



**REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY
UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING**

**COVID-19 PANDEMIC PREDICTION USING ARTIFICIAL
INTELLIGENCE TECHNIQUES IN TÜRKİYE**

**SERPİL KARA
MASTER OF SCIENCE**



REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY
UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING

COVID-19 PANDEMIC PREDICTION USING ARTIFICIAL
INTELLIGENCE TECHNIQUES IN TÜRKİYE

SERPİL KARA
MASTER OF SCIENCE

SUPERVISOR
ASSOC. PROF. DR. MUSTAFA GÖÇKEN

ADANA 2022

I hereby declare that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all information that is not original to this work.

[Signature]

Serpil KARA

ABSTRACT

COVID-19 PANDEMIC PREDICTION USING ARTIFICIAL INTELLIGENCE TECHNIQUES IN TÜRKİYE

Serpil KARA

Department of Industrial Engineering

Supervisor: Assoc. Prof. Dr. Mustafa GÖÇKEN

August 2022, 94 pages

The COVID-19 pandemic has been a serious threat to people's lives since the first case was seen. The rapid spread of the virus around the world and the fact that it creates a global problem reveals the importance of the authorities' making quick and effective decisions. In this respect, detecting the COVID-19 disease and estimating the increasing number of cases has become an important issue. In this study, the number of daily cases, daily deaths, cumulative cases, and cumulative deaths are predicted in TURKEY. Stacked-LSTM, Bi-LSTM, and ANN methods, which are among the artificial intelligence techniques, are used for estimating the number of COVID-19 cases. The hyperparameters of these methods are optimized using the Grey Wolf Algorithm (GWO). In the study conducted in Turkey, MAPE, MAE, and R^2 Score performance metrics were used to determine the accuracy between the actual data and the predicted data for COVID-19. The success of the models was evaluated by comparing the datasets used for the models separately. With this study, the model and data set that give the closest result to the actual data values were determined. Thus, the accuracy of the pandemic was evaluated for the models.

Keywords: Data Analysis, Prediction, Deep Learning, Artificial Neural Networks, LSTM

ÖZET

TÜRKİYE'DE YAPAY ZEKA TEKNİKLERİ KULLANILARAK COVID-19 PANDEMİ TAHMİNİ

Serpil KARA

Endüstri Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Mustafa GÖÇKEN

Ağustos 2022, 94 sayfa

COVID-19 pandemisi, ilk vakanın görüldüğü andan itibaren insanların yaşamları için ciddi bir tehdit oluşturmaktadır. Virüsün dünya çapında hızla yayılması ve küresel bir sorun oluşturması, yetkililerin hızlı ve etkili kararlar almalarının önemini ortaya koymaktadır. Bu açıdan COVID-19 hastalığının tespiti ve artan vaka sayılarının tahmin edilmesi önemli bir konu haline gelmektedir. Bu çalışmada Türkiye'de günlük vaka, günlük ölüm, kümülatif vaka ve kümülatif ölüm sayıları tahmin edilmektedir. Yapay zeka teknikleri arasında yer alan Stacked-LSTM, Bi-LSTM ve ANN yöntemleri COVID-19 vaka sayılarının tahmininde kullanılmaktadır. Bu yöntemlerin hiperparametreleri, Gri Kurt Algoritması (GWO) kullanılarak optimize edilmiştir. Türkiye'de yapılan çalışmada, COVID-19 için gerçek veriler ile tahmini veriler arasındaki doğruluk performansını belirlemek için MAPE, MAE ve R^2 Skoru gibi performans metrikleri kullanılmıştır. Modellerin başarısı, modeller için kullanılan veri setleri ayrı ayrı karşılaştırılarak değerlendirilmiştir. Bu çalışma ile gerçek veri değerlerine en yakın sonucu veren model ve veri seti belirlenmiştir. Böylece pandemi verileri ile yapılan tahmin modellerinin doğruluğu ve başarısı değerlendirilmiştir.

Anahtar Kelimeler: Veri Analizi, Tahminleme, Derin Öğrenme, Yapay Sinir Ağları, LSTM



To my family

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Assoc. Prof. Dr. Mustafa Göçken, who guided me throughout my studies, conveyed his knowledge and experience, patiently answered all my questions and helped me progress through my thesis process meticulously. I am grateful for being a disciplined and inspiring guide in my academic progress by increasing my motivation.

I would like to thank my teacher, Assoc. Prof. Dr. Ayşe Tuğba Dosdoğru, who patiently answered my questions with her advice, guidance, and extensive knowledge at every stage of my research.

I would like to thank my dear father, who always wanted me to study, and my mother, who patiently supported me in my thesis process while I was struggling with cancer, and I wish them mercy. I am glad to have you, my brother Serdar KARA, who always gave me the greatest support and listened to me in all my stress.

And I would like to thank my dear friends who listened to me patiently during this process, did not spare their help and knowledge, did not let me lose my motivation, and always encouraged me.

TABLE OF CONTENTS

TABLE OF CONTENTS	vii
LIST OF FIGURES	viii
LIST OF TABLES	ix
NOMENCLATURE	x
1. INTRODUCTION	1
2. LITERATURE REVIEW	8
3. MATERIALS AND METHODS	22
3.1 DATA ANALYSIS	22
3.1.1 Types of Data Analysis.....	24
3.2 DATA MINING	29
3.2.1 CLASSIFICATION AND REGRESSION.....	30
3.2.2 CLUSTERING	31
3.2.3 ASSOCIATION RULES	32
3.3 ARTIFICIAL INTELLIGENCE	32
3.3.1 ARTIFICIAL NEURAL NETWORK	33
3.3.1.1 ARTIFICIAL NEURAL NETWORK CLASSIFICATION	39
3.3.1.1.1 CLASSIFICATION ACCORDING TO LEARNING METHODS	39
3.3.1.1.2 CLASSIFICATION OF THE ANN STRUCTURE.....	40
3.3.1.1.3 CLASSIFICATION BY PURPOSE OF USE	41
3.5.3 RECURRENT NEURAL NETWORK (RNN)	42
3.5.4 LONG- SHORT TERM MEMORY (LSTM)	43
3.5.4.1 Stacked LSTM.....	47
3.5.4.2 Bidirectional LSTM (Bi-LSTM).....	47
3.5.4.3 Convolutional LSTM (Conv-LSTM)	48
3.5.4.4 Gated recurrent unit (GRU).....	48
3.6 GRAY WOLF OPTIMIZER ALGORITHM	49
3.7 ERROR MEASURES	53
4. RESULTS AND DISCUSSIONS	56
4.4.2 Comparative Analysis of Best Results for Datasets	77
5. CONCLUSIONS	83
REFERENCES	85

LIST OF FIGURES

Figure 1.1: A sample of Table of Weekly Cases in Turkey.....	3
Figure 1.2: Vaccination Rates of Provinces in Turkey	4
Figure 1.3: Number of Cases Per Week by Province	5
Figure 3.1: Data Scale Types	23
Figure 3.2: Disciplines Associated with Data Analysis	29
Figure 3.3: Artificial Intelligence and its sub-branches.....	33
Figure 3.4: General Structure of ANN.....	35
Figure 3.5: Structure of Neuron Network	36
Figure 3.6: Feed-Forward Neural Network Structure	40
Figure 3.7: Back-Forward Neural Network Structure	41
Figure 3.8: Structure of Recurrent Neural Network Cell.....	42
Figure 3.9: Recurrent Neural Network Structure	43
Figure 3.10: Structure of Long-Short Memory Cell	44
Figure 3.11: Structure of LSTM algorithm.....	45
Figure 3.12: Structure of Stacked-LSTM	47
Figure 3.13: The Bi-LSTM Structure	48
Figure 3.14: GWO Algorithm Social Hierarchy	50
Figure 3.15: Pseudo Code of the GWO Algorithm (Mirjalili, Mirjalili, & Lewis, 2014.)	52
Figure 4.1: Flow Chart of the Study	56
Figure 4.2: Predicted Values and Actual Values of COVID-19 Daily Cases.....	60
Figure 4.3: Predicted Values and Actual Values of COVID-19 Daily Deaths	60
Figure 4.4: Predicted Values and Actual Values of COVID-19 Cumulative Cases	61
Figure 4.5: Predicted Values and Actual Values of COVID-19 Cumulative Deaths	62
Figure 4.6: Predicted Values and Actual Values of COVID-19 Daily Cases.....	67
Figure 4.7: Predicted Values and Actual Values of COVID-19 Daily Deaths	67
Figure 4.8: Predicted Values and Actual Values of COVID-19 Cumulative Cases	68
Figure 4.9: Predicted Values and Actual Values of COVID-19 Cumulative Deaths	69
Figure 4.10: Predicted Values and Actual Values of COVID-19 Daily Cases.....	74
Figure 4.11: Predicted Values and Actual Values of COVID-19 Daily Deaths	75
Figure 4.12: Predicted Values and Actual Values of COVID-19 Cumulative Cases	76
Figure 4.13: Predicted Values and Actual Values of COVID-19 Cumulative Deaths	76

LIST OF TABLES

Table 2.1: Summary of Literature Related to AI.....	16
Table 2.2: General Summary Table of the Literature Review on the SIR Model and Other Methods	19
Table 4.1: Descriptive Statistics of COVID-19 Dataset.....	57
Table 4.2: Parameters of GWO	58
Table 4.3: Parameters of LSTM	59
Table 4.4: Optimized Stacked-LSTM Parameter Values	63
Table 4.5: Comparison of Performance Metrics of the Stacked-LSTM.....	64
Table 4.6: CPU Times (in seconds) for All Datasets	66
Table 4.7: Optimized Bi-LSTM Parameter Values	70
Table 4.8: Comparison of Performance Metrics of the Bi-LSTM for COVID-19.....	71
Table 4.9: CPU Times (in seconds) for All Datasets	73
Table 4.10: Optimized ANN Parameter Values	77
Table 4.11: Comparison of Performance Metrics of the ANN for COVID-19.....	78
Table 4.12: CPU Times (in seconds) for All Datasets	80
Table 4.13: Performance Metrics' Comparison of Models for Daily Cases	80
Table 4.14: Performance Metrics' Comparison of Models for Daily Deaths.....	81
Table 4.15: Performance Metrics' Comparison of Models for Cumulative Cases.....	81
Table 4.16: Performance Metrics' Comparison of Models for Cumulative Deaths	82

NOMENCLATURE

SARS	: Severe acute respiratory syndrome
MERS	: Middle East Respiratory Syndrome
WHO	: World Health Organization
R	: Reproduction number
AI	: Artificial Intelligence
ML	: Machine Learning
DL	: Deep Learning
EDA	: Exploratory Data Analysis
VEDA	: Visual Exploratory Data Analysis
SEIR	: Susceptible- Exposed-Contagious- Infected-Recoveres
SIQR	: Susceptible-Exposed-Quarantined-Recovered
SIRD	: Susceptible-Infected-Recovered-Death
SIR	: Susceptible-Infected-Recovered
SEIRD	: Susceptible-Exposed-Infected-Recovered-Death
ARIMA	: Autoregressive Integrated Moving Average
SI	: Serial Range
ANN	: Artifical Neural Networks
RNN	: Recurrent Neural Network
LSTM	: Long-Short Term Memory
BI-LSTM	: Bidirectional LSTM
CONV-LSTM	: Convolutional LSTM
BI-CONV-LSTM	: Bidirectional Convolutional LSTM
GRU	: Gated Recurrent Unit
BI-GRU	: Bidirectional Gated Recurrent Unit

MAE	: Mean Absolute Error
MAPE	: Mean Absolute Percentage Error
MSE	: Mean Squared Error
RMSE	: Root Mean Squared Error
EVS	: Explained Variance Score
MSLE	: Mean Squared Logarithmic Error
VAR	: Vector Autoregressive
CNN	: Convolutional Neural Networks
NARANN	: Nonlinear Autoregressive Artificial Neural Networks
MLP	: Multilayer Perception Neural Network
VAE	: Variable Automatic Encoder
CT	: Computerized Tomography
R^2	: Coefficient of Determination

1. INTRODUCTION

Considering the course of the epidemics from the past to the present, it affects many people in different geographies on various issues. As a result of epidemics, serious negative effects are seen in situations such as people's lives, health, psychology, the economy, agricultural activities, production, education, and life. The way people live is changing, causing their economies to deteriorate. Serious health problems and deaths as a result are the most important issues that individuals are affected by throughout history, many epidemics have occurred, and as a result, people have faced a serious number of deaths. These diseases, which have caused many deaths from the past to the present: Plague Epidemic, which is called the Black Death between 1347-1357, Hemorrhagic Fever that occurred between 1545-1548, Measles that occurred between 1592-1596, Smallpox that appeared in 1630, Yellow Fever that appeared in 1648, Influenza seen between 1775-1776, Typhus epidemic first seen in 1793, Cholera Epidemic that occurred in 1817-1824, Spanish Flu in 1918-1920, HIV/AIDS pandemic that started in the 1960s and still continues, SARS that severe acute respiratory syndrome seen in 2002-2003, Swine Influenza Virus pandemic (H1N1/SIV) that emerged in 2009-2010, MERS that emerged in 2012, Ebola virus epidemic in 2013-2016, Swine flu and Zika virus that emerged in 2015 and still continues to the present day ("List of Epidemics-Wikipedia" Retrieved December 7,2021.) Finally, many patients diagnosed with pneumonia, the cause of which is unknown, applied to the hospital in Wuhan, China in December 2019 ("WHO/Europe | Coronavirus Disease (COVID-19) Outbreak - About the Virus", Retrieved December 7, 2021.) The number of cases increased to a remarkable extent in a very short period of time and forced doctors to investigate the causes of the disease. As a result of examining these samples, Dr. Jianguo Xu, academician of the Chinese Academy of Engineering and leading the scientific team, announced at the national press conference that a new type of coronavirus has emerged, tentatively named 2019-novel coronavirus (2019-nCoV) by the World Health Organization (Lu, Stratton, & Tang, 2020.) This disease, which was later called the New Coronavirus, was named SARS-Cov-2 because of its similarity to the SARS Coronavirus (Bilgiler, Tani, Danışma, & Çalışması,2020.) This new type of betacoronavirus has emerged, unlike both SARS and MERS viruses, posing a serious health problem that can result in deaths to human health worldwide (Zhu et al.,2020.)

Coronaviruses are pathogens of great importance to humans and many vertebrates, such as bats, livestock, birds, and mice. Therefore, it can cause various disorders such as central nervous system, stomach, liver, intestine, respiratory system infections. The outbreaks of severe acute respiratory syndrome (SARS) in 2002-2003 and Middle East respiratory syndrome (MERS) in 2012 demonstrated the possibility of animal-to-human and human-to-human transmission of emerging diseases (Chen, Liu, & Guo,2020.) However, in connection with the fact that the new coronavirus first emerged from the seafood market, the possibility that the virus may have passed from animal to human was considered. Later, the increasing number of patients outside of this seafood market supported the possibility of human-to-human transmission of the virus (Summary,2020.)

Coronaviruses are single-stranded RNA viruses (Yin & Wunderink,2018), and there are four types. These are known as the Alpha, Beta, Gamma, and Delta coronaviruses. There are various types of coronaviruses, such as HCoV-NL63, HCoV-229E, HCoV-OC43, and HKU1-HCoV found in humans (Bilgiler, Tanı, Danışma ve Çalışması, 2020.) These can be transmitted from person to person and cause respiratory diseases such as coughs, sore throats and shortness of breath.

On January 30, 2020, the WHO classified the COVID-19 disease as an "international public (human) health emergency." The epidemic first started in China. As a result of the rapid spread of the virus with the outbreak of cases in 113 countries other than China, it was defined as a global epidemic, that is, a pandemic, on March 11. According to the data of Google News on January 10, 2022, for COVID-19, the new number of cases in the world is 3.188.145 and the new number of deaths is 8.348. The one dose of vaccine is 4.672.538.297 and the number of people who have completed all vaccines is 3.994.715.371 ("Koronavirüs (COVID-19)-Google News," Retrieved December 25,2021.)

In Turkey, the studies for the COVID-19 cases first started on January 10th. The Scientific Advisory Board of the Ministry of Health of the Republic of Turkey held its first meeting on January 22, and various measures were started to be taken. The first cases in Turkey was seen in March ("COVID-19," Retrieved December 25,2021.) A visual table has been prepared by the Turkish Ministry of Health to inform the public about COVID-19 and to make the COVID-19 data more understandable. In this table, information such as the number of daily

cases, the number of deaths per day, the number of recovered patients, the number of tests, the total number of cases, and the total number of deaths can be accessed (Figure 1). Considering the data published by the Ministry of Health, a total of 176 cases were seen between March 11-17, 2020, the week the virus was first seen in Turkey, and the number of deaths was announced to be 1. A sample visual table for the week of 04-10 December 2021 is given in figure 1. From figure 1 it is seen that a total of 2.467.566 tests were carried out within that time period. Also, the number of cases, the number of deaths, and the number of recoveries were 141.582, 1.361, and 175.993, respectively. From the first appearance of the virus in Turkey to December 10, 2021, the total number of cases was 9.002.968 and the total number of deaths is 78.778. Considering the vaccination studies, the number of 1st dose vaccines for the population over the age of 18 is 56.777.545, and the ratio of 1st dose vaccination across the country is 91.47%. When it comes to the second dose of vaccine, the total number is 51.365.908 and the rate of 2nd dose vaccine is 82.75% across the country. The total number of first, second and third dose vaccines is 127.994.715 (“COVID-19,” Retrieved December 25,2021.)

 TÜRKİYE HAFTALIK COVID-19 TABLOSU		TEST SAYISI	2.467.566
04 - 10 ARALIK 2021		VAKA SAYISI	141.582
Hastalarda Zatürre Oranı %	Yatak Doluluk Oranı %	HASTA SAYISI	7.921
4,0	57,7	VEFAT SAYISI	1.361
Erişkin Yoğun Bakım Doluluk Oranı %	Ventilatör Doluluk Oranı %	İYİLEŞEN SAYISI	175.993
71,3	30,9	ORTALAMA AĞIR HASTA SAYISI	1.067
		TOPLAM VAKA SAYISI	9.002.968
		TOPLAM VEFAT SAYISI	78.778

Figure 1.1: A sample of Table of Weekly Cases in Turkey

The weekly number of cases in Turkey table information can be obtained from the website of the Turkish Ministry of Health, as well as the vaccination rate of the provinces and the COVID-19 cases rate in the provinces on the map of Turkey. As seen in Figure 2, the red

areas show the cities with a vaccination rate of less than 55%. The lack of red zone on the map indicates that the vaccination of individuals is progressing in a positive direction. It is seen that provinces such as Şanlıurfa, Mardin, Diyarbakır are in orange color and the rate in these regions is between 55% and 65%, showing that vaccinations have not reached sufficient levels in these regions. The yellow-colored areas show the vaccination rate of individuals between 65% and 75%. The blue colored cities indicate that more than 75% of the individuals have been vaccinated. If we look at the overall picture, the vaccination coverage of individuals tends to positive.

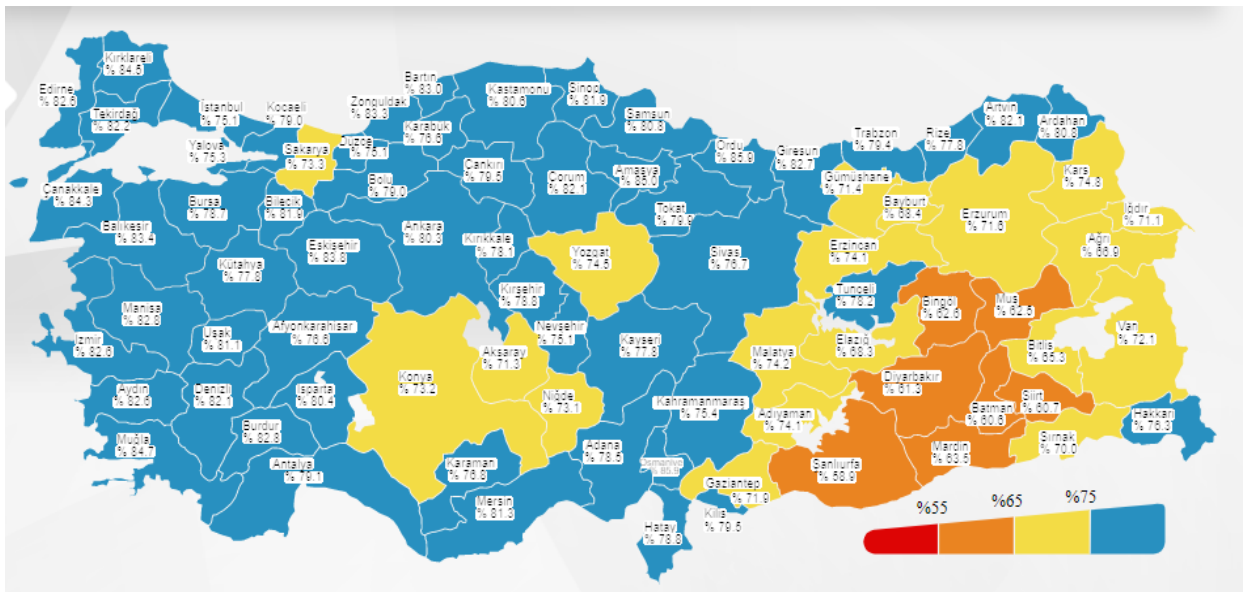


Figure 1.2: Vaccination Rates of Provinces in Turkey

The number of weekly cases by province is shown in Figure 3. The number on each province indicates the number of cases per 100,000 people in that province. The increase in the number of cases is shown from light color to dark color. While Van has the lowest number of cases with 3.92 per 100,000, Kırklareli has the highest number of cases with 467.74 per 100,000 as of 10.12.2021. Note that, all these values are updated weekly by the Turkish Ministry of Health.

that the 95% confidence interval for the mean of R_0 is [2.24-3.58] for COVID-19. This result supports that COVID-19 is a pandemic, since the value is greater than 1 (2020.)

Although many more studies related to epidemic have been published and many are still done, there is still a lack of information and gaps in knowledge on how to properly implement a particular policy. Therefore, these parameters alone do not allow us to obtain sufficient information about the status of the pandemic. The recent literature reveals that the researchers more heavily tend to utilize statistical analysis and mathematical models in their studies related to COVID-19. This analysis may be incomplete, as the studies for the course of the pandemic should be real-time studies. With AI techniques, more practical and useful results are produced in the solution of daily life problems. These techniques can provide effective answers for solving complex a difficult problem. At the same time, the relationship between AI and learning from data, evaluating the learned information, and making decisions from the information is strong. With these techniques, classification, estimation, clustering, generalization, and learning functions are frequently preferred. Looking at deep learning approaches; they can give effective results because it is more flexible, more realistic, easy to implement and economical. Therefore, in our study, it is desired to make predictions by using AI techniques and to determine, the performance of the models.

In this study, COVID-19 data for Turkey was predicted from the Turkovac website. Data groups were categorized as daily cases, daily deaths, cumulative cases, and cumulative deaths. In the study carried out, the data groups were predicted by LSTM, Bi-LSTM, and Stacked-LSTM methods, and model accuracies were compared. Although neural networks are successful in solving complex problems, they are difficult to train. However, more successful results are obtained when AI is trained with optimization algorithms. The Gray Wolf Optimization (GWO) algorithm, which is one of the meta-heuristic algorithms, was chosen to train the model in this study. The model in this studys are also optimized with the GWO, and the performance values are compared and the prediction method that gives the most accurate result is determined.

In the first part of this thesis, information about the subject was given. In the second part, literature research is conducted for the studies on the epidemic. In the third chapter, information about data analysis, artificial intelligence, machine learning, deep learning, the

GWO algorithm, and performance criteria are given, and the techniques used are explained. In the fourth chapter, the results obtained by using LSTM, Bi-LSTM, and Stacked-LSTM techniques using the GWO algorithm are compared and analyzed, and the accuracy of the models is determined by evaluating the performance metrics. In the fifth section, the results of the AI based prediction models are compared.



2. LITERATURE REVIEW

With the development of technology, a large number of data can be accessed quickly. Therefore, there is a need to obtain information from large amounts of data more easily and to produce faster solutions. Various analysis studies are carried out to classify data, produce and develop solutions. While performing data analysis, it is necessary to transform the data into appropriate forms in order to make the most of the data. In this respect, it is important to make accurate and meaningful analyzes in order to obtain meaningful information about the problem at hand (Oğuzlar, 2003.) From the past to the present human life affects the course of the pandemic as well as human life affects the pandemic. Pandemic data is important to determine the course of the pandemic, take precautions, and make decisions. The effects of the pandemic are investigated by establishing models with these data using various analysis techniques. At the same time, in order to obtain effective information from the pandemic data, different types of graphs were examined, including geographic graphs used for COVID-19 (Er & Bal,2020.)

Dey, Rahman, Siddigi & Howlader, (2020) used exploratory data analysis to visualize and analyze the number of confirmed, deaths and recovered cases using different datasets. Exploratory data analysis examines the characteristics of the data in a simpler and more understandable way by summarizing them visually. Khare et al. (2020) examine the status of the infectious disease for raising awareness of the people. Visual EDA has also been used to analyze and visualize spread time of COVID-19 worldwide. The SEIR model (susceptible-exposed-contagious-infected-recovered) was used to predict future cases numbers, deaths cases, and analyze the behavior of the disease. The result of this model was analyzed by performing a sensitivity analysis (2020.)

