Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

PREDICTIVE HEALTHCARE ANALYSIS OF PAKISTAN'S COVID-19 PANDEMIC USING DATA MINING AND TIME SERIES MODELLING

Abdul Karim Kazi [1], Maria Andleeb [2], Saad Ahmed [3], Raheela Asif [4] and Syed Muhammad Nabeel Mustafa [5]

[1,2] Department of Computer Science and I.T., NED University of Engineering and Technology, Pakistan
[3] Department of Computer Science, IQRA University, Pakistan
[4] Department of Software Engineering, NED University of Engineering and Technology, Pakistan

[*]Corresponding author: karimkazi@neduet.edu.pk,

Abstract: **The novel coronavirus known as COVID-19 has become widespread throughout the world and presented new problems to the scientific community. This resulted in severe measures being implemented in several impacted countries, including total lockdowns, trade and business closures, and travel restrictions, all of which had a major negative economic impact. Pakistan has also had five coronavirus waves. Thus, government officials, legislators, business associates, and entrepreneurs will place a high value on knowing and anticipating how a nation might stop the spread of COVID-19. We use AI-based forecasting models, such as time series ARIMA, LSTM, FB Prophet, and VAR, to predict the spread of the COVID-19 pandemic. These techniques support the decisions made by legislators and public health authorities about the war against the epidemic. This paper demonstrates the promising potential of the time series model in forecasting COVID-19 cases and highlights the superior performance of the time series compared to the LSTM.**
Keywords: **Data Mining; Random Forest; Time series; ARIMA; VAR; FB Prophet; LSTM; COVID-19**

I. INTRODUCTION

A complex AI technique called data mining is used to extract novel, useful, and accurate hidden patterns or knowledge from datasets. Time series forecasting is a technique used to predict future values by analyzing past data collected over time. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models are neural COVID-19 types that can effectively handle time series data. Machine learning (ML), a broader domain, encompasses various methods and algorithms, including RNNs and LSTMs, which are employed for forecasting purposes. Facebook created a unique open-source program called Prophet to anticipate time series data. It can handle missing data, seasonality, and other complications and employs a combination of linear and non-linear models to create Forecasting. A statistical model called VAR (Vector Auto Regression) is used to examine the dynamic connection between time series.

The primary aim of research on COVID-19 forecasting utilizing various forecasting approaches is to foresee the virus's future spread and effects. This can help with emergency planning, resource allocation, and public health policy. To examine historical data and forecast future trends in cases, fatalities, and other pertinent metrics, techniques including time series analysis, recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and vector auto-regression (VAR) models can be utilized. To prevent the spread of the virus and lessen its effects on communities, policymakers, and public health authorities can use these projections to help them make more knowledgeable decisions

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

[3]. To find trends and patterns in all the occurrences related to infectious diseases, the majority of the research in this paper makes use of AI forecasting models. This paper will also examine the usage of such methods that are solely concerned with predicting epidemiological variables like cumulative cases, fatalities, and recoveries from the current COVID-19 pandemic.

II. RELATED WORK.

COVID-19- Since January 2020, the number of incidents in the US has increased by 19. Even after the limitations were loosened, the number of instances increased despite social exclusion and lockdowns. Modeling the disease's spread can help governments and healthcare professionals plan ahead and get funding. For the healthcare system, precise short-term case forecasts are essential. Simple Moving Average, Exponentially Weighted Moving Average, Holt-Winters Double Exponential Smoothing Additive, ARIMA, and SARIMA are just a few of the models that have been used since the pandemic's start. In this paper, ARIMA and SARIMA were selected for prediction, and the optimal model parameters for each were found using a grid search. According to the findings, ARIMA performs better than SARIMA in predicting, whereas the Holt-Winters Double Exponential model surpasses the Exponentially Weighted Moving Average and Simple Moving Average [4]. This paper aimed to assess how well the ARIMA model predicted the spread of COVID-19, which the World Health Organization designated a global pandemic in March 2020 after infecting over 4 million people and killing over 300,000 by early May 2020. Although it is believed that ARIMA is not appropriate for complex and dynamic environments, the study evaluated Kuwait as a case study to assess its accuracy over a considerable amount of time. The actual statistics were mainly within the ranges predicted by the selected ARIMA model at a 95% confidence level, despite the disease's unpredictability and changes made by the Kuwaiti government. With a Pearson correlation coefficient of 0.996, the predicted values and observed data showed a significant correlation. This indicates that the predictions made by the ARIMA model are appropriate and sufficiently accurate [5]. A Study was presented to detect COVID-19 misinformation in Swahili language in tweets. A machine learning model was used to carry out this study with highest accuracy was achieved using SVM i.e. 83.67% [6]. A comparison was presented in [7] which addressed the COVID-19 issues using Machine learning, deep learning, and Artificial intelligence techniques. Regarding the COVID-19 pandemic, this work discovered contemporary and up-to-date information to fight against the COVID-19 pandemic using ML, DL, and AI techniques. A survey was conducted in [8] which explores the several deep learning applications in natural language processing for COVID-19. It also presented some limitations i.e. Interpretability, Learning from Limited Labeled Data, Generalization Metrics, and Data Privacy. The study suggested that in the spread forecasting of epidemiology, deep learning has additionally been utilized. Based on the comparative study of multiple research papers, the comprehensive study in [9], explores the numerous data mining algorithms that are used in combination with epidemiological prediction models. The specific risk of COVID-19 using LSTM-based ANN guided by Bayesian optimization was presented in [10]. A study [11] proposed an ML algorithm that accurately predict the mortality risk of COVID-19 patients.

