**, 2222-2232.
- ALAMRI, N. M. (2023). "Optimizing the Parameters of Long Short-Term Memory Networks Using the Bees Algorithm", **Applied Sciences**, 2536.
- KHAN, H. R. (2023). "Early Identification of Crop Type for Smallholder Farming Systems Using Deep Learning on Time-Series Sentinel-2 Imagery", **Sensors**, 1779.
- M. H. ESSAI ALI, A. B.-R. (2022). "Developing Novel Activation Functions Based Deep Learning LSTM for Classification", **IEEE Access**, 97259-97275.
- MERCHANT, C. E. (2019). "Satellite-based time-series of sea-surface temperature since 1981 for climate applications", **Scientific Data**, 223.
- MERCHANT, C. J. (2020). "Adjusting for desert-dust-related biases in a climate data record of sea surface temperature", **Remote Sensing**, 2554.
- Y. FANG, C. Z. (2019). "Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism", **IEEE Access**, 56329-56340.

MANABE, S. A. (1988). "Two stable equilibria of a coupled ocean-atmosphere model". **Journal of Climate**, 841-866.

SARKAR, P., JANARDHAN, P., & ROY, P. (2020). "Prediction of sea surface temperatures using deep learning neural networks", **SN Applied Sciences**, 1458.

ELEKTRONİK KAYNAKLAR

NOAA. (2017, 05 10). "What is eutrophication", National Ocean Service: <https://oceanservice.noaa.gov/facts/eutrophication.html>

KIZRAK, A. (2019, 02 04). Medium. "Derin Öğrenme İçin Aktivasyon Fonksiyonlarının Karşılaştırılması", <https://ayyucekizrak.medium.com/derin-ogrenme-icin-aktivasyon-fonksiyonlarinin-karsilastirilmasi-cee17fd1d9cd>

URL-1 "AVHRR" MetOp. (2022). ESA: https://www.esa.int/Applications/Observing_the_Earth/Meteorological_missions/MetOp/About_AVHRR_3

URL-2 "Europe's first meteorological satellite in polar orbit ends its run", (2021, 11 15). EUMETSAT: <https://www.eumetsat.int/europes-first-meteorological-satellite-polar-orbit-ends-its-run>

DİĞER KAYNAKLAR

SIAMI-NAMINI, S. T. (2018). "A Comparison of ARIMA and LSTM in Forecasting Time Series". 17th IEEE International Conference on Machine Learning and Applications (*ICMLA*), (s. 1394-1401). Orlando, FL, USA.

EKLER

LSTM ALGORITMA

- TensorFlow, NumPy, Pandas, MinMaxScaler ve Matplotlib gibi gerekli kitaplıkları içe aktarın
- Veri setini (zonguldak-bartın-clean.csv) yükleyin ve sıcaklığı Kelvin'den Santigrat'a çevirerek ve tarih sütununu indeks olarak ayarlayarak önceden işleyin
- Verileri aylık sıklığa göre yeniden örnekleyin ve eğitim ve test kümelerine ayırın
- Özellik aralığı (-1, 1) ile MinMaxScaler'ı kullanarak verileri ölçeklendirin
- Ölçeklenmiş verileri ve dizi uzunluğunu girdi olarak alan ve iki sayısal dizi X ve y döndüren bir işlev tanımlayarak LSTM modeli için girdi ve çıktı dizileri oluşturun; burada X, girdi dizisini ve y tahmin edilecek çıktı değerini temsil eder
- 64 birimli bir LSTM katmanından ve tek çıkışlı bir Yoğun katmandan oluşan Keras Sequential API'yi kullanarak LSTM modelini oluşturun
- Modeli, kayıp işlevi ve Adam iyileştirici olarak ortalama karesel hata (MSE) ile derleyin
- Eğitim sırasında her dönemin sonunda MSE'yi yazdırmak için bir geri arama tanımlayın
- Eğitim verileri ve doğrulama verileri ve yukarıda tanımlanan geri arama ile uyum yöntemini kullanarak modeli eğitin
- Son 12 aylık test verileriyle başlatarak ve bir sonraki değeri yinelemeli olarak tahmin ederek sonraki 60 ay için tahminler oluşturun
- MinMaxScaler'ı kullanarak tahmin edilen sıcaklıkların ölçeğini kaldırın ve bunları Santigrat ölçeğine dönüştürün

- Matplotlib kullanarak orijinal ve tahmin edilen sıcaklıkları çizin ve x eksenini etiketlerini yılları gösterecek şekilde biçimlendirin.

ARIMA ALGORITMA

- İlk adımda gerekli kütüphaneleri yüklüyoruz. Bu adımda, kullanılacak fonksiyonları içeren kütüphaneler içeri aktarılır,
- Veri yükleme işlemi ikinci adımda gerçekleştiriliyor. Burada, kullanılacak veri seti yüklenir,
- Verinin index'ini tarih kolonuna ayarlanır. Bu işlem, verinin daha iyi işlenebilmesi için önemlidir.
- Sıcaklık değerlerini index'e göre aylık ortalama sıcaklık değerlerine dönüştürme işlemi dördüncü adımda gerçekleştirilir. Bu işlem, aylık bazda sıcaklık ortalamalarına ulaşılmasını sağlar.
- Verilerin normalizasyonu beşinci adımda gerçekleştirilir. Verilerin farklı ölçeklerde olması durumunda, bu verileri normalleştirerek farklı ölçeklerdeki verileri aynı ölçeğe indirgeyerek daha doğru sonuçlar elde edilir.
- ARIMA modeli için p, d, q değerleri belirlenir. Bu değerler, modelin doğru çalışması için önemlidir. Bu adımda, farklı p, d, q değerleri deneyerek en uygun değerleri belirlenir.
- ARIMA için AIC değerleri hesaplanır. AIC değeri, modelin kalitesini ölçmek için kullanılır. Bu adımda, farklı p, d, q değerleri için AIC değerlerini hesaplayarak en uygun değerler belirlenir.
- ARIMA modeli eğitilir. Bu adımda, en uygun p, d, q değerlerini kullanarak model eğitilir.
- Gelecek dönemler için oluşturulan model ile tahmin işlemi yapılır. Bu adımda, eğitilmiş modeli kullanarak gelecek dönemlerde sıcaklık değerleri tahmin edilir.

- Son olarak, sıcaklık verileri ile grafik oluřturma iřlemi gerekleřtirilir. Bu adımda, tahmin edilen sıcaklık deęerleri ile gerek sıcaklık deęerlerini karřılařtırmak iin bir grafik oluřturulur.

ÖZGEÇMİŞ

Ad-Soyad : Cemal ERDEM

ÖĞRENİM DURUMU:

- **Lisans** : 2021, İstanbul Aydın Üniversitesi, Mühendislik Fakültesi, Yazılım Mühendisliği
- **Yüksek Lisans** : 2023, İstanbul Aydın Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2023, Property Finder, Yazılım Mühendisi
- 2022, Roofstacks – GoArt, Yazılım Mühendisliği
- 2019, Protel Bilgisayar A.Ş, Yazılım Mühendisliği