Amira et al. estimated the recovery and deaths rates of COVID-19 cases using the SEIR model, ignoring the confirmed cases. In addition, this study analyzes the news about the spread of the virus in the media. While selecting the news, care was taken to ensure that it should have at least more than 8 words and be in English. The negative and positive effects of this news on political, economic, social and psychological situations were evaluated by sensitivity analysis (2020.)

SEIR model was replaced by SIQR (susceptible-infected-quarantined-recovered) model to make predictions to understand the spread of COVID-19 in Brazil. SIQR model also considers the number of quarantined data (Crokidakis,2020.)

Fanelli & Piazza, (2020) used SIRD model to analyze the status of the epidemic in China, Italy and France. Spread size and peak time are mainly considered for determining the status of the epidemic. Also, the number of susceptible, infected, recovered, deaths regardless of exposed and quarantined cases are used as inputs to the SIRD model. The study mainly aimed at whether deaths, recovered and confirmed cases in these countries vary from country to country (Fanelli & Piazza,2020.)

In a study conducted in Pakistan, the SIR model was used to accurately predicted the epidemic data. By adding the exponential fit method to the SIR model, the number of infections and the peak of the epidemic are predicted (Waqas, Farooq, Ahmad, & Ahmad,2020.)

Casas, Carrasco, & Subirana, (2020) examined the causality of the model for Korea and Spain using the susceptible-exposed-infected-recovered-death (SEIRD) model. In the study, the suitability of the model is tested by targeting the “Is the model I have defined, correct?” and “Is my model coding correct?” questions. The changes in the parameters of system were examined, and the accuracy of the data sets was compared and evaluated (Casas, Carrasco, & Subirana,2020.)

Tandon, Ranjan, Chakraborty, & Suhag, (2020) predicted parameters using statistical analysis such as ARIMA. They developed an ARIMA model to evaluate the effects of the 60-day restriction on the epidemic in Italy. Daily confirmed and recovered cases are considered as the inputs to the model. The study justified the positive effects of containment and isolation on the spread of the pandemic. In the study, R_0 value of the growth curve of the epidemic is predicted. If $R_0 = 1$, it means that an average person who has this disease can infect another person with this disease. In the study, the R_0 value was confirmed to be between 2.0 and 2.5 (Zhao, Lin, et al.,2020.) Similarly, Aricigil Çilan & Kuzu, examined the serial range information of SARS and MERS viruses and predicted R_0 with the exponential growth curve. R_0 was found to be between 2.24 and 3.58 on average, and since this value was greater than

1, COVID-19 was considered as an epidemic (Arıcigil Çılan, & Kuzu,2013.) In another study, Gaeta (2020), using data from China, Korea, and Italy evaluates the reduction in growth rate. Tapiwa et al., (2020) developed a model that considers the rate of transmission to track the number of cases and the number of contacts. Moreover, serial range, effective reproduction number, and transmission values were examined, and the effects of social distancing measures on the course of the pandemic were evaluated.

SI assists in making strategic decisions for the health of the community. Zhao et al., (2020) tested various transmission scenarios for COVID-19 using Lognormal, Gamma and Weibull distributions while estimating the serial range (SI). Both Pearson correlation and R^2 used to measure the reliability of the serial interval.

AL-Rousan & AL-Najjar examined the relationship between dependent and independent variables using cross-tabulation. The effects of infection on variables such as deceased and recovered cases were analyzed using the chi-square test. While performing the X^2 test, multinomial logistic regression was used to check the accuracy and validity of each variable (AL-Rousan & AL-Najjar,2020.)

Association rules and the Apriori algorithm are also widely used in processing and transforming data into information (Uluyardımcı,2019.)

Irmak, Köksal, Asilkan, (2012) performed the density estimation analysis for estimating the number of patient density in the hospital for the next month by considering the number of patients admitted to the hospital. Three different analyses were performed while estimating the number of patient densities. Firstly, time series analysis with an exponential smoothing model, secondly an analysis using the ARIMA model and finally a prediction model with artificial neural networks were developed. These analyses were examined separately by developing different models within themselves, and then the method with the best model was determined comparatively. Similarly, Picariello & Aliani, (2020) predicted the total number of cases and deaths rates of the provinces of Bergama and Brescia, along with the Lombardy region. Canbay, İsmetoğlu, & Canbay, (2021) utilized artificial intelligence (AI) for detecting COVID-19 cases. Also, they emphasized the importance of patient data privacy for data analysis.

Mohamadou, Halidou, & Kapen, (2020) reported what kind of techniques can be used to obtain information about the general state of the epidemic, diagnosis and treatment of the disease. In the study, AI, mathematical modeling, data set analysis and the effects of social distance, restrictions, closures, and masks on the course of the pandemic were examined. In addition, whether factors such as temperature, age and population have an effect on the pandemic has been explained by comparing the analyses.

Giangreco, (2020) investigated the cases in Italy, South Korea, Germany, Spain and the United Kingdom on the hypothesis that the number of victims could be high due to contact between different age groups.

For the COVID-19 virus, AI and machine learning are the preferred algorithms for decision-making. Such application techniques can be evaluated on issues such as “early diagnosis of infection, estimation of the number of cases and deaths, development of drugs and vaccines, and control of the virus” (Vaishya, Javaid, Khan, & Haleem, 2020.)

Artificial neural networks (ANNs) are effective tools that are used to reveal the previously unknown hidden patterns in the data. In the study, 14 different prediction models are developed to evaluate the daily number of confirmed cases using ANNs. These prediction models were evaluated by examining the effects of the number confirmed COVID-19 cases per day on the forecasting time period for China, Japan, Singapore, Iran, Italy, South Africa, United States (Niazkar & Niazkar, 2020).

Al-Najjar & Al-Rousan, (2020) used ANNs for predicting the number of recovery and deaths cases in South Korea. Also, the effects of variables such as country, gender, cause of infection, year of birth, and region on recovered and fatal cases were examined. Lalmuanawma, Hussain, & Chhakchhuak, (2020) used AI and deep learning approaches measuring the accuracy and the reliability of epidemic predictions.

ANNs and deep learning algorithms are effective tools in modeling from small to large numbers of data. Eroğlu, (2020) proposed six different ANNs and deep learning approaches for predicting COVID-19 cases seven days in advance in Turkey. In the proposed models,

Long-Short term memory (LSTM) method is used in an attempt to increase prediction performances of both daily events and cumulative daily events by considering varying number of hidden layers (Eroğlu, 2020.) In addition, these techniques contribute to the prediction of the number of hospitalized patients, the arrival of patients to the hospital, and the management of decisions made for the hospital (Alimadadi et al.,2020.)

Kim, Kim, Lee, & Gim, (2021) predicted economic effects such as the inflation rate of one-year COVID-19 data considering COVID-19 historical data of weekly confirmed cases. Kim, Kim, Lee, & Gim, (2021) used Recurrent Neural Network (RNN) for examining the accuracy of the future predictions and the suitability of the utilized models.

Wang, Zheng, Ai, Liu & Zhu, (2020) reported that traditional epidemic analyses are sufficient for predicting the trend of the number of cases. They considered Iran, Russia, and Peru in their study and concluded that LSTM technique is more effective than the others in predicting when the spread of virus will decrease and cease. Similarly, in another study, three different cities were selected from three different continents and the recovered cases were predicted using the LSTM model. This study, conducted in the USA, Italy and India, predicted the next total recovered cases over 7 days. While making this estimation, mean absolute error (MAE), the mean squared error (MSE), the mean absolute percentage error (MAPE) and root mean squared error (RMSE) are considered as the performance metrics (Bahri, Kdayem, & Zoghlami,2020.)

Preprint & Sengupta, (2020) predicted how long it takes to reduce the current number of infected patients and the value of the number of infected patients at its peak in India. The statistical ARIMA model is used to analyze the data, the ML regression model is used to predict the future cases. Also, the SIR model is used to track the growth of infected individuals. In addition to the three models used in the study, the results are evaluated with the LSTM method.

Devaraj et al., 2021 predicted the number of COVID-19 cases using ARIMA, LSTM, Stacked-LSTM and Prophet methods. Cumulative number of deaths, confirmed and recovered cases are used as inputs to these models. Besides, they performed a comparative analysis by examining the reliability and applicability of these models. In addition to these models,

factors such as temperature, precipitation, region, population density are considered and the correlation analysis of these variables performed (Devaraj et al.,2021.)

Ayoobi et al., (2021) used LSTM, Bidirectional LSTM (Bi-LSTM), Convolutional LSTM (Conv-LSTM), Bidirectional Convolutional LSTM (Bi-Conv-LSTM), Gated Recurrent Unit (GRU) and Bidirectional gated recurrent unit (Bi-Gru) methods to predict one, three, and seven days ahead for the next 100 days for Austria and Iran. Various scenarios established and the number of deaths and new cases predicted. At the same time, these analyses are used in lower cost analyses.

Gomez-Cravioto, Diaz-Ramos, Cantu-Ortiz, & Ceballos, (2021) used LSTM and Vector autoregressive (VAR) models to analyze the status of the pandemic and to estimate the number of deaths and confirmed cases in Mexico and compared the results with other Latin American countries. As a result, the best model at predicting the number of COVID-19 cases for Mexico determined. Also, the relationship between climate and social mobility characteristics and the number of cases examined by correlation analysis.

Chimmula & Zhang, (2020) developed two, four, six, eight, ten, twelve, and fourteen-day forecast models. In the study, the transmission rate in Canada was compared with Italy and the USA and analyzed using LSTM.

Shahid, Zameer, & Muneeb, (2020) predicted confirmed, recovered, and deaths cases using ARIMA, support vector regression (SVR), LSTM and Bi-LSTM. The best model was determined by evaluating the performances of these models with respect to MAE, RMSE and R^2 Score (Shahid, Zameer, & Muneeb,2020.)

Istaiteh, Owais, Al-Madi, & Abu-Soud, (2020) used using ARIMA, LSTM, Artificial Neural Networks (ANN) and convolutional neural networks (CNN) for making a one-week forecast for the COVID-19 outbreak. In addition, the effective use of hyperparameters was examined and the compatibility and accuracy of the developed models were examined (Istaiteh, Owais, Al-Madi, & Abu-Soud, (2020.)) Guo and He, (1930) used only ANN in their study and predicted the confirmed cases and deceased cases considering R, MAE, RMSE performance measurements.

The system is dynamically modelled using stock flow models and temporal loop modeling techniques. The effect of social measures on the spread of the COVID-19 virus is examined, and the importance of social distance is mentioned (Niwa, Hara, Sengoku, & Kodama, 2020.)

In an Indian study, the region was examined in three groups based on the daily growth rate and the number of positive cases. The region is labeled as light zone if the number of cases is fewer than 200 and the daily increase in the number of cases is less than 2%. Secondly, if the number of positive cases is between 200 and 2000 and the increase in daily cases is less than 5%, then the region is labeled as middle region. Finally, if there are more than 2000 cases and the increase in daily cases is more than 5% then the region is labeled as severe region. The relationship between the number of confirmed cases were modeled with convolutional LSTM, Bi-directional LSTM, and Stacked-LSTM. Also, LSTM structures consisted of 100 hidden layers (Arora, Kumar, & Panigrahi, 2020). In their study, Alakus and Turkoglu (2020) created ANN, RNN, LTSM, and CNN models with 100 hidden layers to predict the number of individuals contracting COVID-19 disease. They also created two hybrid models i.e., CNNLSTM and CNNRNN, and considered precision, F1 score, recall, AUC, accuracy score to measure the performances of the models (2020.)

When the total confirmed, total recovered, and total deaths are examined, Marzouk, Elshaboury, Abdel-Latif, & Azab (2021), in their study in Saudi Arabia, evaluated the performance of ARIMA and NARANN (non-linear autoregressive artificial neural networks) models by comparing them with LSTM. Arun Kumar, Kalaga, Kumar, Kawaji, & Brenza (2021) compare RNN and LSTM models for ten different countries. Elsheikh et al. (2021), examine the COVID-19 outbreak in Egypt using the GRU and examine the performance of the GRU model by comparing LSTM, CNN, and MLP (Multilayer perception neural network) models.

Duan, Y., Yisheng, L. V., & Wang, F. Y (2016) construct an LSTM model by considering travel time and Fu, R., Zhang, Z., & Li, L. considering the traffic flow situation. Altché, F., & de La Fortelle, A., (2017), also predicts the behavior of vehicles using LSTM and GRU models.

Shahid, Zameer, & Muneeb analyzed the time series using Bi-LSTM while estimating. Liu, Y., Zheng, H., Feng, X., & Chen, Z., (2017) using the short-time traffic flow predicting LSTM and Conv-LSTM, they made an evaluation for the travel decision making and traffic jam elimination in studies. Also, the periodic behavior of traffic data is evaluated using Bi-LSTM. Cases prediction of the COVID-19 pandemic in India, using LSTM and Bi-LSTM techniques (Chandra, Jain, & Chauhan, 2022) by creating different scenarios, and Zeroual, Harrou, Dairi, & Sun using RNN, GRU and VAE (variable automatic encoder) in addition to these techniques, makes cases prediction and compares studies (2020.) Chandra, Jain, & Chauhan, 2022 created different scenarios and compared the performance of the analysis with LSTM and Bi-LSTM techniques. Zeroual, Harrou, Dairi, & Sun make cases predictions of the COVID-19 pandemic in India using RNN, GRU, and VAE (variable automatic encoder).

Table 2.1 and Table 2.2 provide an overview of the studies related to COVID-19 pandemic. Table 2.1 contains a summary of forecasting models using AI techniques. It is clearly seen that techniques such as ANNs, RNN, LSTM, CNN, GRU, and deep learning are frequently used as prediction models and for the diagnosis and the treatment of diseases. Table 2.2 consists of articles that examine the SIR model and its derivatives. Such models are used to prediction the number of cases, deaths, recoveries, and infections.

Table 2.1: Summary of Literature Related to AI

	PUBLISHING NAME AND NUMBER	AUTHOR NAME	DATE OF PUBLISHING	METHODS	OBJECTIVE OF THE METHOD	REFERENCES
1	A Classifier Prediction Model to Predict the Status of Coronavirus COVID-19 Patients In South Korea	H Al-Najjar, N Al-Rousan	March 2020	Artificial Neural Networks	It Is Used to Classify and Analyze Them Using Data on Recovering and Fatal Cases.	https://doi.org/10.26355/Eurev.202003.20709
2	Application of Artificial Neural Networks to Predict the COVID-19 Outbreak	Hamid Reza Niazkar & Majid Niazkar	November 2020	Artificial Neural Networks	To Predict Confirmed COVID-19 Cases, 14 Models Are Created and Estimated, and The Accuracy of the Model's Performance Is Examined.	https://ghrp.biomedcentral.com/articles/10.1186/s41256-020-00175-y#:~:text=Conclusions,Led%20to%20more%20accurate%20estimations.
3	Forecasting Models for COVID-19 Cases of Turkey Using Artificial Neural Networks and Deep Learning	Yunus Eroğlu	December 2020	Artificial Neural Networks Deep Learning (Long-Short-Term Memory, Lstm Networks)	By Using Ann And Lstm, The Number of Cases Is Estimated 7 Days in Advance and Analyzed. The Performances Were Evaluated by Comparing the Ann and LSTM Models.	https://doi.org/10.46465/Endustrimuhendisligi.771646
4	The Prediction of COVID-19 Using LSTM Algorithm	Myung Hwa Kim, Ju Hyung Kim, Kyoungjin Lee, Gwang-Yong Gim	November 2020	LSTM RNN Algorithm	Using These Techniques, The Spread of the COVID-19 Disease and The Effect of Economic Conditions on The Spread of The Virus Are Examined and Evaluated in Terms of Making Strategic Decisions.	https://doi.org/10.1016/j.jimu.2021.100566
5	Deep Learning for COVID-19 Prediction	Safa Bahri; Moetez Kdayem;	December	LSTM	Estimation of Recovered Cases Is Made	https://doi.org/10.110

		Nesrine Zoghلامي	2020	RNN Algorithm	in the USA, Italy and India Over 7 Days.	9/Ic_Aset49463.2020.9318297
6	Artificial Intelligence and Machine Learning to Fight COVID-19	Ahmad Alimadadi 1, Sachin Aryal 1, Ishan Manandhar 1, Patricia B Munroe 1 2, Bina Joe 1, Xi Cheng 1	April 2020	Artificial Intelligence Machine Learning	It Is Used for Diagnosing the Disease and Estimating the Number of Installed Patients.	https://doi.org/10.1152/Physiolgenomics.00029.2020
7	Artificial Intelligence (AI) Applications for COVID-19 Pandemic	Raju Vaishyaa, Mohd Javaid, Ibrahim Haleem Khan, Abid Haleem	August 2020	AI	Artificial Intelligence Applications Are Evaluated for The Detection and Treatment of COVID-19 Infections, The Transmission Status of The Disease, The Increase in The Number of Cases and Deaths, The Prevention of The Disease, And the Measures the Taken.	https://doi.org/10.1016/J.Dsx.2020.04.012
8	COVID-19 Hastalığının Teşhisinde Derin Öğrenme ve Veri Mahremiyeti	Yavuz Canbay, Abdullah İsmetoğlu, Pelin Canbay	June 2021	Artificial Intelligence Deep Neural Network Convolutional Neural Network- Cnn Recurrent Neural Network- RNN Long-Short Term Memory- LSTM	In the Literature, Artificial Intelligence Techniques Used in The Diagnosis of COVID-19 Disease, Data Sets, And Studies for Data Privacy of Patients Are Examined.	https://doi.org/10.21923/Jesd.870263
9	Prediction and Analysis of COVID-19 Positive Cases Using Deep Learning Models: A Descriptive Case Study of India	Parul Aroraa, Himans Hu Kumar Bijaya, Ketan Panigrahi	October 2020	Convolutional Lstm Bidirectional LSTM STACKED LSTM	Methods Are Used Separately for The Estimation of The Number of Positive Cases, And the Model with The Least Error Is Determined.	https://doi.org/10.1016/J.Chaos.2020.110017
10	Comparison of Deep Learning Approaches to Predict COVID-19 Infection	Talha Burak Alakus, Ibrahim Turkoglu	November 2020	Artificial Neural Network (ANN)	It Predicts Which Patients Are Likely to Contract COVID-19, Infection	https://doi.org/10.1016/J.Chaos.2020.110120

				Convolutional Neural Network (CNN)	Detection and The Performance of The Models Are Evaluated.	
				Recurrent Neural Network (RNN)		
				Long-Short Term Memory (LSTM)		
				CNNLSTM and CNNRNN		
11	Deep Learning-Based Forecasting Model for COVID-19 Outbreak in Saudi Arabia	Ammar H. Elsheikha, Amal I. Saba, Mohamed Abd Elaziz, Songfeng Lu, S. Shanmugan, T. Muthuramalingam, Ravinder Kumar, Ahmed O. Mosleh, F.A. Essa, Taher A. Shehabeldeen	May 2021	LSTM ARIMA NARANN (Nonlinear Autoregressive Artificial Neural Networks)	Prediction of Total Confirmed, Total Healed, And Total Deaths Are Made.	https://doi.org/10.1016/j.psep.2020.10.048
12	Deep Learning Model for Forecasting COVID-19 Outbreak in Egypt	Mohamed Marzouk, Nehal Elshaboury, Amr Abdel-Latif, Shimaa Azab	September 2021	Long Short-Term Memory Network (LSTM) Convolutional Neural Network, Multilayer Perceptron Neural Network	Methods Are Used to Learn the Extent of The Epidemic and To Estimate the Number of Cases For 7 And 30 Days.	https://doi.org/10.1016/j.psep.2021.07.034
13	Forecasting of COVID-19 Using Deep Layer Recurrent Neural Networks (RNNs) With Gated Recurrent Units (GRUs) And Long Short-Term Memory (LSTM) Cells	K.E. Arun Kumar, Dinesh V. Kalaga, Ch. Mohan Sai Kumar, Masahiro Kawaji, Timothy M Brenza	May 2021	RNN Algorithm Long Short-Term Memory Network (LSTM) GRU	The Importance of Various Factors, Such as Age, Gender, Preventive Measures, Health Facilities, And Population Density, Which Have a Significant Impact on The Spread of COVID-19, Is Examined, And the Future Situation Is Analyzed by Estimating the Number of Cases.	https://doi.org/10.1016/j.chaos.2021.110861

Table 2 2: General Summary Table of the Literature Review on the SIR Model and Other Methods

	PUBLISHING NAME AND NUMBER	AUTHOR NAME	DATE OF PUBLISHING	METHODS	OBJECTIVE OF THE METHOD	REFERENCES
1	Data Analysis for Covid-19	Noopur Khare	May 2020	Exploratory Data Analysis (EDA)	The Number of Confirmed Cases, Deaths, and Recovered Cases Was Used to Record, Analyze and Visualize Data.	https://www.researchgate.net/publication/n/342329585_Data_Analysis_for_COVID-19
		Meenakshi Jha		Visual Exploratory Data Analysis (VEDA)	To Control the Rate of Spread of COVID-19 and The Behavior of The Disease and The Rate of Spread to The World	
		Runjhun Mathur		Susceptible-Exposed-Infection-Recovered (SEIR) Model	It Has Been Used to Estimate the Time and Speed It Took to Spread the Virus Around the World, And to Estimate the Number of Cases and Mortality That May Occur In The Future.	
		Abhimanyu Kumar Jha		Sensitivity Analysis	Used for Keeping Track of Data and Statistical Analysis	
2	Modelling the Early Evolution of the COVID-19 In Brazil: Results from A Susceptible-Infections-Quartined-Recovered (SIQR) Model	Nuno Crokidakis	May 2020	Susceptible-Infectious-Quarantined-Recovered (SIQR) Model	By Analyzing The Data, the Infection Rate, The Rate Of Isolation of Individuals, The Number of Reproduction, The Doubling Time of The Epidemic Were Estimated, And Accordingly, The Change In The Number of Cases Over Time Was Examined	https://doi.org/10.1142/S0129183120501351
3	Corona Tracker: World-Wide COVID-19	Fairoza Amira Binti	March 2020	SEIR Model	The SEIR Model Was Used to Predict Covid-19 Cases, Deaths, And Recoveries.	http://dx.doi.org/10.2471/BLT.20.25569

	Outbreak Data Analysis and Prediction	Hamzaha, Cher Han Laub, Hafeez Nazric, Dominic Vincent Ligotd, Guanhua Lee, etc		Sensitivity Analysis	The Positive and Negative Effects of The News on People Are Examined.	5
4	Analysis and Prediction of COVID-19 Pandemic in Pakistan Using Time-Dependent SIR Model	Muhammad Waqas, Muhammad Farooq, Rashid Ahmad, And Ashfaq Ahmad	March 2020	SIR Models	The Technique Was Used to Generate the Data and Provide Future Predictions.	https://doi.org/10.48550/arXiv.2005.02353
				Experimental Harmony	This Technique Predicts a Reasonable Number Of Infections And Peaks of Pandemics.	
5	SEIRD COVID-19 Formal Characterization and Model Comparison Validation	Pau Fonseca i Casas, Victor Garcia i Carrasco and Joan Garcia i Subirana	JULY 2020	SEIRD Model	Modeling Is Done by Using Data from Korea and Spain, With High Infection Rate and Low Death Rate Parameters.	https://doi.org/10.3390/app10155162
				System Dynamic	Is My Model Coding Correct? And Is the Model I Described Correct? It Is Used to Predict the Causality of The Model by seeking Answer to Its Questions.	

6	Predicting the Progress Of COVID-19: The Case for Turkey	Mesut Ozdinc, Kerem Senel, Selcen Ozturkcan, Ahmet Akgül	APRIL 2020	SIR Model	The Transmission Coefficient of The Virus and The Maximum Number Of Infected People Are Estimated by Analyzing the Model for A 43-Day Data Set from The Beginning of the COVID-19 Epidemic In Turkey.	https://www.researchgate.net/publication/340897846_Predicting_the_Progress_of_COVID-19_The_Case_for_Turkey
7	Forecasting the Peak Of COVID-19 Daily Cases In India Using Time Series Analysis and Multivariate LSTM	Souvik Sengupta	AUGUST 2020	Statistical Arima Model	Statistical Estimation Is Made for Daily Cases.	https://easychair.org/publications/preprint/D9sl
				SIR Epidemical Model,	Case Count Estimate and Highest Case Count Is Estimated for India	
				L Regression Model	With the Regression Model, In Addition to The Number of Daily Cases, It Provides Information About the Effect of Features Such as Temperature, Humidity, And Population on The Number of Cases In Countries.	
				LSTM Model	LSTM Is Trained with The Other Three Techniques Used to Predict Future Case Numbers.	

3. MATERIALS AND METHODS

3.1 DATA ANALYSIS

Data is any number, word, photograph, piece of music, or other object that may be stored on a computer (Tüzüntürk, 2010.) Today, thanks to rapid technological advancements, a vast amount of data is collected. Data analysis is the act of obtaining usable information and results by processing and interpreting raw data using several disciplines, particularly statistical techniques. It is very important why the data analysis is done, how it should be done, and what the size of the data is. While analyzing the data, it is necessary to determine the purpose for which it is done, and as a result, the right analysis method must be clearly chosen. The method chosen for data analysis is based on the fact that historical data reflects future data. When this situation is examined, successful results are achieved when the past data is examined and managed according to the necessary conditions and predicted effectively. The purpose of data analysis is to determine the optimum point by helping to make effective decisions. Using this result data, appropriate and timely strategic decisions and efficient policies may be made, and these analyses evaluate the success of the developed models (Akpınar, 2014.)

The data analysis process steps can be listed as follows:

1. In the first step, we must determine the nature of the problem we want to analyze. In addition, it is necessary to clearly state what our aim is as a result of solving this problem.
2. The second step is to determine whether the data needs to be used for analysis. It should be determined whether the data is numerical data or a categorical data.

Categorical data: It is type of data that we cannot express with numbers. Groups such as gender, education, marital status, and occupation are included in this type of data group. Categorical data is divided into two as nominal scale and ordinal scale according to the scale type.

- Nominal scale: If the value assigned to a variable cannot be measured, such as largeness or smallness, or if it cannot be sorted, if there is only naming, it is the type of scale used in such cases.

- Ordinal scale: If a variable can be listed as large or small, if a variable is superior to another variable, it is the preferred scale type in such cases.

Numerical data: Data that we can measure with numbers such as age, height, weight, success score. It is divided into proportional scale and interval scale. We can convert some categorical data into numerical data, but it is difficult when analyzing. (“Kılınç”, Retrieved February 07, 2022.)

- Interval scale: It is a type of scale where values between numerical data can be measured.
- Ratio scale: It is a type of scale that has all the features of the interval scale and also has zero value (Akpınar, 2014.)

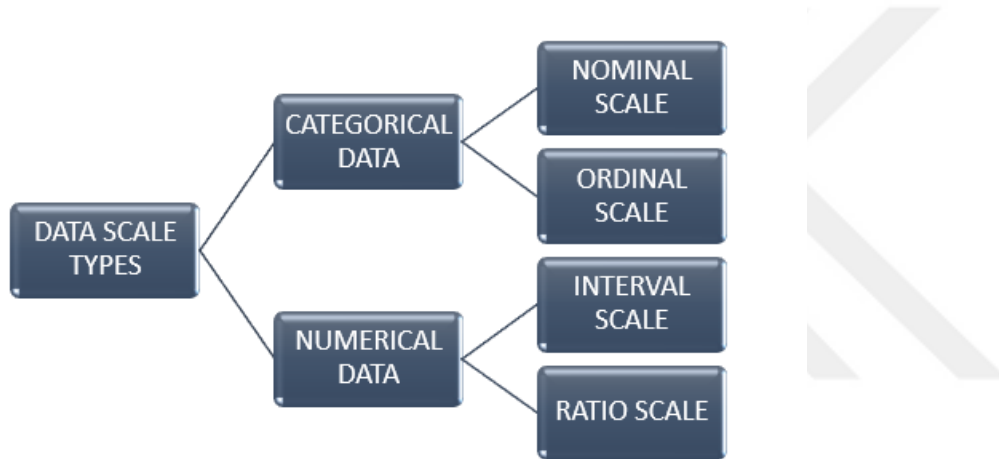


Figure 3.1: Data Scale Types

3. After deciding on the data, we want to analyze; the third step is to determine how to collect this data and from which sources it will be obtained.
4. The data used in the analysis must be regular data in order to obtain effective and efficient results. For this, in the fourth step, the data is organized by analyzing the irregular data with methods such as statistical analysis.
5. In the fifth step, cleaning is done so that the data can be understood. In this data cleaning step, missing, inaccurate and noisy data in the data are removed. This may require data integration, transformation, reduction.
6. When the data is ready for analysis, an appropriate model for the desired outcome is created.

7. Appropriate software to run the model is selected, and the accuracy and effectiveness of the model are evaluated.
8. Finally, the analysis is interpreted and it is decided how to proceed according to the results.
9. Analysis can be optimized by establishing optimization models (“Proente, Retrieved February 07,2022.)

3.1.1 Types of Data Analysis

Data is analyzed by experts, decision makers, companies to make strategic decisions, produce higher quality products and provide better service. When the data is analyzed in the appropriate form, it produces effective results and affects the improvement of performances. Analysis is preferred in order to examine customer behaviors, understand the reason for the decrease in sales, increase the quality of products, reduce costs, and improve processes. It is possible to examine data analysis types as eight headings.

1. Descriptive (Explanatory) Analysis: It is a very simple type of data analysis that can be easily understood by everyone. It is used to explain, summarize, visualize, analyze historical data.(“Veri Analizleri Nedir?”, Retrieved February 07,2022.) It aims to understand what happened by examining historical data. It defines the data, not the cause-effect relationship between the data, and is generally included in the content of all data. The best example of this type of data analysis is Google Analytics. It helps to obtain general information by defining situations such as how many people visit a page, how often they visit, the number of purchases made, how many people review an ad on average. It describes the data, not the relationship between the information in a group of data. In this way, graphs such as pie, bar, line can be used to explain the information (“What Are the Different Types of Data Analysis?” Retrieved February 07,2022.)
2. Exploratory Analysis: While analyzing the data, the state of the data is examined through a cause-and-effect relationship. It is used to understand the relationship between data variables directly or indirectly. Exploratory analysis can identify new relationships between variables and explore what caused the analysis as well as why. While using descriptive analysis to answer the question of what happened in data

analysis, the answer to the question of why did this happen is sought in exploratory analysis (“Types of Data Analysis”, Retrieved February 07,2022.) The obtained results are analyzed in more detail, the correlation between these variables is examined and meaningful results are obtained. For example, it has been observed that there is a decrease in the sales made in a store over the weekend, and the type of analysis performed to understand the reason for the decrease in these sales is called exploratory analysis. When a problem occurs, this type of analysis is used to search for similar results to this problem and to produce new solutions according to these similar models. In addition, the reason for this increase for an increasing product order at a certain time on a website can be investigated and the application of the same policy for similar products can be shown as an example (“What Is Data Analysis”, Retrieved February 07,2022.)

3. Inferential Analysis: Inferential analysis obtains information about future data by examining past data in addition to the previous two types of analysis. It is a type of analysis in which certain results are obtained by examining a small group when it is desired to make inferences about a large group (“By Benedict Neo”, Retrieved February 07,2022.) While performing this type of analysis, it is expected to make predictions by establishing statistical analyses and models. Different results can be obtained by choosing different samples from the same data group (“What Is Data Analysis”, Retrieved February 07,2022.) For example, it is desired to examine the education levels of 500 men and women in a city, and as a result, an idea about the general education level in this city can be obtained.
4. Forecast Analysis: Forecast analysis is used to make comments, make decisions, and predictive analysis for another group of data by using a certain amount of data stacks (“Veri Analizi Süreci ve Veri Analizi Yöntemleri Nedir?”, Retrieved February 07,2022.) It is a type of analysis made to make predictions about the future situation by considering past and current data. In this type of analysis, the accuracy of the predictions depends on the model and the input variables (“By Benedict Neo”, Retrieved February 07,2022.) Although the analysis for the predicted situation is not entirely correct, it is the type of analysis we are looking for an answer to the question of "what is possible" (“What is Data Analysis Types, Methods and Techniques?”,

Retrieved February 07,2022.) For example, consider a stationery store. In the analysis made by looking at the available data for this store, it is predicted that there will be a decrease in sales in the summer months. A decision can be made to develop a campaign in order to minimize this decrease in the sales rate in the summer season. Or take the example of a hospital where it records the number of patients who come to a hospital emergency department. By examining the data on the number of patients in this hospital, the frequency of patients for the weekend should be predicted. According to the result, logical decisions can be made and preparations can be made. With this analysis, data losses, inefficiencies and problems can be revealed, as well as gaining competitive advantage for companies, providing process optimization and revealing the potential of situations (“What Is Data Analysis Methods, Techniques, Types & How-To?”, Retrieved February 07,2022.)

5. Prescriptive analysis: It is the type of analysis preferred when predictive and descriptive analysis are not sufficient. It combines the information obtained from all other analyses when making a decision in an existing situation or solving a problem (“What Is Data Analysis?”, Retrieved February 07,2022.) With the prescriptive analysis, which is one of the most effective types of analysis, the answer to the question "how will it be" and "what will be done" is determined (“What Is Data Analysis? Methods, Techniques, Types & How-To”, Retrieved February 07,2022.) Analysis is done by considering questions such as; "What needs to be done to be able to solve a future problem and determine the appropriate steps?", "How to take advantage of the predicted situation in order to make the most appropriate decision?" (“What Are the Different Types of Data Analysis?”, Retrieved February 07,2022.) In some cases a single analysis type is not sufficient. In such cases, it is desirable to benefit from more than one type of analysis. In such cases, prescriptive analysis is used. For example, it is desired to reach another region from one region. While determining the most suitable route between these two regions; It is analyzed by considering factors such as traffic density, road condition, and multiple criteria are evaluated for this. In other words, prescriptive analysis plays an important role in determining what will be the best decision for the predicted situation in the future, and why it may cause it.

6. **Text Data Analysis:** Text data analysis is a sort of text mining that examines particular text, audio data. It is the simplest and most direct type of analysis. It's the study and analysis of data from text-based sources including social media messages, articles, and survey replies (“Veri Analizi Nedir? - Etkili Veri Analizi Teknikleri”, Retrieved February 07,2022.) It extracts useful and logical information from raw text data. Data mining, database, business intelligence, artificial intelligence tools are used to interpret the data and it is aimed at finding relationships between the data. While performing text analysis, keywords can be determined and frequency of use can be calculated, and people's feelings and thoughts can be analyzed (“What Is Data Analysis?”, Retrieved February 07,2022.)
7. **Mechanistic Analysis:** It aims to determine the results of the differences caused by another variable in the variables. It is a preferred type of analysis for measuring situations that require less error and high precision, such as engineering, physical sciences, behavioral and biological fields (“By Benedict Neo”, Retrieved February 07,2022.) It is used when a rigorous study is required to obtain accurate, high quality and clear experimental analyzes in scientific research.
8. **Casual analysis:** Correlation relationship between the data reveals the behavior of the relationships, the reason, the possible effect of the relationship, the reason for the result. Since the correlation cannot obtain a fully experimental result, it is desirable to investigate the cause of the results obtained using the causal analysis. It is the type of analysis used in cases where it is desired to obtain such more concrete results. e.g, A pharmaceutical company has developed a new drug and is being asked to test this drug. It is to compare the results with the experimental measurements and examine the possible effects of the drug (“By Benedict Neo” Retrieved February 07,2022). It is the type of analysis used not in the cases of competition between two or more models, but in situations where what causes which situation (Peyrot, 1996.)

It should be noted that all of these data analysis types are interconnected. Each analysis is used for a different purpose. It is important to choose the right method for the correct analysis and to benefit efficiently from the data. It is also possible for us to use more than one analysis

type together, when necessary, when we want to analyze a subject. While analyzing the data, knowledge and expert opinion are needed, whether in simple analyses or in complex analyses. Data analysis achieves important result in future predictions, behavioral analysis, strategic decisions taken for customer satisfaction, diagnosis and treatment of diseases. By using the relevant techniques, knowledge can be created such as making meaningful and efficient results from these large data groups (“Veri Analizi Nedir? Nasıl Yapılır? -By Can Cankıran,”, Retrieved February 08,2022.)

Although there are many medical studies on the COVID-19 pandemic, which is a serious struggle today, individuals want to learn the answers to many questions such as where the virus spread, how it spread, what is the rate of spread, and what is the number of cases. To assess and evaluate the situation, distribution, and behavior of the COVID-19 epidemic, detailed analyses are required. These analyses have a great impact on determining the course of the pandemic, the policies to be implemented, and the measures to be taken.

To inform people about the disease, analyzes can be made with models that compare data from different countries (Khare et al.,2020.)

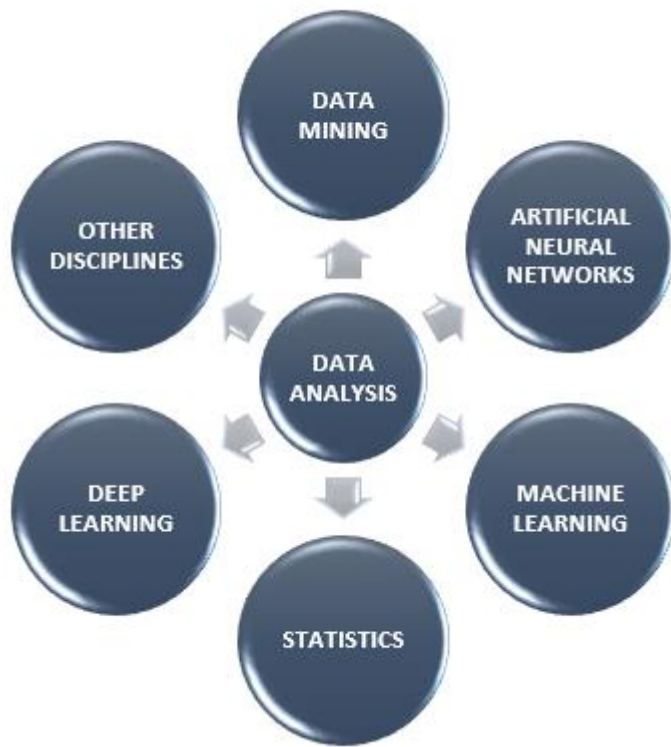


Figure 3.2: Disciplines Associated with Data Analysis

Data analysis is related to the fields such as data mining, artificial intelligence, and deep learning, as shown in Figure 3.2.

3.2 DATA MINING

Data mining refers to tools that aid decision-making by allowing users to swiftly access data and extract meaningful and usable information from massive amounts of ambiguous data (Baykal, 2006.) As a result of recording the data that has been going on for years, complex data stacks are formed with the accumulation of a large amount of data. Nowadays, with traditional data analysis approaches, it is impossible to extract relevant and effective information from these massive data stacks. As a matter of fact, in order to overcome this difficulty, data mining techniques are widely used in various fields such as insurance, marketing, astronomy, health, textile, trade, risk analysis, government policies (Clifton, C. 2022.) Data mining is a decision-making process that helps to determine future data predictions by analyzing historical data by finding the connections and relationships between various analyzes and data (Tüzüntürk, 2010.)

Models used in data mining are evaluated in two different ways as predictive and descriptive approaches (Özekes, S. 2003.)

Predictive models; This type of model is used for datasets with specific results. The developed models are predictive models. Various prediction approaches are made for future data, usually using historical data (Çelik M., 2009.) For example, these are models that are established to predict the number of patients who will arrive as a result of the examination of this data by recording the data on the number of patients coming to a hospital.

Descriptive models; Descriptive models describe and explain the relationships between hidden and complex data in data groups. Models that examine whether the purchasing relationship is similar between families with children who have more than one house with an income in the A-B range and those with a low income in the A-B range without children are examples (Özekes, S. 2003.)

Data mining models are examined in three different ways: classification and regression, clustering and association rules, according to the function they are used in. Classification and regression models are classified as predictive models, clustering and association rules are classified as descriptive models (Akpınar, 2000.)

3.2.1 CLASSIFICATION AND REGRESSION

It is the definition and classification of data according to some characteristics and qualities of the data in the data sets. The aim is to establish models where previously unknown, unlabeled data are correctly labeled and placed in the right class. Classification models can often be used in prediction models, and class labels can be given in advance (Tüzüntürk, 2010.)

The difference between classification and regression is the type of dependent variable. The predicted dependent variable in classification has a categorical value, while in regression it has a continuous value (Ayık, Özdemir, & Yavuz, 2007.)

Methods like multinomial logistic regression, which may also estimate categorical values, use classification and regression models in close proximity to one another. Almost the same techniques may be applied to classification and regression models (Akpınar, 2000.)

Classification is a type of model that is generally applied to many models to solve various problems in data mining. Some of the techniques used in classification are as follows:

- Decision Trees,
- Discriminant Analysis
- Rule Making Methods
- Neural Networks
- K Nearest Neighborhood
- Case-by-case reasoning
- Genetic Algorithms
- Naive Bayes

3.2.2 CLUSTERING

Clustering is a discovery process used in data mining to create strong similarities in data, to create this similarity in the best way and to minimize the difference between data sets (Karypis Eui-Hong Han, Vipin Kumar, 1999.) The clustering method can also be considered as the classification approach. It is known to which class the data examined in the classification belong (Özekes, S. 2003.) However, in clustering methods, there is no predetermined class formation stage as in classification. In the cluster method, a relationship between dependent and independent variables is not examined. In clustering methods, clustering function is performed in groups according to whether the features between the data are similar or not (Akpınar, 2014.)

There are many clustering algorithms in the literature (Özekes, S. 2003.) The main ones are;

- Partitioning methods
- Hierarchical methods (Hierarchical methods)
- Density-based methods
- Grid-based methods
- Model-based methods

3.2.3 ASSOCIATION RULES

Association rules establish association relations between datasets in order to reveal association relations between analyzed and stored datasets and to benefit from analysis and decision-making processes (Özekes, S. 2003.) One of the most basic usage examples of association rules is market basket analysis. This method analyzes which products they buy together when customers buy products, and as a result, it helps managers make decisions by creating efficient sales strategies. Association rules are used in the analysis of relationships that occur simultaneously and are interdependent, such as the probability of a customer buying yogurt to buy eggs. On the other hand, it is used to evaluate the relationship between the situations that are related to each other and the situations that occur at successive time (Akpınar, 2014.) Examples of the relationships between conditions such as the probability of developing disease B in a patient who has undergone surgery A, the possibility of purchasing paper in the first two months of a customer who buys a printer, and the probability of purchasing paper in the next two months can be given as examples.

3.3 ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is defined as “a field of science and engineering for the numerical understanding of what is commonly called intelligent behavior and helping to create artifacts that artificially exhibit such behavior (Ramesh, Kambhampati, Monson, & Drew, 2004.) In other words, it is called a program that can intelligently show people's behavior. For this purpose, the 'Turing Test' developed by Alan Turing started to be applied to determine the performance and behavior of people in an intelligent way. With the popular application of this development by machines and computers, the concept of AI continues to develop rapidly (Ramesh et al.,2004.) AI is a branch of science that integrates computer science by examining the function and way of working of the human brain, human experiences, and how people obtain information (Bayır F., 2006.) AI has the ability to gather the data needed to solve complex problems and provide solutions quickly (Uğur, A., & Kınacı, A. C.,2006.) It can be used effectively in many fields such as modern medicine, energy, economy, production and banking.

AI can be applied to many types of problems, such as making predictions, classifying data, identifying relationships between data, image recognition, disease diagnosis and treatment,

data interpretation, and analysis. AI technologies such as expert systems, machine learning, deep learning, ANNs, genetic algorithms, and fuzzy logic are used to solve such problems.

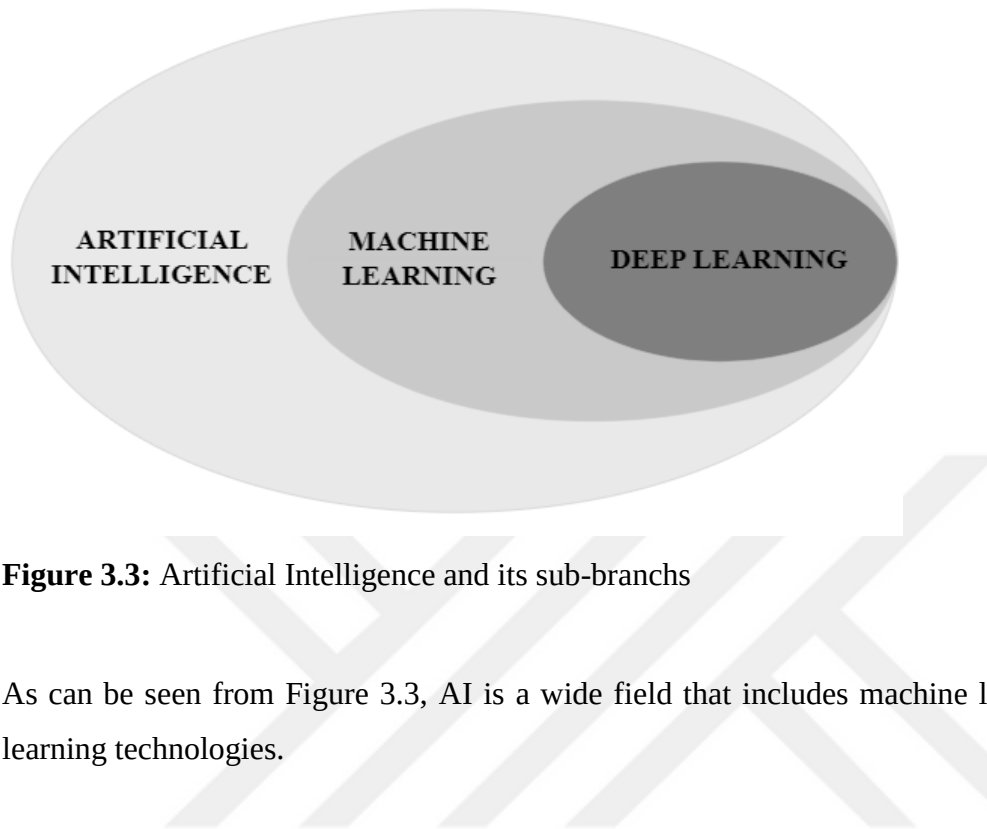


Figure 3.3: Artificial Intelligence and its sub-branches

As can be seen from Figure 3.3, AI is a wide field that includes machine learning and deep learning technologies.

3.3.1 ARTIFICIAL NEURAL NETWORK

With the rapid development of technology, the use of machines in our lives is increasing. Due to the increasing need for this use, new methods and new processes are being developed. The machine's ability to think like a human was first suggested in the 1980s. Later, in the 1990s, ANN methods progressed very rapidly (Yadav N., Yadav A., & Kumar M., 2015.)

Traditional programming methods may not be sufficient for solving faulty, missing, noisy, complex, ambiguous, multidimensional problems in ANN. ANN algorithms and models could be used to solve such kind of problems. As a result, ANNs became widespread quickly, as reliable information could be obtained, effective relationships can be established, and nonlinear complex behaviors can be detected (Yangin A.,2017.) Although ANNs are widely used for solving problems in the field of engineering, they also clarify various problems in the field of medicine. ANN is used effectively in all fields of medicine such as diagnosis and determination of diagnoses, radiology, pathology, CT (Computerized tomography), MR. (Ramesh et al., 2004). It aims to transfer human-specific features such as understanding, ideation, decision making, reasoning to machines with ANN methods (“Yapay Sinir Ağları ve

Uygulamaları | GTech,” Retrieved March 3,2022.) The biological nervous system is considered while achieving this goal. Neural networks in the biological nervous system consist of biological nerve cells. As in the biological neural network, there are artificial nerve cells in the ANN.

As a result, the ANN models based on the ways in which ANNs work were created by observing the human brain with nerve cells called neurons, as an example. ANNs transmit information to each other, receive information from each other, in short, they work in relation to each other. (Zontul, M., & Yangın A., 2017.) And by giving weights to the elements in the model, they can perform a study in connection with each other (Uğur, A., & Kınacı, A. C. 2006.)

As seen in Figure 3.4, after the inputs are determined, the network is trained to make decisions and generalize. As a result, the output is produced (Furkan Keskenler & Fahri Keskenler, 2017.) Its general structure consists of input, hidden and output layers. Each layer consists of many synaptic neural networks in itself (Bahri, Kdayem, & Zoghlami, 2020.) The input layer is the layer that consists of the inputs of the data, that is, the variables used for analysis. The output layer is the layer where the results and outputs are located. The hidden layer is the layer where the connections are established to transform the outputs using the input data (Niazkar & Niazkar, 2020.) Models can be developed not only using a single hidden layer but also using multiple hidden layers. The training process of ANNs is carried out by applying the optimization method. The synaptic weights are updated by considering the back-propagation approach (Elsheikh, 2021.)

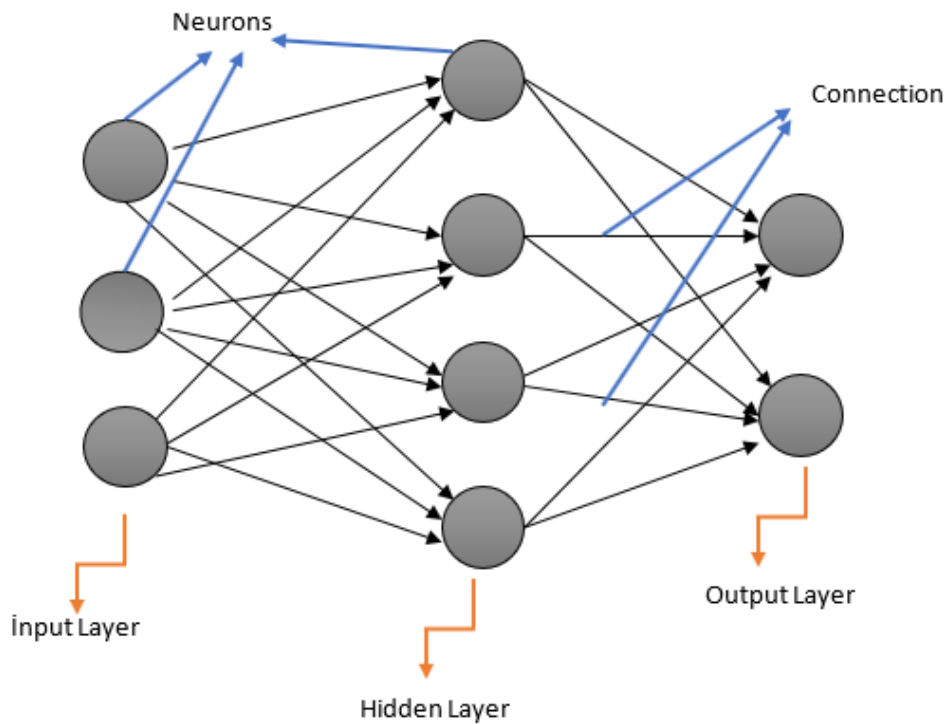


Figure 3.4: General Structure of ANN

Neurons form ANNs by connecting to other neurons with weights. The data coming from the outside is held according to the effect of the weights and connected to the neurons, and analyses are made from the data in the neurons for learning to take place. These weights determine the effect (i.e., contribution) of the input data. The aggregation function calculates the net value of the input data. The net input value is found by multiplying the input value with the weights. The activation function calculates what the net value of the output data is.

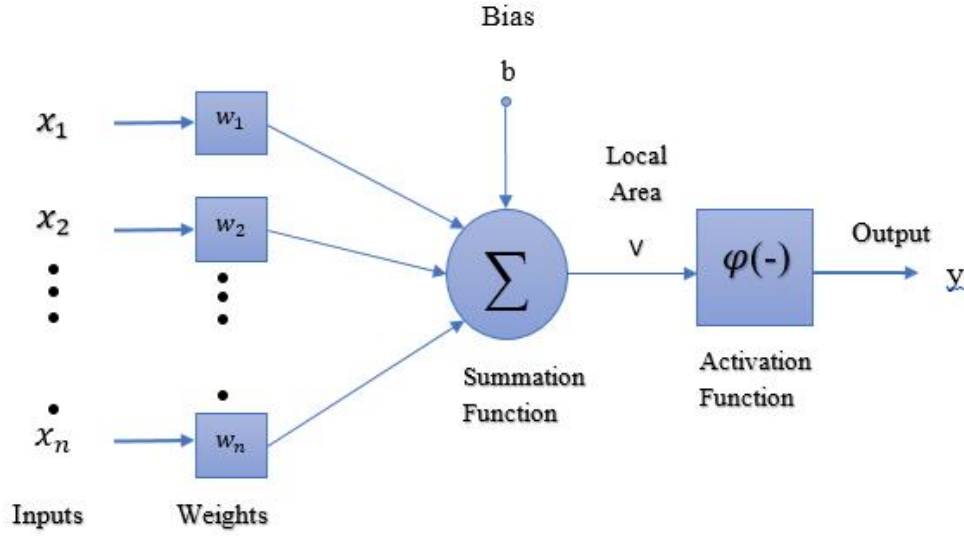


Figure 3.5: Structure of Neuron Network

As seen in Figure 3.5, an ANN cell consists of inputs, weights, summation function and activation function and outputs.

Inputs: information or data coming from outside of a neural network. A neural network can have an unlimited number of inputs.

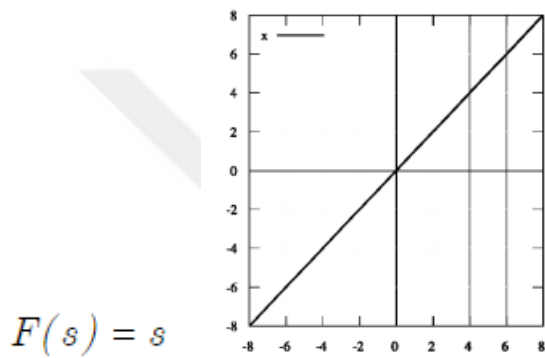
Weights: Affect the binding of incoming input values to neurons. It shows the importance of incoming input data and its effect on the nerve cell. Input data is connected to neurons according to the effect of weights (Yangın A. ,2017.) It does not specify whether the weights are important or unimportant, depending on whether they are large or small. A weight of zero may mean that it is much more important to that network (“Yapay Sinir Ağları - Derin Öğrenme | Deep Learning”, Retrieved March 3,2022.) The initial value of the weights is usually between -1 and +1 randomly (Yangın, A. ,2017.)

The summation function: It calculates the net value of the input data. The net input value is found by multiplying the input value with the weights (“Yapay Sinir Ağları - Derin Öğrenme | Deep Learning”, Retrieved March 3,2022.)

Activation function: The activation function determines the output data by processing the input data coming to the nerve cell. The activation function calculates what the net value of

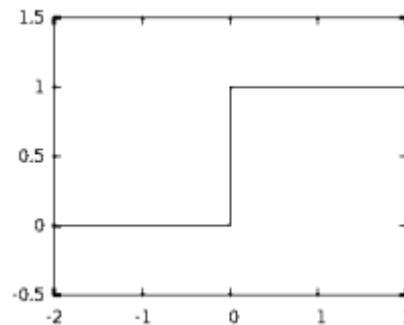
the output data is. When choosing the activation function, a nonlinear function is chosen, and for some models, it is expected that the derivative of the selected function will be easy to calculate (Yangın, A. ,2017.) Since the activation function affects performance, it is important to choose it correctly. While selecting the activation function, it is determined according to what is desired to be learned from the input data. Accordingly, the types of activation functions are as follows: (Bayır F., 2006.)

Linear activation function: it is the preferred function for solving linear problems. The result of the addition function is multiplied by a certain coefficient and the output is found.



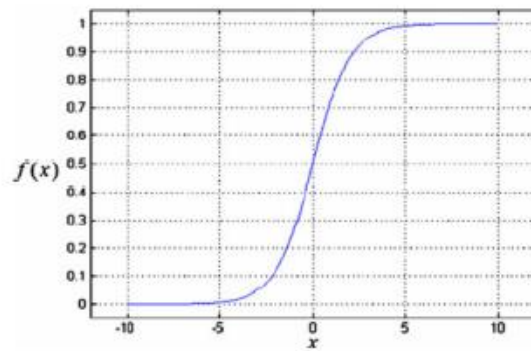
Step (step function): The cell output takes the value of 1 or 0 according to the value of the incoming net input data, below or above a specified threshold value.

$$y = F(s) = \begin{cases} 1 & s > 0 \\ -1 & s \leq 0 \end{cases}$$



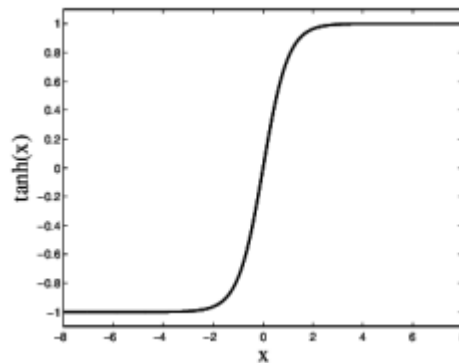
Sigmoid activation function: This function is continuous and can be differentiated. Since it is a nonlinear function, it is a very preferred activation function in ANN models. The sigmoid function finds a value between 0 and 1 for each value of the input data.

$$y = F(s) = \frac{1}{1 + e^{-s}}$$



Tangent hyperbolic activation function: The sigmoid function is similar to this function. In the tangent hyperbolic function, each output value created for the input data takes a value between -1 and +1.

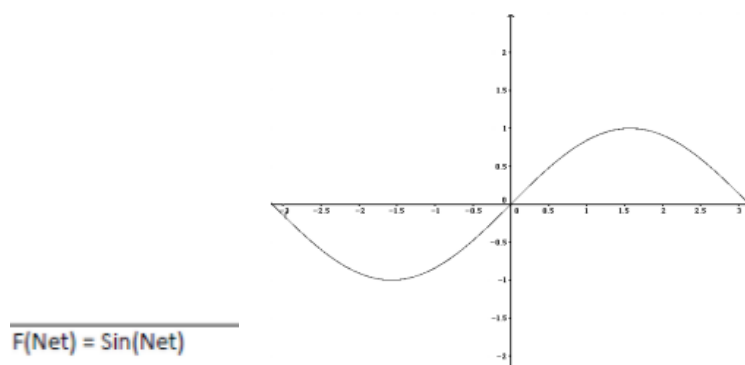
$$y = F(s) = \frac{e^s + e^{-s}}{e^s - e^{-s}}$$



Threshold function: The threshold function is the function that outputs 0 when the input information is less than or equal to zero, takes its own value when it is between zero and one, and returns 1 if the input data is greater than or equal to 1.

$$y = F(s) = \begin{cases} 1 & s > \theta \\ 0 & s \leq \theta \end{cases}$$

Sinus activation function: It is the function used when the information to be learned is distributed in accordance with the sinus function.



Output: Although a neural network has unlimited inputs, it only has one output.

ANNs have models that are self-learning even without a programmer. (Zontul, M., & Yangın, A. 2017).

3.3.1.1 ARTIFICIAL NEURAL NETWORK CLASSIFICATION

ANN models have been classified according to various purposes by scientists from past to present.

3.3.1.1.1 CLASSIFICATION ACCORDING TO LEARNING METHODS

In general, it is examined in three ways as supervised, unsupervised and reinforcement learning.

Supervised learning: The data used in supervised learning are historical data, which is a real system data (Öztanır O., 2018.) In this type of learning, the network weights are initially calculated randomly, then adjusted by the neural network, and a match is made by comparing the desired output value with the actual output in the network loop (Anderson & Mcneill, 1992.) Changes, mistakes, and inconsistencies in the neural network are determined according to the learning process (Sydenham & Thorn, 2005.) Neural networks need to be trained before they can be used. Input and output information sets are used as training sets. A supervised learning method is generally preferred for use in feed-forward networks. In this learning method, the expected output value for the input variable is known (Tolon & Güneri Tosunoğlu, 2008.)

Unsupervised learning: In unsupervised learning, only the input variables are known. The output data model does not need to be known. The neural network produces an output by training the output itself according to the input data. Therefore, external data works without assistance. The purpose of this learning is to reveal the model in the dataset in the neural network (Tolon & Güneri Tosunoğlu, 2008.) In this learning method, no external influence is needed to calculate the weights. The weights are adjusted in accordance with the neural network's performance. It is a less common type of learning and is often used in self-organizing neural networks (Anderson & Mcneill, 1992.)

Reinforcement learning: In this learning method, historical data is not used, instead the analysis of the data is done by trial and error. A result based on the data in the real system is obtained and evaluated when this result matches with the expected value (Öztanır, O., 2018.) It explores determining the most rewarding state by deciding what to do with the information or how to match input and output data (Sydenham & Thorn, 2005.)

3.3.1.1.2 CLASSIFICATION OF THE ANN STRUCTURE

The ANN structure is divided into two categories: feed-forward and feed-back, considering the connection of neurons to each other.

Feed-forward neural networks: In this type of network structure, there is a forward connection from one layer to the other, but there is no connection between neurons in the same layer. In such neural networks, the information in the input layer is transmitted to the next layer without any calculation (Yangin, A. 2017.) Output data from each layer is sent to the next layer as input with weights. Finally, the information is processed in the output layer and the targeted output is obtained. The number of neurons in the input and output layers depends on the problem to be solved. The number of hidden layers and the number of neurons in each layer can be determined according to the problem or can be found by trial and error (Ataseven, B.,2013.)

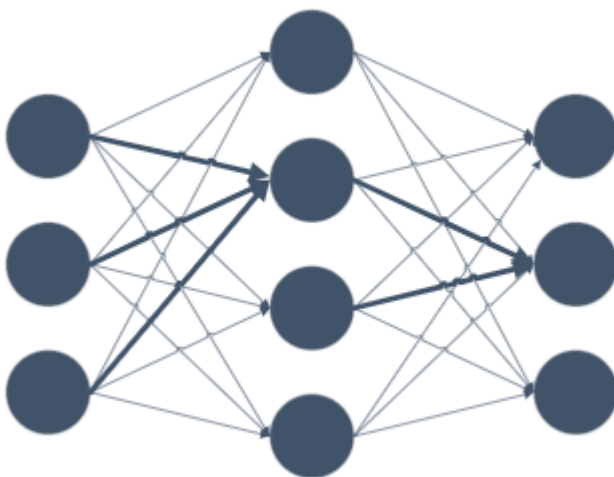


Figure 3.6: Feed-Forward Neural Network Structure

Neural networks with feed-back: There are a back synaptic link from the output layer to the input layer (Sazlı M., 2006.) In contrast to the feed-forward network structure, the input data in these networks is not only transmitted to the next layer. In this type of network structure, learning takes place by comparing each input with the target output data, and by calculating the error, it runs the error function backwards in the neural network model. This network model can be easily used for input and output. However, this type of network learns slowly and requires a lot of data for learning to take place (Kutlu, B., & Badur, B. 2009.)



Figure 3.7: Back-Forward Neural Network Structure

3.3.1.1.3 CLASSIFICATION BY PURPOSE OF USE

Since ANNs are generally organized into layers, we can classify these models as single-layer and multi-layer ANNs (Sazli, M. H., 2006.)

Single-layer neural network: This type of neural network consists only of input and output layers. ANNs consisting of a single layer can predict linear functions (Ataseven, B., 2013.) In single-layer neural networks, the output values take the values of 1 and -1. If the output is 1, it is accepted to the first class. If it is -1, it is accepted to the second class (Öztürk, K., & Şahin, M. E. ,2018.)

Multilayer neural networks: These are neural network models with at least one hidden layer. In these multi-layer neural network models, the hidden layer complex problems are preferred. At the same time, the fact that it can be used with this model in many learning networks

causes this model to be preferred (Ataseven, B., 2013) and classifies the data according to its properties (Doğan, F., & Türkoğlu, İ.,2019.)

3.5.3 RECURRENT NEURAL NETWORK (RNN)

Recurrent neural networks (RNNs) are especially preferred in the fields of linguistics and psychoanalysis. It is preferred in structures used to predict the next state from a sequential state that occurs sequentially. For example, you want to make a sentence. It can be used when it is desired to predict what kind of sentence will continue after words come one after another (Doğan, F., & Türkoğlu, İ., 2019.)

RNN performs the same analysis for each item in an input dataset. The items in each output dataset depend on all the calculated analyses, considering the inputs. In the RNN architecture, it has a structure that allows the information used in the past to be stored and used again for future analysis. RNN aims to obtain information at time t in one dimension, and it is also used to obtain both past and future information bidirectionally (Karakuş, B. A. 2018.)

In the RNN architecture, a loop is created to determine the behavior of a time series through connections between nodes (Aydın G. & Önün F., 2019.)

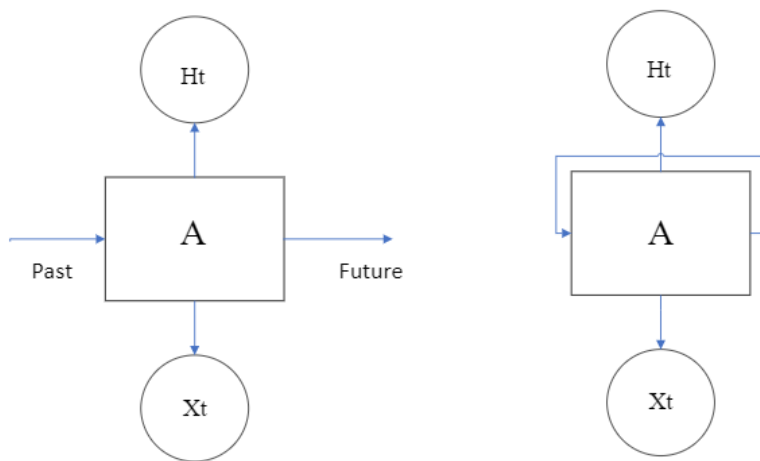


Figure 3.8: Structure of Recurrent Neural Network Cell

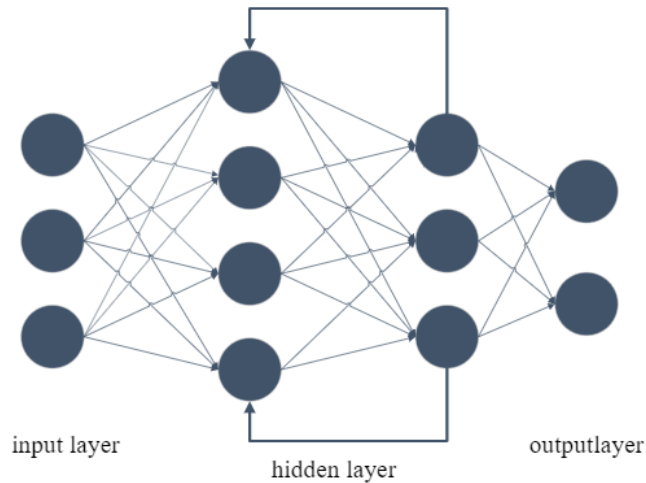


Figure 3.9: Recurrent Neural Network Structure

3.5.4 LONG- SHORT TERM MEMORY (LSTM)

LSTM gives successful results in applications such as analysis and forecasting in areas such as natural language recognition, voice recognition, health and economics. LSTM was developed to eliminate the time gaps in the RNN architecture. Unlike RNN, LSTM contains a memory layer. The LSTM technique has been developed because the amount of information that is required to be remembered for a long time is insufficient with RNN. With LSTM, long-term information is remembered, and it is decided which part of this information should be remembered and which part should be forgotten (Karakuş, B. A. 2018.)

Since the desired result in the RNN is not related to the current input but to the previous result, the errors are calculated by back propagation. As a result of this backpropagation, the vanishing gradient problem limits the models, so the training capacity of the system is lacking. In order to eliminate this problem, it is aimed to solve the more complex structure with the forget gate developed with LSTM (Öztanır, O., 2018.)

Unlike standard feed-forward neural networks, LSTM has feedback connections. It has the advantage of processing the entire data array rather than single data points. LSTM is considered an efficient tool as memory cells encode information and evaluate long-term relationships (Bahri et al., 2020.)

While there is generally one tanh network layer in the RNN structure, there are four tanh network layers in the LSTM structure and these networks are responsible for remembering long and short time periods in the model (Canbay et al., 2021.)

The existing cells in the LSTM evaluate the information coming into the network and decide whether it is necessary for the model or not. If the information is not appropriate, it is forgotten. As a result, the hidden layer is improved and the dependency problems between the data arrays are solved (Wang, Zheng, Ai, Liu, & Zhu, 2020.)

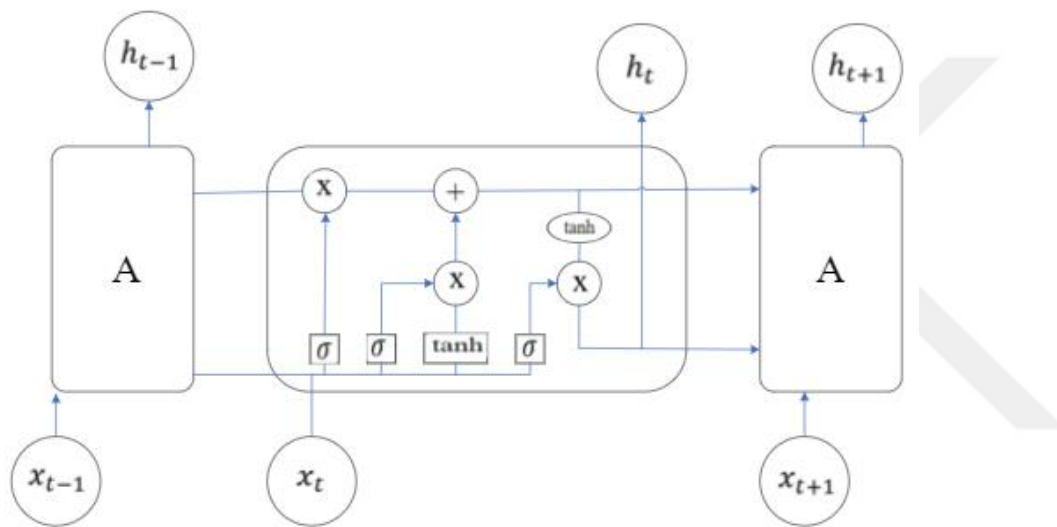


Figure 3.10: Structure of Long-Short Memory Cell

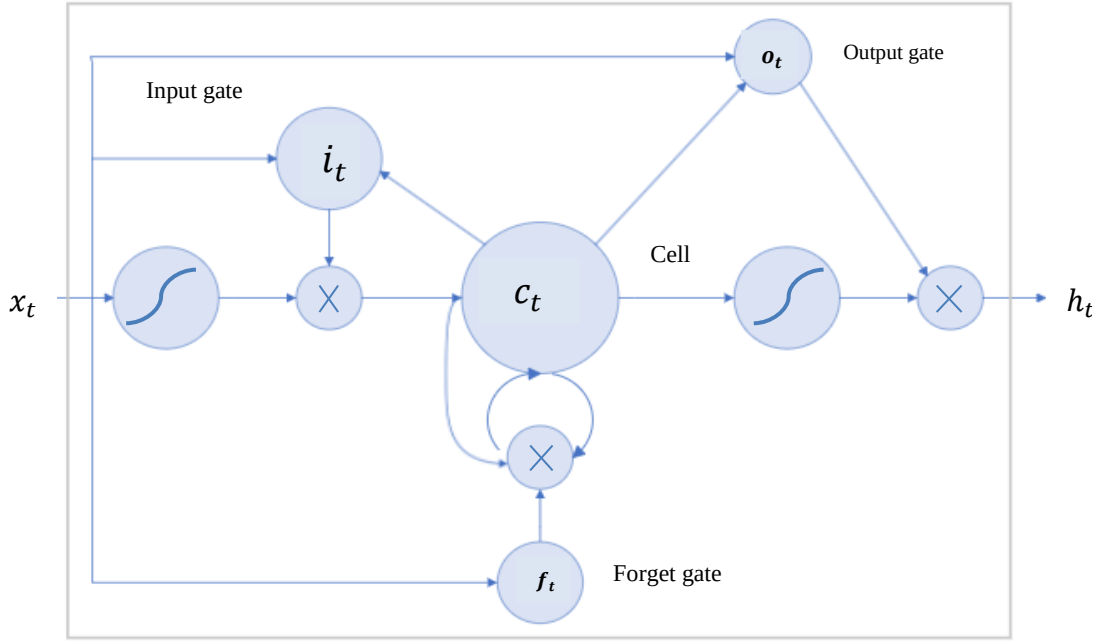


Figure 3.11: Structure of LSTM Algorithm

LSTM cell architecture consists of memory cell, input gate, forget gate and output gate (Devaraj et al., 2021.) The input, forget, and output gates control the flow of information in the neural network through the cells. In the memory cell, the information is stored and helps to be transmitted to the next cell unchanged (Er & Işık. 2021.) The unnecessary information in the information coming to the neural network is deleted and the important information is kept (Karakuş, B. A. 2018.)

The input gate decides to select the information coming into the cell and to receive the necessary information. The input gate formula is as follows:

$$i_t = \sigma(w_i * [h_{t-1}, x_t] + b_i) \quad (3.1)$$

The '*' operator expresses element multiplication. The sigmoid function decides the input (x_t) at time t and the output (h_{t-1}) of the last unit at time $t-1$. In other words, the sigmoid function determines the information taken into or deleted from the cell.

As shown in Equation 4, it is used to determine the information that will not be evaluated in the previous memory with the Forget gate (f_t). In short, it decides which information to forget with the sigmoid function. h_{t-1} ranges from 0 to 1. In the cell state (c_t) it is a vector with a value of c_{t-1} for each number. w_f is the weight matrix of the forget gate and b_f is the bias (Er & Işık, 2021.)

$$f_t = \sigma(w_f * [h_{t-1}, x_t] + b_f) \quad (3.2)$$

The next step consists of two parts, including the sigmoid function and the tanh function. Sigmoid function: between 0 and 1 values, it decides whether to update the information. The Tanh function decides the importance level of the information between -1 and 1 and assigns weight to the values of the information.

The cell state is updated by the forget gate of the current cell state by processing the information.

$$\tilde{c}_t = \tanh(w_c * [h_{t-1}, x_t] + b_c) \quad (3.3)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (3.4)$$

The output port (o_t) decides how much output will be updated, and the information of the hidden layer are formed with the output data.

$$o_t = \sigma(w_o * [h_{t-1}, x_t] + b_o) \quad (3.5)$$

$$h_t = o_t * \tanh(c_t) \quad (3.6)$$

Briefly, LSTM consists of four gates and two memory cells. The input gate receives information from short-term memory and transfers it to other gates for learning. The forget gate takes the data from the long-term memory and forgets the unnecessary data, while updating the new information by adding the learned data to the long-term memory. The output gate processes the learned information and makes analysis and predictions with long-term memory. Short-term memory is then updated using this information (Karakuş, B. A. 2018.)

3.5.4.1 Stacked-LSTM

Stacked-LSTM is expressed as an extension of LSTM. Stacked-LSTM consists of multiple memory cells and multiple hidden layers in addition to standard LSTM. This method, which consists of multiple layers, increases the depth of neural networks. Stacked-LSTM also provides a separate output for each cell, rather than a single output for the entire structure (Arora, Kumar, & Panigrahi, 2020.) Stacked-LSTM is used to make predictions using time series data (Devaraj et al., 2021.)

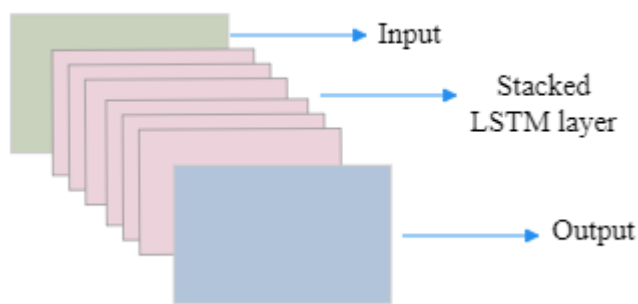


Figure 3.12: Structure of Stacked-LSTM

(Arora, P., Kumar, H., & Panigrahi, B. K.2020.)

3.5.4.2 Bidirectional LSTM (Bi-LSTM)

When the standard LSTM are examined, the inputs only function one-sided. To solve this problem, Bi-LSTM usage has been developed. The Bi-LSTM model is a modified model of LSTM so that it can be trained both forward and backward. So Bi-LSTM has two layers, forward and backward (Betül et al.,2017.) In addition, Bi-LSTM aims to reach the information of the present time by evaluating both the past and the future. Bi-LSTM learns faster than traditional LSTM (Ayoobi et al., 2021.)

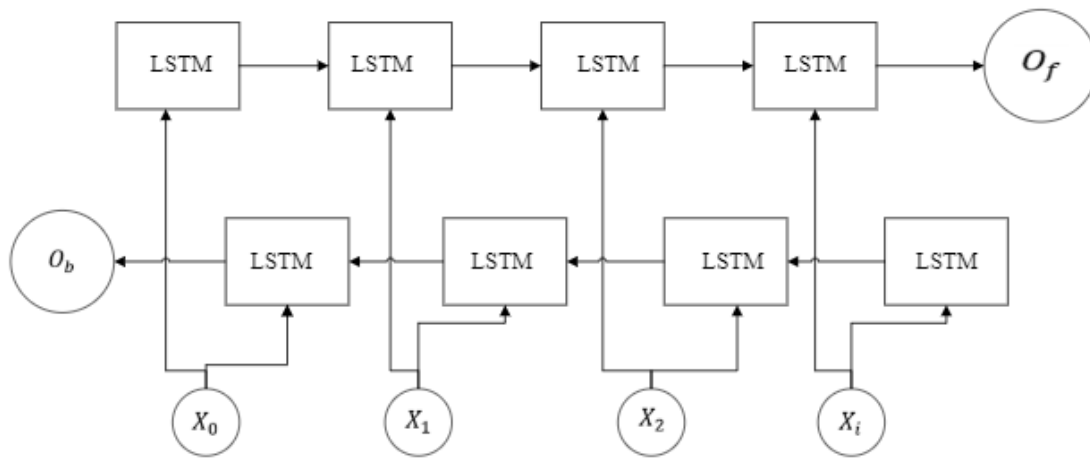


Figure 3.13: The Bi-LSTM Structure

3.5.4.3 Convolutional LSTM (Conv-LSTM)

The structure of this model is similar to biological neural networks. In Conv-LSTM, neurons in the network have certain weights and these neurons are more closely connected to each other, unlike traditional neural networks. Connections between neurons are simplified. The structure of this method consists of convolutional, batch normalization, rectified linear unit (relu), pooling and full link layer.

Convolutional layer: responsible for feature extraction and feature map generation by filtering inputs.

Pooling layer: it provides easier computation of data processing by reducing the feature map size and obtaining a subset for processing the data.

Relu layer: it is a kind of enabler to preserve non-linear data and size.

Fully connected layer: it provides the output by mapping the attribute outputs in the appropriate form (Marzouk, Elshaboury, Abdel-Latif, & Azab, 2021.)

3.5.4.4 Gated recurrent unit (GRU)

A gated recurrent unit is a simpler version of the LSTM in terms of less training time and a lower total number of gates. Calculations when making predictions with GRU are simpler than with LSTM. In GRU architecture, there is a hidden state cell where an update gate is used, which includes a forget gate and an input gate. Architecture works with reset gate. In short, it uses the group cell state and the hidden state as a single state.

The hidden state of the GRU cell is as follows:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (3.7)$$

The update gate decides how much of the GRU cell to update. With the update gate, it indicates how much of the previous information will be used as the next information. If the calculated value of the update gate is close to zero, the old information is not considered, only the new information data is considered (Karakuş, B. A. 2018.) And it is calculated as follows:

$$z_t = \sigma(w_z * [h_{t-1}, x_t]) \quad (3.8)$$

The reset gate determines which of the past information will be forgotten. The reset gate is calculated very similarly to the update gate (ArunKumar, Kalaga, Kumar, Kawaji, & Brenza, 2021.)

$$r_t = \sigma(w_r * [h_{t-1}, x_t]) \quad (3.9)$$

3.6 GRAY WOLF OPTIMIZER ALGORITHM

The Gray wolf optimization algorithm is a type of herd-based meta-heuristic algorithm developed by modeling the behavior of gray wolves, such as herd hunting and social hierarchy (Mirjalili, Mirjalili, & Lewis, 2014.) It was developed by Mirjalili in 2014 and was used for optimization in ANNs in 2015 (“Gri Kurt Algoritması Hakkında Bilgi”, Retrieved March 17, 2022.)

Many AI algorithms are probability-based and use control parameters such as population size and repetition parameters when optimizing problems. In order for these algorithms to yield effective results, control parameters must be used correctly. In the GWO algorithm, population size and the number of iterations is considered as control parameters. It is also aimed at minimizing the control parameters of the GWO algorithm (Uzlu, 2019.)

Considering the herd characteristics of gray wolves, which are at the top of the food chain, four basic behavior patterns were determined. These; social hierarchy is to contain prey, hunt and attack prey (Karakoyun, M. 2021.)

1. Social Hierarchy: While creating the GWO algorithm, the individuals in this herd population are from the top to the bottom; It is represented as alpha (α), beta (β), delta (δ), and omega (Ω) (Arikan & Latifoğlu, 2021.) Each individual in the population must have only one of these represented roles. If the hierarchy is considered in the algorithm, the team consisting of 3 individuals with the strongest, that is, the most suitable solution, is called the leader team (Karakoyun, M.2021.)

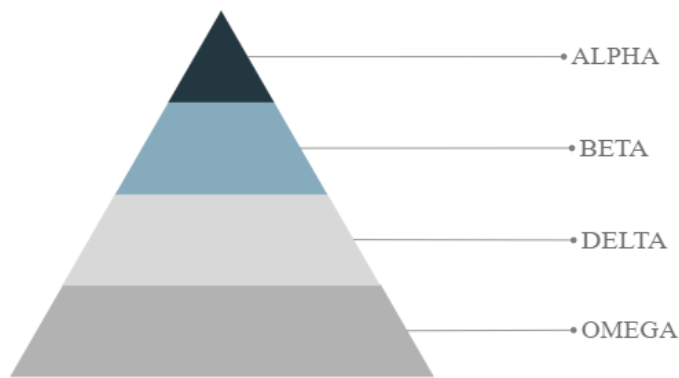


Figure 3.14: GWO Algorithm Social Hierarchy

The most suitable first-ranked individuals in the leader team are alpha, second and third-ranked beta and delta individuals, respectively. Another remaining member of the hierarchy is the omega individual. Alpha individuals are individuals with a leadership spirit. They play a decisive role in determining behaviors such as hunting, sleeping places, and waking places. The decisions made by the alpha individual are implemented for the herd. Alpha individuals cannot be said to be the best for the herd, but they are strong individuals in terms of herd management and discipline. Beta individuals, on the other hand, take part in making decisions, reinforcing the commands of alpha individuals, and counseling alpha individuals. They can replace alpha individuals when they age and die. Omega individuals are the lowest ranked wolves, and other lead wolves dominate omega wolves. Omega individuals seem unimportant to the herd, but they have a stabilizing role in turmoil and problems in the herd. While delta individuals dominate omega wolves, they cannot dominate alpha and beta wolves. This group includes scouts, sentries, the elderly, hunters, and caretakers. This group includes scouts, sentries, the elderly, hunters, and caretakers. Scouts are responsible for monitoring the boundaries of the area and alerting the group when danger occurs. The guards are in charge of ensuring and protecting the safety of the herd. Old wolves were previously alpha and beta and

are experienced wolves. Prey wolves assist alphas and betas in hunting prey to provide food for the flock. Nurse wolves, on the other hand, are responsible for the treatment of sick, old, and injured wolves in the herd. In the GWO algorithm, the hunting optimization is done by alpha, beta, and delta individuals. The omega individual, on the other hand, follows these three leading wolves (Mirjalili, Mirjalili, & Lewis, 2014.)

2. Encircling Prey: Gray wolves show prey engaging in behavior during hunting.

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \quad (3.10)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad (3.11)$$

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (3.12)$$

$$\vec{C} = 2\vec{r}_2 \quad (3.13)$$

In equations, the positions of the gray wolves around the prey in any random area are updated.

3. Hunting: Gray wolves have features such as recognizing their prey, locating the prey, and encircling the prey. The hunt is led by alpha, but beta and delta individuals can sometimes join the hunt. Beta has knowledge of the best candidate solution, while Delta has knowledge of the potential location of the hunt.

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \quad (3.14)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \quad (3.15)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (3.16)$$

4. Attacking Prey: The hunting process ends when the gray wolves attack the prey. To model the gray wolf approaching the prey, the value of is reduced by considering iteration. It is

linearly reduced from 2 to 0 from the range $[2a, 2a]$ and takes a random value. As $\vec{A}$ decreases, the $\vec{A}$ value decreases and in cases of $|A| < 1$ the gray wolf can attack the prey (Mirjalili, Mirjalili, & Lewis, 2014).

5. Search for Prey: The gray wolves move away from each other to search for prey. While modeling the prey search $\vec{A}$, the search is conducted using values greater than 1 and less than -1. The $|A| > 1$ status helps make the hunt for prey more efficient. The C parameter is used as another parameter for the search. The value of $\vec{C}$ takes a random value between 0 and 2. Values of $C > 1$ are provided to indicate the importance of the prey, and $C < 1$ to reduce the importance of the prey. Because of this, GWO can work with random values, allowing for exploration and moving away from the local optimum (Doğan, C. 2019.)

```

Initialize the grey wolf population  $X_i (i = 1, 2, \dots, n)$ 
Initialize  $a, A$ , and  $C$ 
Calculate the fitness of each search agent
 $X_\alpha$  = the best search agent
 $X_\beta$  = the second best search agent
 $X_\delta$  = the third best search agent
while ( $t < \text{Max number of iterations}$ )
    for each search agent
        update the position of the current search agent by equation
    end for
    update  $a, A$ , and  $C$ 
    Calculate  $X_\alpha, X_\beta$ , and  $X_\delta$ 
     $t = t + 1$ 
end while
return  $X_\alpha$ 

```

Figure 3.15: Pseudo Code of the GWO Algorithm (Mirjalili, Mirjalili, & Lewis, 2014.)

3.7 ERROR MEASURES

A_i = actual values

$\bar{A}$ = the mean of the actual values

P_i = predicted values

$e_i = y_i - \hat{y}_i$ error

n = size of the dataset

(“MSE, RMSE, MAE, MAPE ve Diğer Metrikler – Veri Bilimcisi”, retrieved from July 04, 2022.)

EVS (Explained Variance Score)

It is used when evaluating whether there is a fit between model data and real data. In short, it is measured whether there is variability in the estimates. High EVS values, indicates a strong relationship between the data and good prediction. The EVS formula is shown in equation 3.17.

$$EVS = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i - \bar{e})^2} \quad (3.17)$$

MAE (Mean Absolute Error)

The absolute mean of error is the measured distance between each true value and the appropriate value. It has negative biased performance that measures the average size of errors in the data prediction. Measurements with a lower order perform better. The MAE formula is shown in equation 3.18.

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (3.18)$$

MAPE (Mean Absolute Percentage Error)

The mean absolute percent error is used to measure and compare the accuracy of model estimators. The MAPE formula is shown in equation 3.19.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|e_i|}{|y_i|} \quad (3.19)$$

MSE (Mean Squared Error)

The mean squared error measures the performance of the estimator. Predictive values with MSE values close to zero show better performance. The MSE formula is shown in equation 3.20.

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (3.20)$$

MSLE (Mean Squared Logarithmic Error)

It is expressed as a measure of the ratio that exists between actual values and predicted values. It only cares about the percentage difference between the actual value and the predicted value. The MSLE formula is shown in equation 3.21 (“Mean Squared Logarithmic Error (MSLE) | Peltarion Platform,” Retrieved from July 4, 2022).

$$L(A, P) = \frac{1}{n} \sum_{i=0}^n (\log(y_i + 1) - \log(\hat{y}_i + 1))^2 \quad (3.21)$$

RMSE (Root Mean Square Error)

It is a performance measure that measures the size of the error between the predicted value and the true value for a dataset. RMSE is the standard deviation of the prediction errors. Calculates the magnitude of the density between the appropriate values, considering the values that fit the data. When this performance criterion, takes the value of zero, we can say that the model is almost error-free. The RMSE formula is shown in equation 3.22, 3.23.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad (3.22)$$

$$RMSE = \sqrt{MSE} \quad (3.23)$$

R^2 SCORE

R^2 score interprets to what extent the independent variable of one model affects the variance value of the other dependent variable (Tüzüntürk, 2010). The R^2 score indicates how much the independent variable of one model affects the variance value of the other dependent

variable (“R-Squared Definition”, Retrieved from July 04, 2022). It is preferred when deciding how accurate the predictive value of the model used is. The R^2 score value can be at most 1, and the model is successful when the measured value is close to 1. The R^2 score formula is shown in equation 3.24, 3.25.

$$R^2 = 1 - \frac{\text{sum of squares of error}}{\text{average sum of differences}} \quad (3.24)$$

$$R^2(y, \hat{y}) = 1 - \frac{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2}{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \bar{y})^2} \quad (3.25)$$

4. RESULTS AND DISCUSSIONS

In this thesis, a predictive analysis for Turkey is performed using COVID-19 data from the Turkovac website. Figure 4.1 shows graphical summary of the study. First of all, the data to be used in the study and the source of the data is determined. Afterwards, preprocessing of the data is performed to bring the data format suitable for analysis. After training with 70% training and 30% test data, the model selection is made and the algorithms to be used is determined. The predicted results are evaluated by comparing them with the actual data. In this section, the dataset, obtained results, and the comparative analysis are given. The accuracy of the models is examined in the prediction analysis of COVID-19 using the Stacked-LSTM, Bi-LSTM and LSTM algorithms. A single input variable was used for each model analysis and a single output variable was obtained. Briely, each data group in the datasets was examined for individual models. In addition, the GWO algorithm is used to optimize the LSTM, Bi-LSTM and Stacked-LSTM algorithms. The analysis is performed by using a computer Intel(R) Core (TM) i7-8700K CPU @ 3.70GHz 64,0 GB RAM.

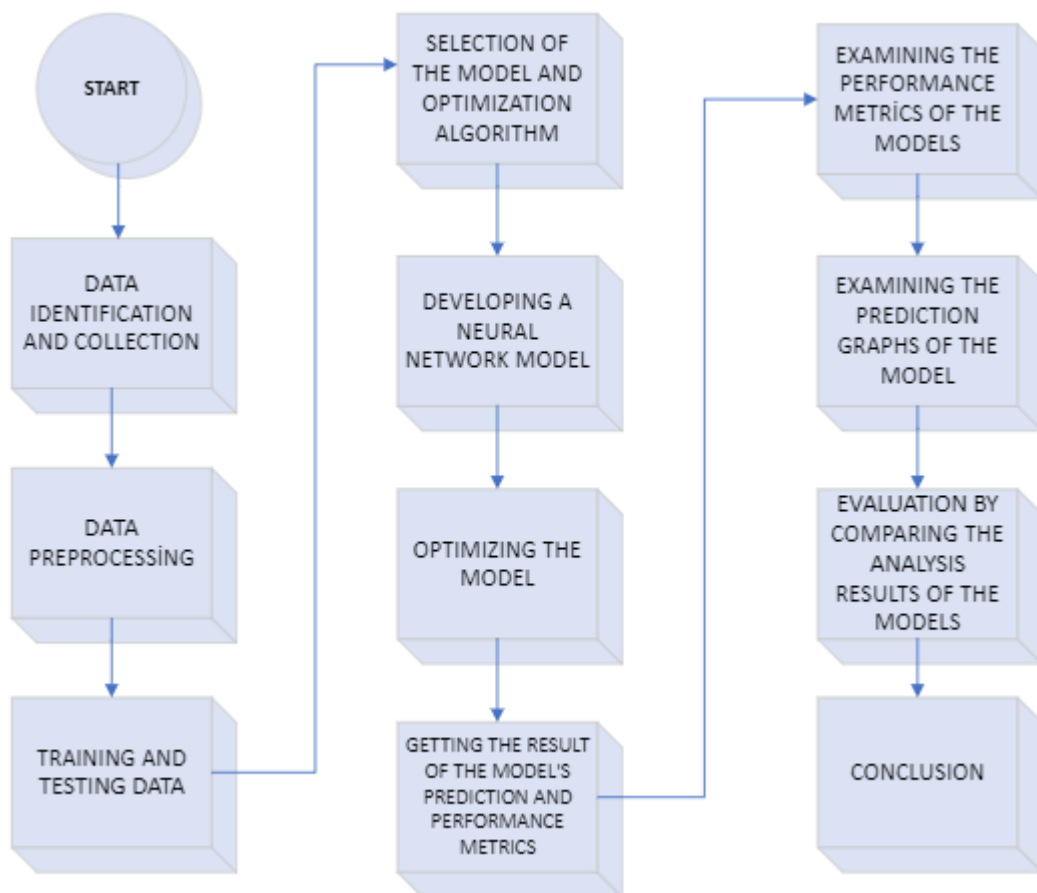


Figure 4.1: Flow Chart of the Study

4.1 Description of Datasets

The study dataset consists of daily cases, daily deaths, cumulative cases and cumulative deaths numbers for Turkey. The first COVID-19 cases in Turkey was seen on March 11, 2020. However, in this study, the data between June 13, 2021 and December 7, 2021 is considered for daily deaths data. The data between August 24, 2021 and December 7, 2021 were selected for daily cases. Cumulative deaths data is from June 11, 2021 to December 7, 2021. The data between August 22, 2021 and December 7, 2021 were analyzed for the cumulative cases.

Table 4.1: Descriptive Statistics of COVID-19 Dataset

Statistics					
		Daily cases	Daily deaths	Cumulative cases	Cumulative deaths
Mean		25559	165	6303259	60357
Median		25694,5000	197	6317215	58913
Std. Deviation		3583,0260	77	832666	9524
Variance		12838071	5996	693332938319	90714976
Skewness		-0,2670	-0,4980	-0,0010	0,3490
Kurtosis		-0,7460	-1,2780	-1,3070	-1,3310
Range		16528	255	2705487	29424
Minimum		17332	35	4961946	48593
Maximum		33860	290	7667433	78017
Sum		2709317	29547	674448731	10803997
Percentiles	25	23052,7500	70,5000	5540983	50934
	50	25694,5000	197,5000	6317215	58913
	75	28483	226,2500	7061707	68917

Table 4.1 shows the descriptive statistics for the considered dataset. Descriptive statistics includes numerical measurements such as mean, standard deviation, variance, kurtosis, skewness, range, minimum and maximum values. Descriptive statistics are used to describe data and help transform data into usable information. Considering mean values, daily cases,

daily deaths, cumulative cases and cumulative deaths data are concentrated around 25559, 165, 6303259, 60357, respectively.

The standard deviation indicates how far each piece of information deviates from the mean. Looking at the standard deviation values in the table, there is a slight difference between the mean and standard deviation for the daily cases, while the difference between the standard deviation and the mean for daily deaths, cumulative deaths, and cumulative cases are greater. Kurtosis and skewness values indicate that the dataset is normally distributed since the range of values in each dataset is between -1 and +1.

The variance indicates variability of the dataset. From table 4.1, it is clearly seen that the number of cumulative cases dataset has the highest variability. Also, the daily deaths dataset has the lowest variability and is more homogeneously distributed compared to others cases data.

While the smallest number of daily cases is 17332, the highest number of daily cases is 33860. Among the considered dates, the lowest highest number of deaths per day are 35 and 290, respectively. According to the table, the maximum number of cumulative cases is 7667433, while the maximum deaths is 78017.

Table 4.2: Parameters of GWO

Parameters	Parameters values of the GWO
Number of iterations	100
Number of populations	20

In GWO algorithm the number of iterations and the number of populations are set to 100 and 20, respectively. Parameters of GWO algorithm are listed in Table 4.2. The population is the results obtained for each iteration. It represents the herd group for the GWO algorithm. The number of populations used for GWO is 20. The iteration parameter to be used for the GWO algorithm is 100. In this study, the same parameter values, given in table 4.2, are used while optimizing the LSTM, Bi-LSTM, and Stacked-LSTM techniques.

Table 4.3: Parameters of LSTM

Parameters	Parameter Values of the LSTM Algorithm
Node size	[8, 16, 32, 64, 128]
Optimizers	['Adadelata', 'Adagrad', 'Adam', 'Adamax', 'Ftrl', 'Nadam', 'RMSprop', 'SGD']
Activation functions	['relu', 'sigmoid', 'softplus', 'softsign', 'tanh']
Step size	[2, 4, 6, 8, 10, 12, 14]
Epoch size	[100, 500, 1000]
Batch size	[8, 16, 32, 64]
Learning rate	[0.01, 0.001, 0.0001, 0.00001]

The parameters used to optimize LSTM techniques with the GWO algorithm are listed in Table 4.3. In table 4.3, possible values for the parameters, node size, epoch size, step size, and batch size are listed. Also, relu, sigmoid, softplus, softsign, and tanh functions are considered as the possible activation functions. The table also gives the lists of the possible optimizers and the learning rate. As mentioned, these parameters are optimized using GWO algorithm.

4.2 PREDICTION RESULTS with STACKED-LSTM

Firstly, the prediction of COVID-19 cases with the Stacked-LSTM algorithm is examined. COVID-19 data is analysed in terms of daily cases, daily deaths, cumulative deaths, and cumulative cases.

4.2.1 Prediction Charts

In this section, a graph of the predicted number of cases using Stacked-LSTM technique is given. The graphs reveal the relationship between the predicted values and the actual data values for each data set clearly. The blue-colored lines show the predicted values and the red-colored lines show the actual values. Note that, dataset is considered to be a time series dataset.

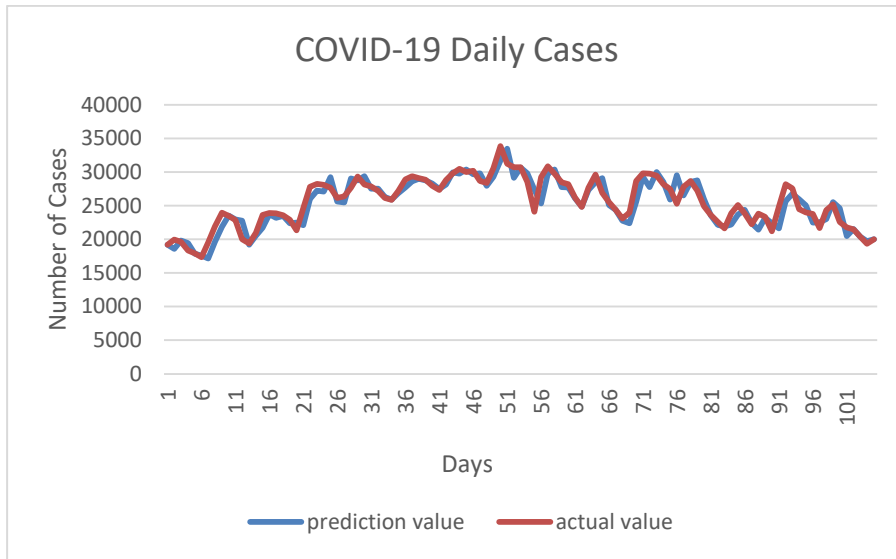


Figure 4.2: Predicted Values and Actual Values of COVID-19 Daily Cases

Figure 4.2 shows the graph of the actual number of daily cases and the predicted values for the number of daily cases. If the MAPE value is examined for daily cases, the Stacked-LSTM MAPE value is %10. The fact that this value, which measures the accuracy of the prediction values, is low indicates that the performance accuracy of the model is strong. Consequently, the graph reveals that one can get accurate and robust prediction of COVID-19 data using the Stacked-LSTM algorithm.

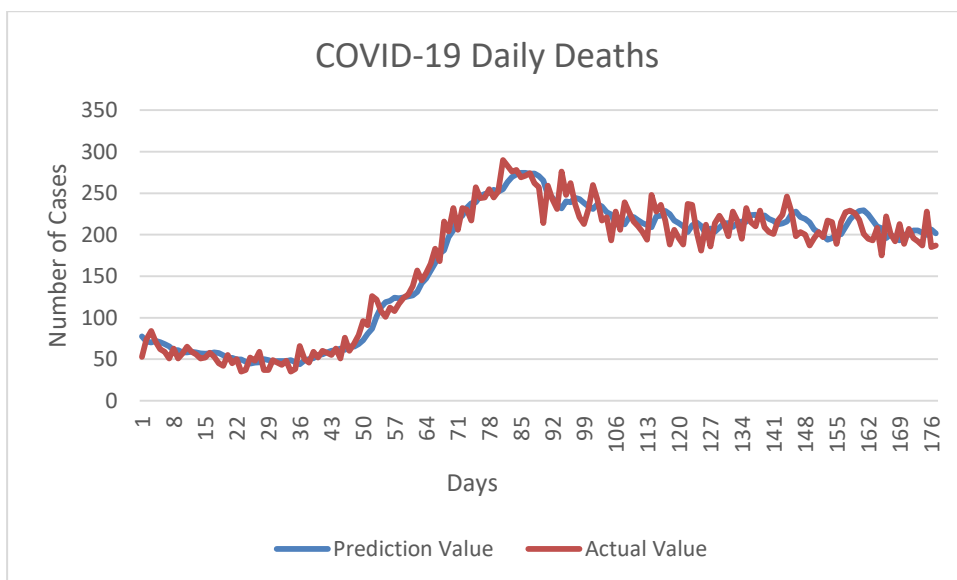


Figure 4.3: Predicted Values and Actual Values of COVID-19 Daily Deaths

Figure 4.3 compares daily deaths to their predicted values. For the daily deaths examined, the dataset shows very close deaths numbers on some days. Although there are no significant differences between days in the number of deaths, from the figure it is observed that it reaches a peak value on some days. The MAPE value related with the daily deaths' prediction is 0.09 which indicates that the prediction error significantly low. Consequently, the daily deaths prediction accuracy with Stacked-LSTM model is considerably high which means the predicted daily deaths and the actual values are very close to each other. Daily deaths have fewer error values when comparing daily cases and daily deaths according to error values.

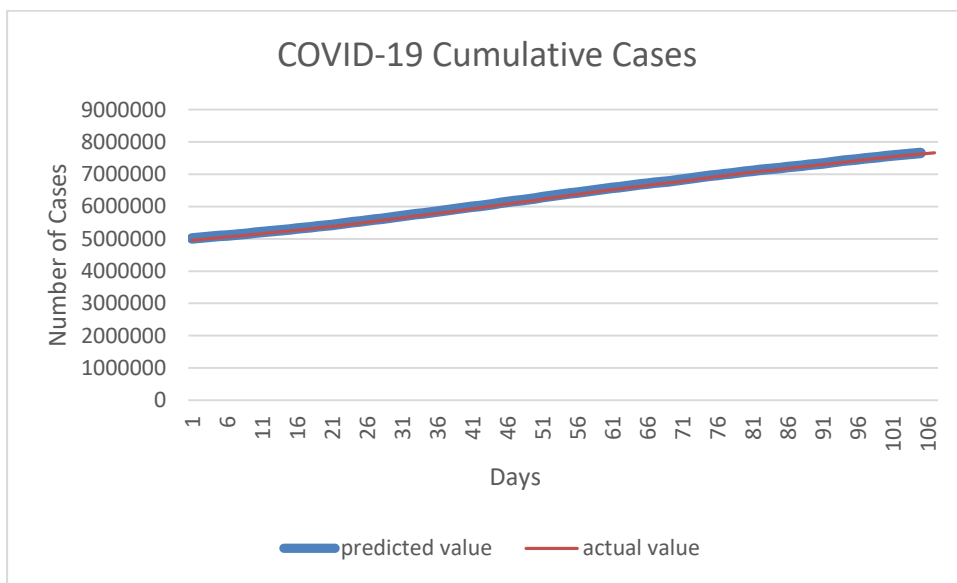


Figure 4.4: Predicted Values and Actual Values of COVID-19 Cumulative Cases

Figure 4.4 shows actual and predicted values for of cumulative cases. From the graph, the predicted values are quite close to the actual values, and it is clearly seen that there is a cumulative increase. The MAPE value of cumulative cases is smaller than daily cases and daily deaths and has a value of %0.3. As a result, the prediction performance of the Stacked-LSTM technique with all considered datasets are superior. Still, the best prediction performance is achieved on the of cumulative cases dataset.

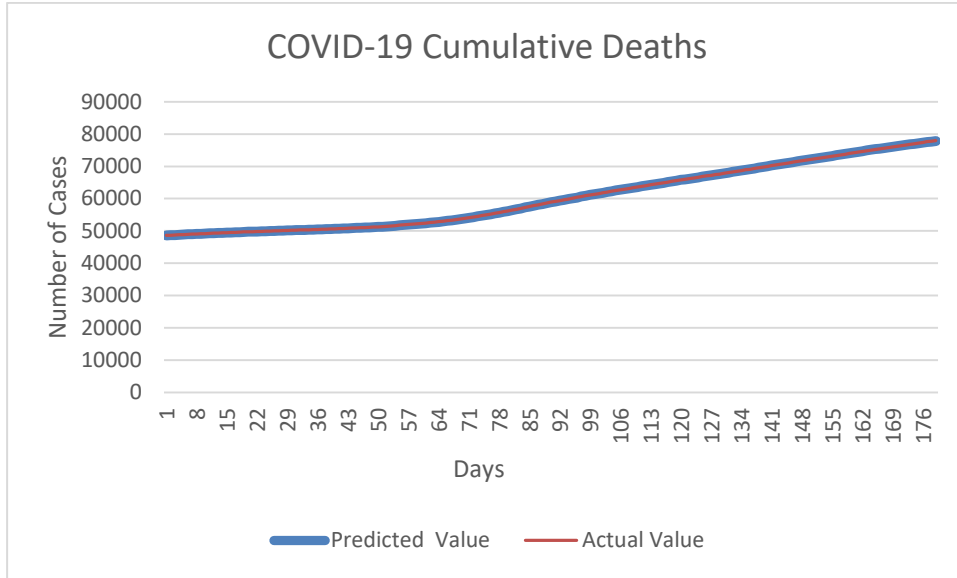


Figure 4.5: Predicted Values and Actual Values of COVID-19 Cumulative Deaths

When Figure 4.5 is examined, it is understood that the predicted cumulative deaths values are very close to the actual values. Among the performance metrics for cumulative cases, the MAPE value is 0.002 and has the least error value. Therefore, when compared to other datasets, the best performance in the prediction technique using the Stacked-LSTM analysis was found for cumulative deaths.

4.2.2 Comparative Analysis of Best Results for Datasets

In this section, the best results of the parameter values of the Stacked-LSTM model, which were first optimized with the GWO algorithm, are given. In addition, performance metrics obtained using Stacked-LSTM are compared in each dataset.

Table 4.4: Optimized Stacked-LSTM Parameter Values

Datasets				
	Daily cases	Daily Deaths	Cumulative cases	Cumulative Deaths
1. Hidden Layer	8	16	16	16
2. Hidden Layer	128	16	64	8
3. Hidden Layer	64	8	128	32
Optimizers	Nadam	Adam	Adam	Adam
1. Activation Functions	Softplus	Relu	Softsign	Sigmoid
2. Activation Functions	Tanh	Sigmoid	Tanh	Softsign
3. Activation Functions	Tanh	Sigmoid	Tanh	Tanh
Epoch Size	500	100	500	500
Step Size	8	12	8	10
Batch Size	8	8	8	8
Learning Rate	0,001	0,001	0,01	0,01

Table 4.4 shows the parameter values of Stacked-LSTM optimized with the GWO algorithm. From table 4.4 it is seen that three hidden layers are determined for the Stacked-LSTM algorithm. Number of hidden layers for of daily cases are optimized to be 8, 128 and 64 units, respectively. Number of hidden layers for of daily deaths dataset are optimized to be 16, 16, 8. While the hidden layer numbers are optimized to be 16, 64, 128 for of cumulative cases dataset and 16, 8, 32 for of cumulative deaths dataset.

While optimizing the data with the GWO algorithm, the Nadam optimizer is chosen for of daily cases dataset, while the Adam optimizer is chosen for of daily deaths, of cumulative cases, and of cumulative deaths datasets.

While analyzing the models, 3 activation functions and 3 hidden layers are used for each dataset. The model chose softplus and tanh activation functions as the best parameters for daily cases. Activation function parameters for daily deaths are relu and sigmoid functions. In the Stacked-LSTM model, the best activation function parameters for cumulative cases are determined as sigmoid and tanh. In addition, sigmoid, softsign and tanh activation functions were selected for cumulative deaths. The size of the epoch is determined to be 100 for the

daily deaths' dataset, while it is 500 for of daily cases, cumulative cases, and cumulative deaths datasets.

Step size is determined to be 8 for of daily cases and of cumulative cases datasets and it is 12 and 10 for of daily deaths and of cumulative deaths datasets, respectively. Batch size is determined to be 8 for each dataset. The learning rate is set to be 0.001 for of daily cases and of daily deaths datasets, and it is 0.01 for of cumulative cases and of cumulative deaths datasets.

Table 4.5: Comparison of Performance Metrics of the Stacked-LSTM

	Datasets			
	Daily Cases	Daily Deaths	Cumulative cases	Cumulative Deaths
EVS	0,8627360	0,9564890	0,9999617	0,9999776
MAE	0,0479100	0,0398590	0,0011718	0,0008315
MAPE	0,1070100	0,0968130	0,0033430	0,0020151
MSE	0,0042670	0,0025730	0,0000020	0,0000010
MSLE	0,0019580	0,0010470	0,0000010	0,0000005
RMSE	0,0653200	0,0507270	0,0014194	0,0011751
R²	0,8572580	0,9563600	0,9999616	0,9999776

EVS values for daily cases, daily deaths and for both cumulative cases and cumulative deaths are 86%, 95%, 99%, respectively. Since higher EVS values indicates the higher compatibility between the cumulative model data and the cumulative actual data.

MAE values for the daily cases, the daily deaths, the cumulative cases, and the cumulative deaths are 0.047, 0.039, 0.001 and 0.0008, respectively. It should be noted that, lower MAE values are achieved with the cumulative values. Since the MAE is a performance metric with avoidant Stacked-LSTM models with cumulative cases and cumulative deaths datasets outperform Stacked-LSTM models with daily cases and daily deaths datasets. On the other hand, Stacked-LSTM model with cumulative deaths dataset performs better than the Stacked-LSTM model with cumulative cases dataset.

Considering the MAPE values, it is seen that the cumulative cases and cumulative deaths numbers both have very low MAPE values. While cumulative deaths have a percentage of 0.2%, the cumulative cases percentage is 0.3%. Thus, it can be said that the Stacked-LSTM model best performed on the of cumulative deaths dataset in terms of MAPE values which is an indicator of prediction accuracy.

Predictive performance of the model improves as the MSE value gets closer to zero. Since all MSE values are very close to 0 the prediction performance of the Stacked-LSTM algorithm on each dataset is very similar to each other. In this regard, it can be said that the prediction performance of the Stacked-LSTM algorithm on each dataset is very good based on the MSE value, but it is still the best performance achieved on the cumulative deaths' dataset.

MSLE values are 0.1% for daily cases and daily deaths, while MSLE values for cumulative deaths and cumulative cases are lower than 0.1% which means the ratio between the actual values and the predicted values is very small for all datasets.

The RMSE value is pretty close to zero for all COVID-19 datasets. As a result, the prediction errors are acceptable and the proposed model gives robust and acceptable predictions, especially with the cumulative cases and the cumulative deaths datasets.

The R^2 values for the daily cases, for the daily deaths, and for both the cumulative cases and the cumulative deaths are 0.85, 0.95, and 0.99, respectively. The predictive performance of the model increases as the R^2 value gets closer to 1. Looking at the R^2 value of the data, except for the daily cases, the others are around 0.9. In spite of being worse than the others the R^2 value with the daily cases dataset is still in acceptable limits. It should be noted that R^2 value is an indicator for the strength of the relationship between dependent and independent variables and the correctness of the model expected to increase with higher R^2 values. Thus, according to these R^2 values more accurate predictions are expected with the cumulative cases and the cumulative deaths rates datasets.

Table 4.6 CPU Times (in seconds) for All Datasets

Datasets			
Of Daily Cases	Of Daily Deaths	Of Cumulative Cases	Of Cumulative Deaths
131584,5	122934,7	126515,4	249028,9

Table 4.5 gives CPU times for Stacked-LSTM optimized with the GWO optimized algorithm. From Table 4.5, it is apparently seen that the minimum CPU time is required to process of daily deaths dataset while the maximum one is required by of cumulative deaths dataset. The required CPU time directly depends on the topology of the Stacked-LSTM algorithm for each dataset which is optimized by GWO algorithm.

4.3 PREDICTION RESULT with Bi-LSTM

In this section the predictive analysis performed using the Bi-LSTM model is discussed. As in the previous section, the results of daily cases, daily deaths, cumulative cases, and cumulative deaths datasets are considered. Firstly, for each dataset, the topology and the parameters of the Bi-LSTM algorithm is optimized using GWO algorithm. Then, the predetermined performance metrics are obtained for each dataset using the corresponding optimized Bi-LSTM prediction model.

4.3.1 Prediction Charts using Bi-LSTM prediction model

In this section, the actual data values and their corresponding predicted values are graphed and discussed for each dataset.

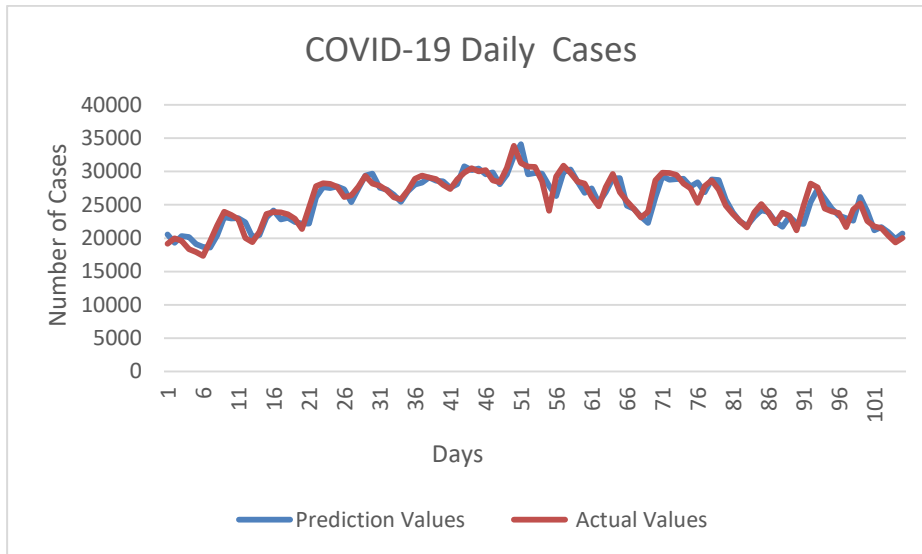


Figure 4.6: Predicted Values and Actual Values of COVID-19 Daily Cases

Figure 4.6 shows the actual and the predicted values for daily cases over the considered time period. From the figure, it is seen that the actual and the predicted values are very close to each other. Also, in the dataset it is clearly seen that there exist sharp increases and decreases over time which means the data is not smooth. From the dataset, the actual data value and its corresponding predicted value using the Bi-LSTM model for 81st day is the same and is 22583. Similarly, looking at day 50 of the data, the actual value of the daily number of cases and corresponding predicted value is 31,248 and 34,110, respectively. These sample results reveal that actual values and predicted values using the Bi-LSTM algorithm are pretty close each other.

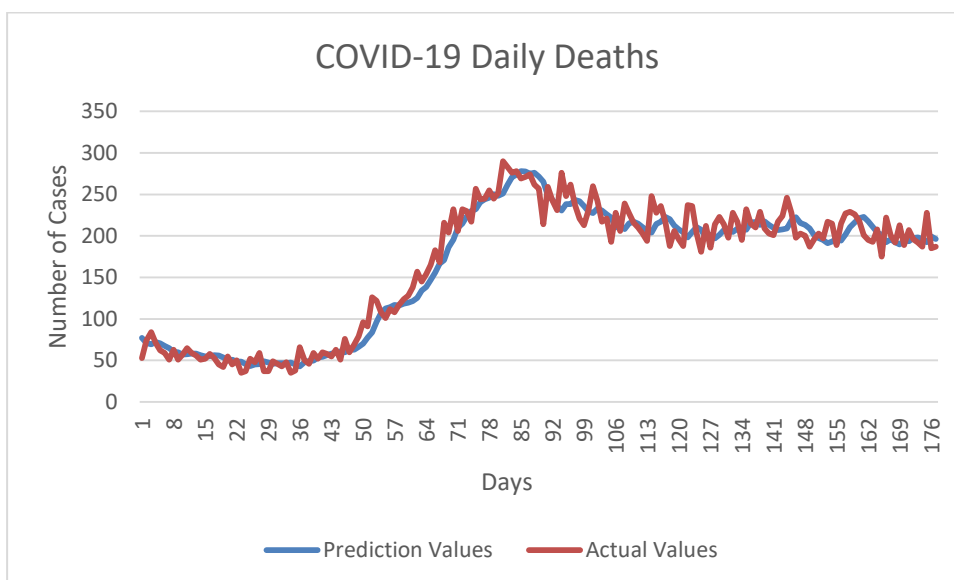


Figure 4.7: Predicted Values and Actual Values of COVID-19 Daily Deaths

In Figure 4.7 the time series plots of the actual values and the predicted values for the daily deaths are shown on top of each other. The number of deaths stabilizes at around 50 deaths during the first 40 days and this number gradually increases between 40 and 85 days. From the figure it is clearly seen that the actual deaths values fluctuate while the predicted values are more stable. Similarly, looking at day 80 of the data, the actual value of the daily number of cases and corresponding predicted value is 290 and 250, respectively. Thus, for this specific day, Bi-LSTM underestimates the daily deaths. However, this does not apply to the whole dataset. Looking at two more samples from the dataset, i.e., on the 76th day, the actual and its corresponding predicted values are 245 and 243, respectively. And, on the 73rd day, the actual and the predicted of daily deaths are 217 and 230, respectively. These sample values reveal that the actual and predicted values using Bi-LSTM algorithm are very close to each other. MAPE value which is calculated to be 0.09 also supports this prediction accuracy.

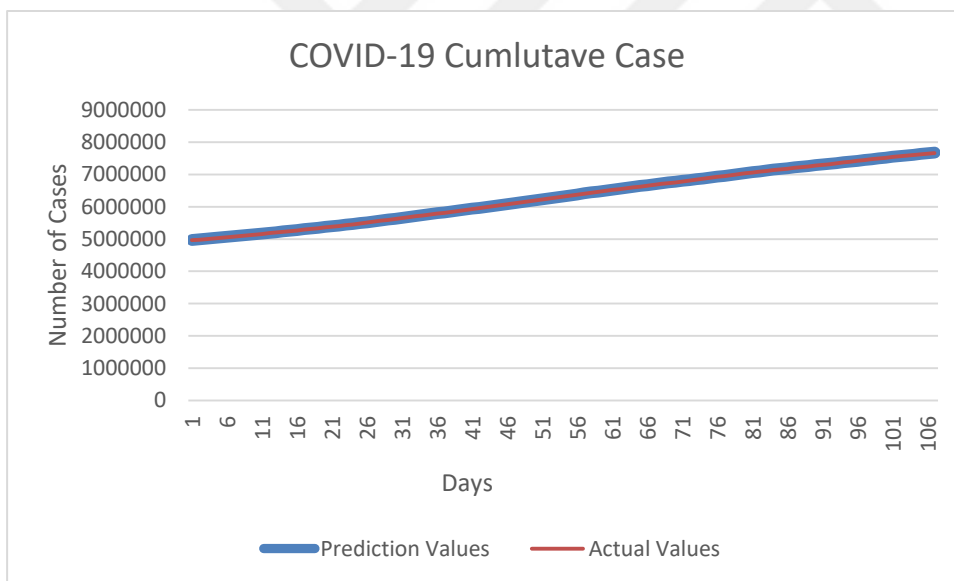


Figure 4.8: Predicted Values and Actual Values of COVID-19 Cumulative Cases

Figure 4.8 shows the plot of actual and predicted values for COVID-19 of cumulative cases dataset. The values shown with the blue line represents predicted of cumulative cases and the values shown with the orange line represents the actual of cumulative cases. From the figure, it is clearly seen that the predicted values of the cumulative cases using the Bi-LSTM model are very close to actual values.

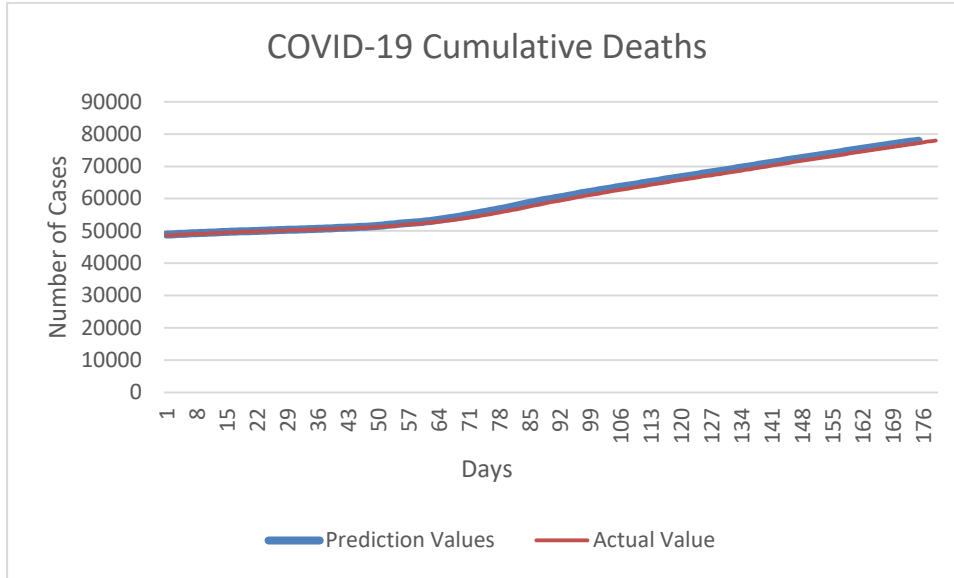


Figure 4.9: Predicted Values and Actual Values of COVID-19 Cumulative Deaths

Figure 4.9 graph the actual and the predicted values of the cumulative measurements. From the figure a smooth positive trend line is seen up to 55th day and it is getting steeper after 78th day.

If the graph is examined, it is understood that the predicted values obtained using the Bi-LSTM model are very close to the actual values. Until the 50th day, the blue and orange lines are plotted just on top of each other but after this time the blue lines remains slightly over the orange line which indicates that the Bi-LSTM model overestimates the cumulative deaths for this time period. Considering the MAPE values, it is seen that the cumulative deaths has very low MAPE values. Cumulative deaths have a percentage of 0.2%. Thus, it can be said that Bi-LSTM model best performed on the of cumulative deaths dataset in terms of MAPE values which is an indicator of prediction accuracy.

4.3.2 Comparative Analysis of Best Results for Datasets

In this section, the best parameter values of the Bi-LSTM model, which is optimized with the GWO algorithm, are examined one by one for the datasets. In addition, the best performance metric values of the COVID-19 dataset are given with the Bi-LSTM technique, and the results of the datasets are evaluated comparatively.

Table 4.7: Optimized Bi-LSTM Parameter Values

Datasets				
	Daily Cases	Daily Deaths	Cumulative Cases	Cumulative Deaths
1.Hidden Layer	128	32	16	64
2.Hidden Layer	128	128	32	32
3.Hidden Layer	64	128	64	128
Optimizers	RMSprop	RMSprop	RMSprop	Adamax
1.Activation Functions	Relu	Relu	Relu	Sigmoid
2.Activation Functions	Tanh	Tanh	Softsign	Tanh
3.Activation Functions	Tanh	Tanh	Sigmoid	Tanh
Epoch Size	100	100	100	1000
Step Size	8	12	6	14
Batch Size	8	8	8	8
Learning Rate	0,01	0,00001	0,0001	0,001

Table 4.7 shows the parameter values of Bi-LSTM optimized with the GWO algorithm. From table 4.7 it is seen that three hidden layers are determined for the Bi-LSTM algorithm. Number of hidden layers for of daily cases are optimized to be 128, 128, and 64 units, respectively. Number of hidden layers for of daily deaths dataset are optimized to be 32,128,128. While the hidden layer numbers are optimized to be 16, 32, 64 for of cumulative cases dataset and 64, 32, 128 for of cumulative deaths dataset.

In the Bi-LSTM model, the RMSprop optimizer is chosen for daily cases, daily deaths, cumulative cases, and the epoch size is known as 100. For cumulative deaths, Adamax optimizer for of cumulative deaths dataset is chosen and epoch size is 1000.

While analyzing the models, three activation functions and three hidden layers are used for each dataset. The model chose relu and tanh activation functions as the best parameters for daily cases and daily deaths. The activation functions parameters for the cumulative cases are determined relu, softsign, and sigmoid functions. In addition, sigmoid and tanh activation functions parameters were chosen for cumulative deaths. The step size is determined to be 8, 12, 6, and 14 for daily cases, daily deaths, cumulative cases, and cumulative deaths, respectively. The batch size is the same for the all of dataset and is set to 8. The learning rates

for daily cases, daily deaths, cumulative cases, and cumulative deaths are 0.01, 0.00001, 0.0001, and 0.0001 respectively.

Table 4.8: Comparison of Performance Metrics of the Bi-LSTM for COVID-19

Datasets				
	Daily Cases	Daily Deaths	Cumulative Cases	Cumulative Deaths
EVS	0,8881370	0,9544270	0,9999680	0,9999850
MAE	0,0443870	0,0405390	0,0030610	0,0007520
MAPE	0,1100270	0,0959470	0,0083210	0,0020530
MSE	0,0033480	0,002774	0,0000110	0,0000009
MSLE	0,0015280	0,0011300	0,0000053	0,0000004
RMSE	0,0578620	0,0526710	0,0033340	0,0009740
R²	0,8879940	0,9529520	0,9997950	0,9999840

Table 4.8 gives the performance metric obtained using optimized Bi-LSTM algorithm.

First of all, EVS values for daily cases, daily deaths and for both cumulative cases and cumulative deaths are %88, %95, %99, respectively. Since higher EVS values indicates the higher compatibility between cumulative model data and the cumulative actual data.

MAE values for daily cases and daily deaths are approximately 0.4. The MAE value for cumulative deaths is 0.003, but the MAE value for cumulative deaths is 0.0007. It should be noted that lower MAE values are achieved with the cumulative deaths. In the Bi-LSTM model, the MAE value of cumulative cases is smaller and has better performance than the MAE values of dail cases and daily deaths. Consequently, in the Bi-LSTM model, the MAE performance value of cumulative deaths outperforms daily cases, daily deaths, and cumulative cases involving another dataset.

MAPE value indicates the accuracy of prediction values. The MAPE values for the cumulative cases and the cumulative deaths are approximately 0.008 and 0.002, respectively. Thus, it can be said that the prediction accuracies of Bi-LSTM models on the cumulative cases and the cumulative deaths datasets are better than those of others. Within

daily values, it can be said that daily deaths show better performance accuracy than daily cases.

Prediction performance improves as the MSE value gets closer to 0. From Figure 4.6, it is seen that the MSE value is very close to 0 for all datasets. However, it is much lower for of cumulative cases and of cumulative deaths datasets. Then, the datasets with the best prediction performance according to the MSE value are of cumulative deaths and of cumulative cases.

MSLE value gives the ratio between the actual and the predicted values. These values are 0.001 for both of daily cases and daily deaths datasets. And, it is much smaller for of cumulative cases and cumulative deaths datasets. As can be seen from these results, the difference between the predicted and the actual values is insignificant which is evidence in favor of the performance of the Bi-LSTM model's prediction performance. In terms of the prediction performance, it is desired that the RMSE values are close to zero. According to Table 4.6, the RMSE value for both of daily cases and daily deaths datasets is 0.05. The RMSE values of of cumulative cases and of cumulative deaths datasets are 0.003 and 0.0009, respectively. Thus, the prediction error predicted for of cumulative deaths dataset based on RMSE value is very low. So, with respect to RMSE value Bi-LSTM algorithm's prediction performance is superior on of cumulative deaths dataset.

The R^2 values for the daily cases, for the daily deaths, and for both the cumulative cases and the cumulative deaths are 0.88, 0.95, and 0.99 respectively. The predictive performance of the model increases as the R^2 value gets closer to 1. Looking at the R^2 value of the data, expect for the daily cases, the others are around 0.9. In spite of being worse than the others the R^2 value with the daily cases dataset is still in acceptable limits. It should be noted that R^2 value is an indicator for the strength of the relationship between depend and indepent variables and the correctness of model expect to increase with higher R^2 values. Thus, according to these R^2 values more accurate predictions are expected with the cumulative cases and the cumulative deaths rates datasets.

Table 4.9: CPU Times (in seconds) for All Datasets

Datasets			
Of Daily Cases	Of Daily Deaths	Of Cumulative Cases	Of Cumulative Deaths
227045,8	233771,8	158317,5	458506,6

Table 4.8 shows the CPU time of the Bi-LSTM model optimized with the GWO algorithm using COVID-19 data. When evaluated for COVID-19 datasets, it is seen that the fastest results are obtained for cumulative cases. The slowest results were achieved for cumulative deaths.

4.4 PREDICTION RESULTS with ANN

In our study, prediction models were created using LSTM models, Stacked-LSTM models, and Bi-LSTM models. In this section, a different model was created by using the ANN method and predictive analysis was performed. As the previous section, the result of daily cases, daily deaths, cumulative cases, and cumulative deaths datasets are considered. Firstly, for each dataset, the topology and the parameters of the ANN algorithm is optimized using GWO algorithm. Then, the predetermined performance metrics are obtained for each dataset using the corresponding optimized ANN prediction model.

Daily cases, daily deaths, cumulative cases and actual and prediction data graphs of cumulative deaths, optimized parameter values, and performance metrics were compared in tables and evaluated.

4.4.1 Prediction Charts Using ANN Prediction Model

In this section, a graph of predicted number of cases using ANN technique is given. The graphs reveal the relationship between the predicted values and the actual data values for each dataset clearly. The blue-colored lines show the predicted values and the red-colored lines show the actual values. Separate graphs for daily cases, daily deaths, cumulative cases, and cumulative deaths are analyzed.

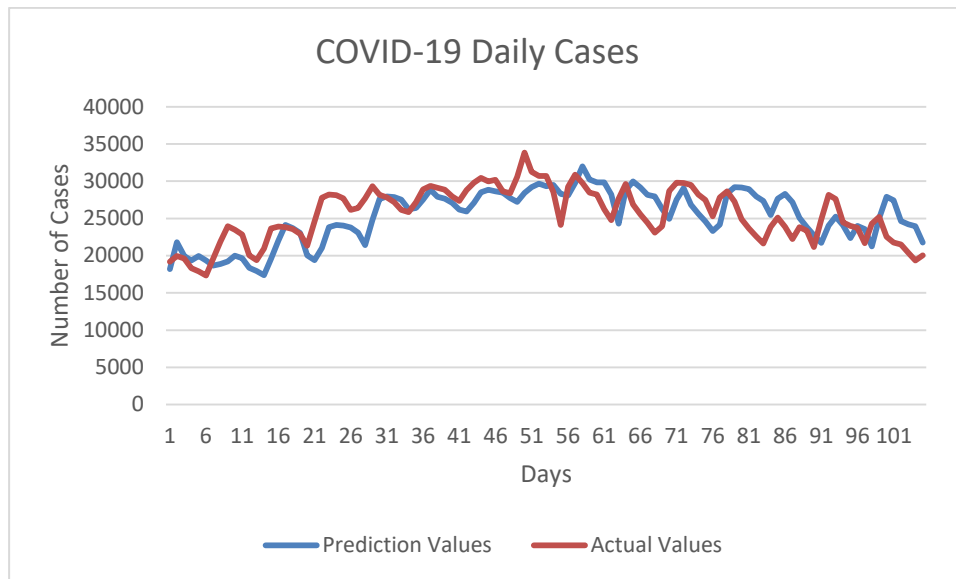


Figure 4.10: Predicted Values and Actual Values of COVID-19 Daily Cases

Figure 4.10 shows the graph of the predicted daily cases and the actual daily cases with the ANN model. Looking at the graph, it is seen that the real data has fluctuating values. The predicted values follow the actual data, but it is understood that the forecast values are predicted as a farther value on some days. Looking at the first day of data, it is known that the actual daily cases value is 19191 and the predicted value is 18193, which means it is predicted to be 998 days less. When the 49th day of data is examined, it is known that the value of actual daily cases is 33860. The value of the predicted daily cases on the 49th day is 28429. The predicted value of 5431 was underestimated. As a result, when the graph and data values are examined, it is seen that the predicted values are very close to the actual value as well as very far.

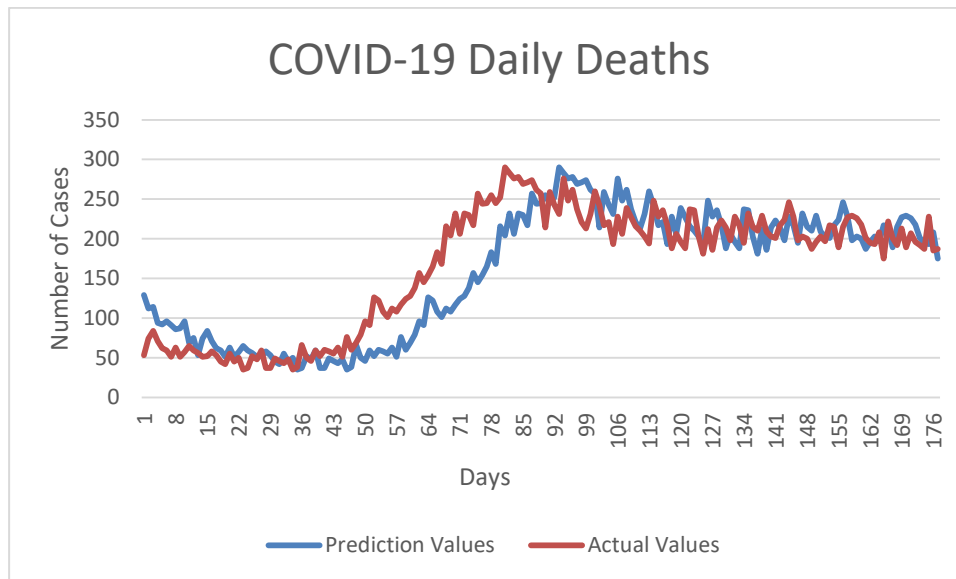


Figure 4.11: Predicted Values and Actual Values of COVID-19 Daily Deaths

Figure 4.11 shows the graph of predicted values and actual values using the ANN model for daily deaths. It is seen that the predicted values between Day 50 and Day 90 are far from the real values compared to the whole dataset. It is seen that the predicted values between Day 100 and Day 176 are very close to the actual values.

While the daily number of deaths on the 53rd day was 126, the predicted value was calculated as 52. In other words, the predicted result was 74 less than the actual value. Looking at the 60th day data, the actual value is 124, while the predicted value is 60. As a result, the predicted value of daily deaths on day 60 was underestimated by 64 difference. The predicted value and the actual value in the 120th day data are 120. The model made one-to-one predictions.

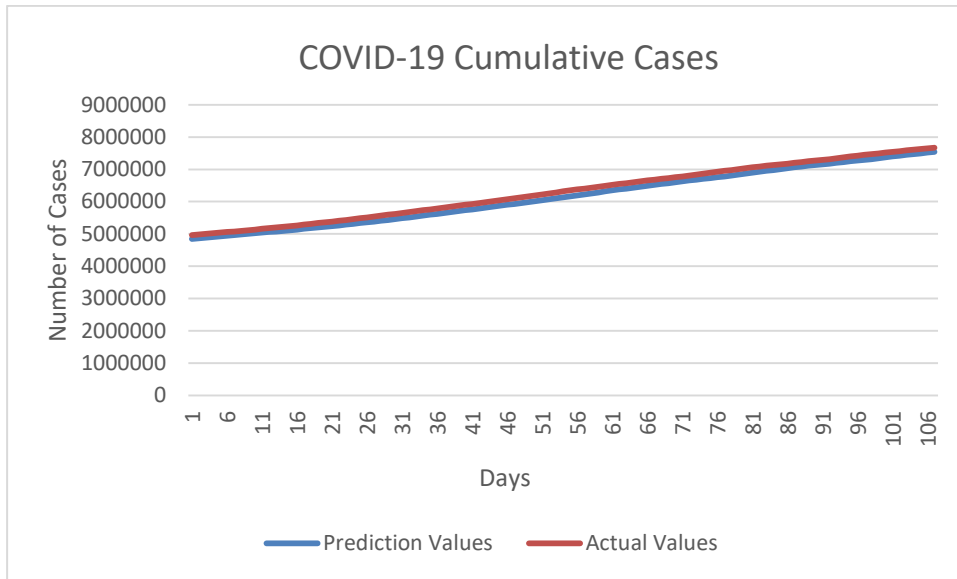


Figure 4.12: Predicted Values and Actual Values of COVID-19 Cumulative Cases

Figure 4.12 shows the graph that combines the actual cumulative cases values with the predicted values of cumulative cases with the ANN model using COVID-19 data. When the graph is examined, it is seen that there is an increasing trend. The blue lines show the predicted cumulative cases; the red lines show the actual value of the cumulative cases. In this cases, where these lines are adjacent in the graph, it shows that the predictive value makes very close predictions.

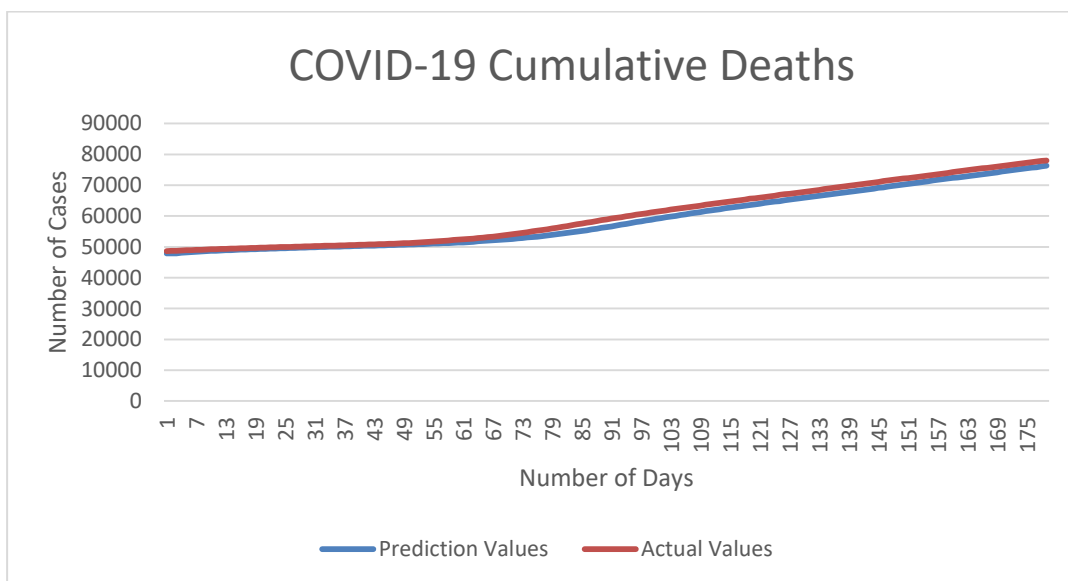


Figure 4.13: Predicted Values and Actual Values of COVID-19 Cumulative Deaths

Figure 4.13 shows a graph of the actual and predicted values of the cumulative death numbers for COVID-19 modeled by ANN. When the graph is examined, it is observed that a fairly stable number of deaths cases have progressed for the first 55 data points, and it is understood that the predicted values are almost the same as the actual values. It is understood that there is an increase in the number of 60th cases more rapidly after the data, and it can be understood that the predicted values are less than the actual values.

4.4.2 Comparative Analysis of Best Results for Datasets

In this section, the best parameter values of the ANN model, which is optimized with the GWO algorithm, are examined one by one for the datasets. In addition, the best performance metric values of the COVID-19 dataset are given with the ANN technique, and the results of the datasets are evaluated comparatively.

Table 4.10: Optimized ANN Parameter Values

	Datasets			
	Daily Cases	Daily Deaths	Cumulative Cases	Cumulative Deaths
1.Hidden Layer	128	128	128	128
2.Hidden Layer	32	16	16	8
3.Hidden Layer	8	16	16	8
Optimizers	Adadelata	Adamax	Adamax	Adam
1.Activation Functions	Relu	Softsign	Softsign	Softplus
2.Activation Functions	Sigmoid	Softsign	Softsign	Relu
3.Activation Functions	128	128	128	128
Epoch Size	Softsign	Softplus	Softplus	Softplus
Step Size	500	1000	1000	1000
Batch Size	8	8	8	8
Learning Rate	0,01	0,01	0,01	0,01

Table 4.10 shows the best parameter values of the ANN model optimized with the GWO algorithm. While optimizing the ANN model, 3 hidden layers and 3 activation functions were selected. For the daily cases, the hidden layer is selected as 128, 32, and 8 respectively. The number of hidden layers for daily deaths and cumulative cases is determined as 128, 16, 16, respectively. For cumulative deaths, the hidden layer is set to 128, 8, and 8, respectively.

In the ANN model, the Adamax optimizer is chosen for daily deaths and cumulative cases. Adadelta and Adam optimizers were used as optimizers for daily cases and cumulative deaths, respectively.

Activation functions Relu, Sigmoid, and Softsign are selected for the daily cases. The activation function for daily deaths and cumulative cases has been chosen as Softsign, Softsign, and Softplus. The activation function for cumulative deaths has been chosen as Softplus, Relu, and Softplus.

Epoch size is set to 500 for daily cases, 1000 for daily deaths, cumulative deaths and cumulative cases. Batch size is calculated as 8 for the entire dataset. The learning rate was taken as 0.01 for the entire dataset.

Table 4.11: Comparison of Performance Metrics of the ANN for COVID-19

	Datasets			
	Daily Cases	Daily Deaths	Cumulative Cases	Cumulative Deaths
EVS	0,7970980	0,9999120	0,9999120	0,9999720
MAE	0,0645110	0,0019930	0,0019930	0,0011480
MAPE	0,1563750	0,0047860	0,0047860	0,0049800
MSE	0,0067280	0,0000054	0,0000054	0,0000018
MSLE	0,0031620	0,0000025	0,0000025	0,0000011
RMSE	0,0820250	0,0023240	0,0023240	0,0013470
R^2	0,7958460	0,9999070	0,9999070	0,9999710

In Table 4.11 gives the performance metric obtained using optimized ANN algorithm. Performance metrics in the table are examined one by one. First of all, EVS value is examined, it is 0.79 for the daily cases. The EVS value for daily deaths, cumulative cases, and cumulative deaths is approximately 0.99. Since higher EVS values indicates the higher compatibility daily cases.

The MAE and MAPE values for the daily cases are 0.06 and 0.15, respectively. Considering the MAE and MAPE values for daily deaths, cumulative cases, cumulative deaths are 0.001 and 0.004, respectively. It should be noted that lower MAE values and MAPE values are achieved. Consequently, in the ANN model, the MAE and MAPE performance values for daily deaths, cumulative cases, cumulative deaths outperform daily cases.

If the MSE value is examined, it is seen that it has very small values for the input dataset. It is even smaller, especially for daily deaths, cumulative cases, and cumulative deaths. According to the result of this performance metric, it is said that the prediction performance of the model is very strong.

Another performance metric, the MSLE value, has very small implications for datasets when examined. Daily cases dataset is slightly higher than other datasets. But in general, this low MSLE value for the input dataset shows that the ratio between the actual value and the predicted value is very low. Thus, we can conclude that very close prediction results can be achieved.

When the RMSE performance value is examined, it is 0.08 for the daily cases. It is 0.002 for daily deaths and cumulative cases. For cumulative deaths, the RMSE value is 0.001. The cumulative deaths have the lowest error according to the RMSE value, which expresses the smallness of the error estimate value. In general, although the datasets are low, daily cases have the highest error rate.

The R^2 values for the daily deaths, cumulative cases, cumulative deaths are 0.99 and the R^2 values for daily cases is approximately 0.79. the predictive performance of the model increases as the R^2 value gets closer to 1. Looking at the R^2 value of the data, expect for the daily cases, the others are around 0.9. In spite of being worse than the others the R^2 value with the daily cases dataset is still in acceptable limits. It should be noted that R^2 value is an indicator for the strength of the relationship between depend and indepent variables and the correctness of model expect to increase with higher R^2 values. Thus, according to these R^2 values more accurate predictions are expected with the daily deaths, the cumulative cases, the cumulative death rates datasets.

Table 4.12: CPU Times (in seconds) for All Datasets

Datasets			
Of Daily Cases	Of Daily Deaths	Of Cumulative Cases	Of Cumulative Deaths
74128,81	107528,9	107528,9	161778,8

Table 4.11 provides information about CPU time as a result of optimizing COVID-19 data with ANN. When Table 4.11 is examined, it is said that the model worked the fastest for daily deaths and cumulative cases. A very slow performance was shown for the daily cases.

Table 4.13: Performance Metrics' Comparison of Models for Daily Cases

Datasets			
	Stacked-LSTM	Bi-LSTM	ANN
EVS	0,862736	0,888137	0,797098
MAE	0,04791	0,044387	0,064511
MAPE	0,10701	0,110027	0,156375
MSE	0,004267	0,003348	0,006728
MSLE	0,001958	0,001528	0,003162
RMSE	0,06532	0,057862	0,082025
R^2	0,857258	0,887994	0,795846

In Table 4.12, performance metrics for the daily cases are examined according to the models. When the EVS and R^2 values are examined, it is seen that the Bi-LSTM model has the highest compatibility compared to other models. The Stacked-LSTM has the lowest value of 0.10701 for the MAPE performance metric. The MAPE value for Bi-LSTM is 0.110027 and the MAPE value for ANN is 0.156375. However, when the table is examined in general, performance metrics such as MAE, MSE, MSLE, RMSE have lower values for Bi-LSTM. According to these metrics with negative slope performance, it is understood that the most appropriate estimator for the daily cases is the Bi-LSTM model.

Table 4.14: Performance Metrics' Comparison of Models for Daily Deaths

Datasets			
	Stacked-LSTM	Bi-LSTM	ANN
EVS	0,9564890	0,9544300	0,9999120
MAE	0,0398590	0,0405400	0,0019930
MAPE	0,0968130	0,0959500	0,0047860
MSE	0,0025730	0,0027700	0,0000054
MSLE	0,0010470	0,0011300	0,0000025
RMSE	0,0507270	0,0526700	0,0023240
R^2	0,9563600	0,9529500	0,9999070

Table 4.14 gives a comparison of the performance metrics of the models for daily deaths. By examining the table, the model that makes the strongest prediction is determined. Looking at the R^2 value in the table, it is seen that it is approximately 0.99 for the ANN model. This value, which is very close to 1, shows that the prediction measure of the ANN model is quite strong. When all other values are examined, it is seen that the most appropriate values are for the ANN model.

Table 4.15: Performance Metrics' Comparison of Models for Cumulative Cases

Datasets			
	Stacked-LSTM	Bi-LSTM	ANN
EVS	0,9999620	0,9999680	0,9999120
MAE	0,0011720	0,0030610	0,0019930
MAPE	0,0033430	0,0083210	0,0047860
MSE	0,0000002	0,0000111	0,0000054
MSLE	0,0000010	0,00000533	0,0000025
RMSE	0,0014190	0,0033340	0,0023240
R^2	0,9999620	0,9997950	0,9999070

Table 4.14 gives a comparison of the performance metrics of the models for the cumulative cases. When the EVS value for cumulative cases is examined, the Bi-LSTM model has a higher value than other models. Looking at the MSE value, the Bi-LSTM model has the

lowest value. For cumulative cases, the lowest error prediction based on negatively biased values such as MAE, MAPE, MSLE, RMSE is successful for Stacked-LSTM.

Table 4.16: Performance Metrics' Comparison of Models for Cumulative Deaths

Datasets			
	Stacked-LSTM	Bi-LSTM	ANN
EVS	0,9999780	0,9999850	0,9999720
MAE	0,0008320	0,0007520	0,0011480
MAPE	0,0020150	0,0020530	0,0049800
MSE	0,0000014	0,0000009	0,0000018
MSLE	0,0000005	0,0000005	0,0000011
RMSE	0,0011750	0,0009740	0,0013470
R²	0,9999770	0,9999840	0,9999710

Table 4.15 shows the comparative results of the models' performance metrics for cumulative cases. The prediction measure of the Bi-LSTM model, which has the highest result by looking at the EVS and R^2 values, is very successful. When the MSLE value is examined, the ANN model has the lowest value. According to this value, the value of the amount between the actual value and the predicted value of the ANN model is very small. When MAE, MAPE, MSE values are compared for models, it can be said that the Stacked-LSTM model has the lowest value and according to these values, the predictive performance of the model is very good.

5. CONCLUSIONS

The current COVID-19 epidemic is a great threat to human health. In the first cases, the disease spread very quickly around the world. Various studies are carried out to obtain information about the disease, such as the prevention of the pandemic, the diagnosis of the disease, the number of cases, and the number of deaths. Analysis techniques such as AI, machine learning, and deep learning are at the forefront of the studies carried out to get information about the pandemic and make strategic decisions.

In this study, we used the COVID-19 data taken from the Turkovac website for Turkey. Using these data, the number of COVID-19 cases and deaths were predicted using Bi-LSTM, Stacked-LSTM and ANN techniques. At the same time, the techniques were optimized by estimating the GWO algorithm with 100 iterations and 20 population parameter values. The number of cases and deaths were examined separately for each technique, both daily and cumulatively, and the actual cases values and the predicted cases numbers were shown in comparison.

When the MAPE values for the Stacked-LSTM are examined, the predicted accuracy is 0.10, 0.9, 0.003, and 0.002 for daily cases, daily deaths, cumulative cases, and cumulative deaths, respectively. The dataset with the best prediction accuracy is the dataset with cumulative measurements for the Stacked-LSTM. The MAPE values for Bi-LSTM are 0.11, 0.09, 0.008, and 0.002 for daily cases, daily deaths, cumulative cases, and cumulative deaths, respectively. According to this value, where the prediction accuracy of the model is measured, the closest predicted dataset for Bi-LSTM is the dataset with cumulative measurements.

For the ANN model, the daily cases value is 0.15. The MAPE value of daily deaths, cumulative cases, and cumulative deaths is 0.004. According to the ANN model, it is concluded that the least successful predicted dataset is daily cases.

When the prediction accuracy of the models is evaluated in general according to the MAPE value, it is seen that the cumulative measurements have the best prediction values compared to all the models.

The Stacked-LSTM, Bi-LSTM, and ANN models had R^2 values of 0.85, 0.88, and 0.79, respectively, for daily cases. When the daily measurements are compared according to the models, it is seen that the ANN model gave the best estimation results. It is seen in table 4.13 that MSE and MSLE values have very small results. Also, when the R^2 value is compared for daily deaths, it is 0.956, 0.952 and 0.999 for Stacked-LSTM, Bi-LSTM and ANN, respectively. According to the R^2 result, since the ANN model has the highest value, it is seen that the model accuracy is quite high.

Considering the MAPE accuracy value for cumulative cases, the MAPE values for Stacked-LSTM, Bi-LSTM, and ANN are 0.003, 0.008, and 0.004. It is seen that the Stacked-LSTM model gives the most successful result according to the MAPE value for cumulative cases. According to cumulative measurements, the MAPE values for Stacked-LSTM, Bi-LSTM, and ANN models are 0.00201, 0.00205, and 0.00498 respectively. Although the values for the Stacked-LSTM and Bi-LSTM cumulative measurements are relatively close, the Stacked-LSTM gives more successful results.

While the COVID-19 datasets were analyzed with prediction models, they were optimized with the GWO algorithm. While applying the GWO optimization algorithm in our model, 3 different hidden layers and 3 different activation functions were selected.

In general, artificial intelligence-based studies such as machine learning, deep learning, and ANNs are productive studies to take strategic measures and make decisions in unexpected situations such as pandemics. As a result of the results of the LSTM, Bi-LSTM, and ANN models used in the study, it was observed that successful results were achieved for the pandemic data. In fact, each technique gave more successful results in different datasets. Bi-LSTM for daily cases and ANN for daily deaths made very successful predictions. The ANN model shows the least success for cumulative cases and cumulative deaths. When the study is evaluated as a whole, the LSTM techniques and the ANN method produce effective results for estimating the epidemic situation.

REFERENCES

- Alimadadi, A., Aryal, S., Manandhar, I., Munroe, P. B., Joe, B., & Cheng, X. (2020). Artificial intelligence and machine learning to fight COVID-19. *Physiological genomics*, 52(4), 200-202.
- Al-Najjar, H., & Al-Rousan, N. (2020). A classifier prediction model to predict the status of Coronavirus COVID-19 patients in South Korea. *European Review for Medical and Pharmacological Sciences*.
- Al-Rousan, N., & Al-Najjar, H. (2020). Data analysis of coronavirus COVID-19 epidemic in South Korea based on recovered and death cases. *Journal of Medical Virology*, 92(9), 1603-1608.
- Anabilim Dalı Danışman, E., & Galip AYDIN, D. (n.d.). *DERİN ÖĞRENME İLE YÜKSEK ÇÖZÜNÜRLÜKLÜ HAVA GÖRÜNTÜLERİNDE YOLLARIN TESPİT EDİLMESİ Figen ÖNÜN Yüksek Lisans Tezi*.
- Anderson, D., & McNeill, G. (1992). Artificial neural networks technology. *Kaman Sciences Corporation*, 258(6), 1-83.
- ARIKAN, S., & LATİFOĞLU, F. Gri Kurt Optimizasyon Algoritması ile Kosinüs Modüleli Süzgeç Bankası Tasarımı. *Avrupa Bilim ve Teknoloji Dergisi*, (26), 472-476.
- Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive cases study of India. *Chaos, Solitons & Fractals*, 139, 110017.
- ArunKumar, K. E., Kalaga, D. V., Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2021). Forecasting of COVID-19 using deep layer recurrent neural networks (RNNs) with gated recurrent units (GRUs) and long short-term memory (LSTM) cells. *Chaos, Solitons & Fractals*, 146, 110861.
- Ataseven, B. (2013). Yapay sinir ağları ile öngörü modellemesi. *Öneri Dergisi*, 10(39), 101-115.
- AYIK, Y. Z., Özdemir, A., & Yavuz, U. (2007). LİSE TÜRÜ VE LİSE MEZUNİYET BAŞARISININ, KAZANILAN FAKÜLTE İLE İLİŞKİSİNİN VERİ MADENCİLİĞİ TEKNİĞİ İLE ANALİZİ. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(2), 441-454.

- Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J. M., Moosaei, H., ... & Mosavi, A. (2021). Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. *Results in Physics*, 27, 104495.
- Bahri, S., Kdayem, M., & Zoghlami, N. (2020a). Deep Learning for COVID-19 prediction. *Proceedings of the International Conference on Advanced Systems and Emergent Technologies, IC_ASET 2020*, 406–411. Institute of Electrical and Electronics Engineers Inc.
- Bahri, S., Kdayem, M., & Zoghlami, N. (2020b). Deep Learning for COVID-19 prediction. *Proceedings of the International Conference on Advanced Systems and Emergent Technologies, IC_ASET 2020*, 406–411. Institute of Electrical and Electronics Engineers Inc.
- Bastos, S. B., & Cajueiro, D. O. (2020). Modeling and forecasting the early evolution of the COVID-19 pandemic in Brazil. *Scientific Reports*, 10(1), 1-10.
- Baykal, A. (2006). VERİ MADENCİLİĞİ UYGULAMA ALANLARI Application Fields of Data Mining. *Ziya Gökalp Eğitim Fakültesi Dergisi*, 7, 95–107.
- Betül, Y. M., Karakuş, A. Y., Tezi, D., Mühendisliği, B., Dalı, A., Doç, D. :, & Aydın, G. (2017). *DERİN ÖĞRENME VE BÜYÜK VERİ YAKLAŞIMLARI İLE METİN ANALİZİ*.
- Bilgiler, G., Tani, V. E., Danışma, B., & Çalışması, K. (n.d.). *COVID-19 (SARS-CoV-2 ENFEKSİYONU)*.
- Boccaletti, S., Ditto, W., Mindlin, G., & Atangana, A. (2020). Modeling and forecasting of epidemic spreading: The cases of COVID-19 and beyond. *Chaos, solitons, and fractals*, 135, 109794.
- CANBAY, Y., İSMETOĞLU, A., & CANBAY, P. (2021). COVID-19 HASTALIĞININ TEŞHİSİNDE DERİN ÖĞRENME VE VERİ MAHREMİYETİ. *Mühendislik Bilimleri ve Tasarım Dergisi*, 9(2), 701-715.
- Çelik, M. (2009). Veri madenciliğinde kullanılan sınıflandırma yöntemleri ve bir uygulama. *Yayımlanmamış Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul*.
- Chandra, R., Jain, A., & Singh Chauhan, D. (2022). Deep learning via LSTM models for COVID-19 infection forecasting in India. *PloS one*, 17(1), e0262708.
- Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, 92(4), 418–423. <https://doi.org/10.1002/jmv.25681>

- Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, 135, 109864.
- ÇİLAN, Ç. A., & Sultan, K. U. Z. U. (2013). KİŞİSEL E-TİCARET UYGULAMALARININ KATEGORİK VERİ ANALİZİ YÖNTEMLERİ İLE DEĞERLENDİRİLMESİ. *Alphanumeric Journal*, 1(1), 27-32.
- COVID-19. (n.d.-a). Retrieved December 7, 2021, from <https://covid19.saglik.gov.tr/>
- COVID-19. (n.d.-b). Retrieved December 8, 2021, from <https://covid19.saglik.gov.tr/>
- data mining Origins and early applications Modeling and data-mining approaches Model creation Data-mining techniques.* (n.d.). Retrieved from <https://web.archive.org/web/20150908002953/http://www.britannica.com/print/article/10561501/3TheWaybackMachinehttps://web.archive.org/web/20150908002953/http://www.britannica.com/print/article/1056150>
- Devaraj, J., Elavarasan, R. M., Pugazhendhi, R., Shafiullah, G. M., Ganesan, S., Jeysree, A. K., ... & Hossain, E. (2021). Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant ?. *Results in Physics*, 21, 103817.
- Dey, S. K., Rahman, M. M., Siddiqi, U. R., & Howlader, A. (2020). Analyzing the epidemiological outbreak of COVID-19: A visual exploratory data analysis approach. *Journal of Medical Virology*, 92(6), 632–638. <https://doi.org/10.1002/jmv.25743>
- DOĞAN, F., & TÜRKOĞLU, İ. (2019). Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme. *DÜMF Mühendislik Dergisi*, 10(2), 409–445. <https://doi.org/10.24012/dumf.411130>
- Elsheikh, A. H., Saba, A. I., Abd Elaziz, M., Lu, S., Shanmugan, S., Muthuramalingam, T., ... & Shehabeldeen, T. A. (2021). Deep learning-based forecasting model for COVID-19 outbreak in Saudi Arabia. *Process Safety and Environmental Protection*, 149, 223-233.
- Er, M. B., & İbrahim, I. Ş. I. K. (2021). LSTM tabanlı derin ağlar kullanılarak diyabet hastalığı tahmini. *Türk Doğa ve Fen Dergisi*, 10(1), 68-74.
- Eroğlu, Y., & Kelimeler Öz, A. (2020). FORECASTING MODELS FOR COVID-19 CASES OF TURKEY USING ARTIFICIAL NEURAL NETWORKS AND DEEP LEARNING YAPAY SİNİR AĞLARI VE DERİN ÖĞRENME KULLANARAK TÜRKİYE’NİN COVID-19 VAKALARI İÇİN TAHMİN MODELLERİ. In *Journal of Industrial Engineering* (Vol. 31). Retrieved from <https://orcid.org/0000-0002-8354-6783>

- Ferdi DOĞAN, İ. T. (n.d.). *Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme*.
- fırat BAYIR. (n.d.). *yapay Sinir Ağları ve Tahmin Modellemesi Üzerine Bir Uygulama Fırat BAYIR*.
- Fikret, E. R., & Cengiz, B. A. L. (2020). COVID-19 ve Grafiksel Veri Analizi. *Osmangazi Tıp Dergisi*, 42(4), 450-461.
- Fonseca i Casas, P., García i Carrasco, V., & Garcia i Subirana, J. (2020). SEIRD COVID-19 formal characterization and model comparison validation. *Applied Sciences*, 10(15), 5162.
- Furkan Keskenler, M., & Fahri Keskenler, E. (2017). *Geçmişten Günümüze Yapay Sinir Ağları ve Tarihçesi*. (5), 8–18.
- Fu, R., Zhang, Z., & Li, L. (2016, November). Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)* (pp. 324-328). IEEE.
- Gaeta, G. (2020). *Data analysis for the COVID-19 early dynamics in Northern Italy*. Retrieved from <http://arxiv.org/abs/2003.02062>
- Giangreco, G. (2020). cases fatality rate analysis of Italian COVID-19 outbreak. *Journal of Medical Virology*, 92(7), 919–923. <https://doi.org/10.1002/jmv.25894>
- Gomez-Cravioto, D. A., Diaz-Ramos, R. E., Cantu-Ortiz, F. J., & Ceballos, H. G. (2021). Data Analysis and Forecasting of the COVID-19 Spread: A Comparison of Recurrent Neural Networks and Time Series Models. *Cognitive Computation*. <https://doi.org/10.1007/s12559-021-09885-y>
- Gri Kurt Algoritması Hakkında Bilgi. (n.d.). Retrieved March 17, 2022, from <https://birbakmali.blogspot.com/2019/12/gri-kurt-algoritmas-hakknda-bilgi-slayt.html?m=1>
- Guo, Q., & He, Z. (2021). Prediction of the confirmed cases and deaths of global COVID-19 using artificial intelligence. *Environmental Science and Pollution Research*, 28(9), 11672-11682.
- IEEE ITSS, & Institute of Electrical and Electronics Engineers. (n.d.). *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) : 1-4 Nov. 2016*.
- Irmak, S., Köksal, C. D., & Asilkan, Ö. (2012). Hastanelerin gelecekteki hasta yoğunluklarının veri madenciliği yöntemleri ile tahmin edilmesi. *Uluslararası Alanya İşletme Fakültesi Dergisi*, 4(1), 101-114.

- Istaiteh, O., Owais, T., Al-Madi, N., & Abu-Soud, S. (2020). Machine Learning Approaches for COVID-19 Forecasting. *2020 International Conference on Intelligent Data Science Technologies and Applications, IDSTA 2020*, 50–57. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/IDSTA50958.2020.9264101>
- Javor, D., Kaplan, H., Kaplan, A., Puchner, S. B., Krestan, C., & Baltzer, P. (2020). Deep learning analysis provides accurate COVID-19 diagnosis on chest computed tomography. *European journal of radiology*, 133, 109402.
- Karakoyun, M. (2021). Çok amaçlı optimizasyon problemlerinin çözümü için Kurbağa Sıçrama ve Gri Kurt Optimizasyonu algoritmaları tabanlı hibrit bir yöntemin geliştirilmesi.
- Karypis, G., Han, E. H., & Kumar, V. (1999). Chameleon: Hierarchical clustering using dynamic modeling. *Computer*, 32(8), 68-75.
- Khare, N., Jha, M., Mathur, R., & Jha, A. K. (2020). Data analysis for COVID-19. *The International journal of analytical and experimental modal analysis*, 12(5), 960-974.
- KILINÇ, H. Ç., & ÖZTÜRK, Y. Hibrit Gri Kurt Optimizasyonu ile Kapılı Tekrarlayan Birim Modeli Kullanılarak Günlük Akım Tahmini. *Avrupa Bilim ve Teknoloji Dergisi*, (35), 259-267.
- Kim, M. H., Kim, J. H., Lee, K., & Gim, G. Y. (2021, January 1). The Prediction of COVID-19 Using LSTM Algorithms. *International Journal of Networked and Distributed Computing*, Vol. 9, pp. 59–74. Atlantis Press. <https://doi.org/10.2991/IJNDC.K.201218.003>
- Koronavirüs (COVID-19) - Google Haberler. (n.d.). Retrieved December 25, 2021, from <https://news.google.com/covid19/map?hl=tr&gl=TR&ceid=TR%3Atr>
- Kutlu, B., & Badur, B. (2009). Yapay Sinir Ağları İle Borsa Endeksi Tahmini. *Yönetim-İstanbul Üniversitesi İşletme Fakültesi İşletme İktisadi Enstitüsü Dergisi*, 20(63).
- Lalmuanawma, S., Hussain, J., & Chhakhuak, L. (2020, October 1). Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons and Fractals*, Vol. 139. Elsevier Ltd. <https://doi.org/10.1016/j.chaos.2020.110059>
- Lu, H., Stratton, C. W., & Tang, Y. W. (2020). Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *Journal of Medical Virology*, Vol. 92, pp. 401–402. <https://doi.org/10.1002/jmv.25678>

- Makine öğrenmesi kullanılarak kestirimci bakım. (2018). *OLCAY ÖZTANIRPROF, DR. ALİ ZİYA ALKAR*.
- Marzouk, M., Elshaboury, N., Abdel-Latif, A., & Azab, S. (2021a). Deep learning model for forecasting COVID-19 outbreak in Egypt. *Process Safety and Environmental Protection*, 153, 363–375. <https://doi.org/10.1016/j.psep.2021.07.034>
- Marzouk, M., Elshaboury, N., Abdel-Latif, A., & Azab, S. (2021b). Deep learning model for forecasting COVID-19 outbreak in Egypt. *Process Safety and Environmental Protection*, 153, 363–375. <https://doi.org/10.1016/j.psep.2021.07.034>
- Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey Wolf Optimizer. *Advances in Engineering Software*, 69, 46–61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>
- Mohamadou, Y., Halidou, A., & Kapen, P. T. (2020). A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. *Applied Intelligence*, 50(11), 3913–3925. <https://doi.org/10.1007/s10489-020-01770-9>
- Niazkar, H. R., & Niazkar, M. (2020a). Application of artificial neural networks to predict the COVID-19 outbreak. *Global Health Research and Policy*, 5(1). <https://doi.org/10.1186/s41256-020-00175-y>
- Niazkar, H. R., & Niazkar, M. (2020b). Application of artificial neural networks to predict the COVID-19 outbreak. *Global Health Research and Policy*, 5(1). <https://doi.org/10.1186/s41256-020-00175-y>
- Niwa, M., Hara, Y., Sengoku, S., & Kodama, K. (2020). Effectiveness of social measures against COVID-19 outbreaks in selected Japanese regions analyzed by system dynamic modeling. *International Journal of Environmental Research and Public Health*, 17(17), 1–12. <https://doi.org/10.3390/ijerph17176238>
- Oğuzlar, A. (2003). *VERİ ÖN İŞLEME*. (21), 67–76.
- Özekes, S. (2003). Veri madenciliği modelleri ve uygulama alanları.
- Öztürk, K., & Şahin, M. E. (2018). Yapay sinir ağları ve yapay zekâ'ya genel bir bakış. *Takvim-i Vekayi*, 6(2), 25-36.
- Peyrot, M. (1996). Causal Analysis: Theory and Application. *Journal of Pediatric Psychology*, 21(1), 3–24. <https://doi.org/10.1093/JPEPSY/21.1.3>
- Picariello, M., & Aliani, P. (2020). COVID-19: Data analysis of the Lombardy region and the provinces of Bergamo and Brescia. *arXiv preprint arXiv:2003.10518*.

- Ramesh, A. N., Kambhampati, C., Monson, J. R. T., & Drew, P. J. (2004, September). Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*, Vol. 86, pp. 334–338. <https://doi.org/10.1308/147870804290>
- Salgınlar listesi - Vikipedi. (n.d.). Retrieved December 7, 2021, from https://tr.wikipedia.org/wiki/Salg%C4%B1nlar_listesi
- Sazli, M. H. (2006). A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, 50(01).
- Sengupta, S. (2020). Forecasting the peak of COVID-19 daily cases in india using time series analysis and multivariate lstm. *Technical report*.
- Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons and Fractals*, 140. <https://doi.org/10.1016/j.chaos.2020.110212>
- Summary, S. (2020). *2019 Novel Coronavirus (2019-nCoV), Wuhan , China*.
- Sydenham, P. H., & Thorn, R. (2005). *Handbook of measuring system design*. Wiley.
- Tandon, H., Ranjan, P., Chakraborty, T., & Suhag, V. (2020). Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future. *arXiv preprint arXiv:2004.07859*.
- Tapiwa, G., Cécile, K., Dongxuan, C., Andrea, T., Christel, F., Jacco, W., & Niel, H. (2020). Estimating the generation interval for COVID-19 based on symptom onset data. *MedRxiv*. <https://doi.org/10.1101/2020.03.05.20031815>
- The Six Types of Data Analysis. Data Science is about asking questions... | by Benedict Neo | Towards Data Science. (n.d.). Retrieved February 7, 2022, from <https://towardsdatascience.com/the-six-types-of-data-analysis-75517ba7ea61>
- Tolon, M., & TOSUNOĞLU, N. G. (2008). Tüketici tatmini verilerinin analizi: yapay sinir ağları ve regresyon analizi karşılaştırması. *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 10(2), 247-259.
- Tüzüntürk, S. (2010). Veri madenciliği ve istatistik. *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 29(1), 65-90.
- Types of Data Analysis | Tutorial by Chartio. (n.d.). Retrieved February 7, 2022, from <https://chartio.com/learn/data-analytics/types-of-data-analysis/#diagnostic-analysis>
- ULUYARDIMCI, M. M., & ZONTUL, M. Veri Madenciliği Yöntemleri ile Uçuş Biletleme Analizi. *AURUM Journal of Engineering Systems and Architecture*, 3(2), 153-168.

- UZLU, E. (2019). Türkiye için gri kurt optimizasyon algoritması ile yapay sinir ağılarını kullanarak enerji tüketiminin tahmini. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 7(2), 245–262. <https://doi.org/10.29109/gujsc.519553>
- Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 14(4), 337–339. <https://doi.org/10.1016/j.dsx.2020.04.012>
- Veri Analizi Nedir? - Etkili Veri Analizi Teknikleri - Turhost. (n.d.). Retrieved February 8, 2022, from <https://www.turhost.com/blog/veri-analizi-nedir-nasil-yapilir/#serp>
- Veri Analizi Nedir? Nasıl Yapılır? - Can Cankıran. (n.d.). Retrieved February 8, 2022, from <https://www.cancankiran.com/veri-analizi-nedir-nasil-yapilir/>
- Veri Analizi Nedir? | Önemi | Modelleri | Yararları | Türleri ve Çeşitleri. (n.d.). Retrieved February 7, 2022, from <https://proente.com/veri-analizi-nedir/>
- Veri Analizi Süreci ve Veri Analizi Yöntemleri Nedir? | GTech. (n.d.). Retrieved February 7, 2022, from <https://www.gtech.com.tr/veri-analizi-ve-etkili-veri-analiz-yontemleri/>
- Veri Analizleri Nedir? (n.d.). Retrieved February 7, 2022, from <https://www.intel.com.tr/content/www/tr/tr/analytics/what-is-data-analytics.html>
- Veri Bilimi ve Yapa Zeka Yazı Serisi | by Deniz KILINÇ | Medium. (n.d.). Retrieved February 7, 2022, from <https://medium.com/@denizkilinc/veri-bilimi-ve-yapa-zeka-yaz%C4%B1-serisi-902ef59fdfab>
- Wang, P., Zheng, X., Ai, G., Liu, D., & Zhu, B. (2020). Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos, Solitons and Fractals*, 140. <https://doi.org/10.1016/j.chaos.2020.110214>
- Waqas, M., Farooq, M., Ahmad, R., & Ahmad, A. (2020). *Analysis and Prediction of COVID-19 Pandemic in Pakistan using Time-dependent SIR Model*. Retrieved from <http://arxiv.org/abs/2005.02353>
- What Are the Different Types of Data Analysis? (n.d.-a). Retrieved February 7, 2022, from <https://careerfoundry.com/en/blog/data-analytics/different-types-of-data-analysis/>
- What Are the Different Types of Data Analysis? (n.d.-b). Retrieved February 8, 2022, from <https://careerfoundry.com/en/blog/data-analytics/different-types-of-data-analysis/>
- What Is Data Analysis? Methods, Techniques, Types & How-To. (n.d.). Retrieved February 7, 2022, from <https://www.datapine.com/blog/data-analysis-methods-and-techniques/>

- What is Data Analysis? Research | Types | Methods | Techniques. (n.d.). Retrieved February 7, 2022, from <https://www.guru99.com/what-is-data-analysis.html>
- What is Data Analysis? Types, Methods and Techniques [2022 Edition] | Simplilearn. (n.d.). Retrieved February 7, 2022, from <https://www.simplilearn.com/data-analysis-methods-process-types-article>
- WHO/Europe | Coronavirus disease (COVID-19) outbreak - About the virus. (n.d.). Retrieved December 7, 2021, from <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov>
- Yadav, N., Yadav, A., & Kumar, M. (2015). *An introduction to neural network methods for differential equations* (Vol. 1, p. 114). Berlin: Springer.
- Yapay Sinir Ağları - Derin Öğrenme | Deep Learning. (n.d.). Retrieved March 3, 2022, from <https://www.derinogrenme.com/2017/03/04/yapay-sinir-aglari/>
- Yapay Sinir Ağları ve Uygulamaları | GTech. (n.d.). Retrieved March 3, 2022, from <https://www.gtech.com.tr/yapay-sinir-aglari-ve-uygulamalari-1/>
- Yapay Zeka Teknikleri ve Yapay Sinir Ağları Kullanılarak web sayfalarının sınıflandırılması.* (n.d.).
- Yin, Y., & Wunderink, R. G. (2018). MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology*, Vol. 23, pp. 130–137. <https://doi.org/10.1111/resp.13196>
- Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons and Fractals*, 140. <https://doi.org/10.1016/j.chaos.2020.110121>
- Zhao, S., Gao, D., Zhuang, Z., Chong, M. K. C., Cai, Y., Ran, J., ... Wang, M. H. (2020). Estimating the Serial Interval of the Novel Coronavirus Disease (COVID-19): A Statistical Analysis Using the Public Data in Hong Kong From January 16 to February 15, 2020. *Frontiers in Physics*, 8. <https://doi.org/10.3389/fphy.2020.00347>
- Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... Wang, M. H. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases*, 92, 214–217. <https://doi.org/10.1016/j.ijid.2020.01.050>
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., ... Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, 382(8), 727–733. <https://doi.org/10.1056/nejmoa2001017>

Zontul, M., & YANGIN, A. (2017). Yapay Sinir Ağı Teknikleri Kullanarak Eğitim Yayıncılığı Sektöründe Veri Madenciliği= Data Mining On Education Publishing Sector By Artificial Neural Network Tecniques. *AURUM Journal of Engineering Systems and Architecture*, 1(2), 1-15.