Millions of people are affected by the most current coronavirus outbreak (COVID-19), which has expanded widely and caused severe diseases. Modeling was done using data acquired between January 30 and April 26, 2020, while forecasting was done using data received between April 27 and May 11, 2020. The spatial distribution of illness 29 risk was examined on a GIS platform using weighted overlay analysis. [12] The paper analyses the COVID-19 situation in Pakistan, which is presently dealing with the virus' fourth wave. The paper examines COVID-19 data for the nation using epidemiological models. The article evaluates both Bayesian and time-series SIR (tSIR) techniques while considering the fundamental susceptible (SIR) model infected (SIR) model, and recovered (SIR) model. Due to the government's successful strategy, the paper also discovered that the global assumption of a 14-day incubation time is inappropriate for Pakistan's data and that COVID-19 was not a pandemic. According to the study, the posterior-based SIR (pSIR) model with a 34-uniform prior for R0 and Poisson distribution yields superior outcomes. The reporting rate (ρ) is less than 1, indicating underreporting of instances, according to the time-series SIR (tSIR) study [13]. To diagnose and detect the COVID-19 pandemic, the overview of state of art applications and algorithms was presented in [14]. The authors have proposed a Fake News Encoder Classifier (FNEC) [15] for online published news

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

related to COVID-19 vaccines. It uses an ELECTRA model to classify news articles into real or fake and creates a new dataset called COVAX-Reality for evaluation. Estimation of the strengths of negative and positive sentiments was carried out in [16]. The text analytics of Twitter data using tweets, retweets, and hashtags were used to detect the strength.

III. METHODOLOGY

This research utilizes Facebook Prophet, VAR, ARIMA, and LSTM models to predict upcoming COVID-19 cases and deaths. Models are created for each approach using the COVID-19 dataset, and their performance is assessed by comparing predictions via graphs and performance rates on a country-specific level.

A. Design Experiment

Before conducting experiments, several processes are completed, such as model building, data exploration, and data preparation. Data transformation includes data aggregation, extrapolation, creation of dummies, data time transformation, and variable reduction. Data combining entails merging or combining datasets. The novelty in the next section lies in the comprehensive mathematical explanations of both Facebook Prophet and LSTM, making it a valuable resource for readers seeking a deep understanding of these forecasting techniques. Additionally, it provides practical implementation guidance for Facebook Prophet, enhancing its usability in real-world scenarios.

B. Exploratory Data Analysis

Exploratory data analysis (EDA), inspects data to identify its unique characteristics. As part of EDA, data summaries and visualizations are carried out, and interesting data points are considered. Before any model is fitted to the data, exploratory data analysis should always be performed. The flow diagram of the EDA of our research work is shown in Figure 1.

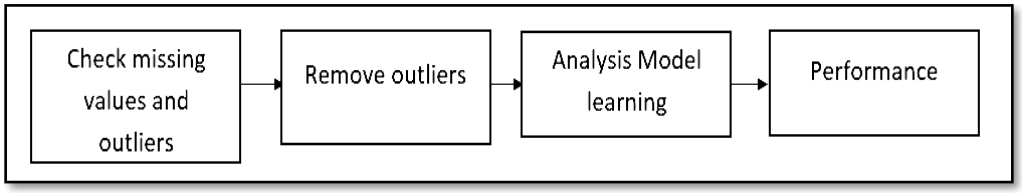

Figure 1: Data Preprocessing Flow Diagram.

Data cleaning is the process of identifying parts of the data that are missing, inaccurate, or irrelevant and then removing, recreating, or cleaning up the damaged or impure data. After cleaning, a dataset should be consistent with all other connected datasets in operation. The principal reasons for the disparities discovered or eliminated may have been user-input errors, data corruption during storage or transmission, or various data dictionary definitions of the same entity in other repositories.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

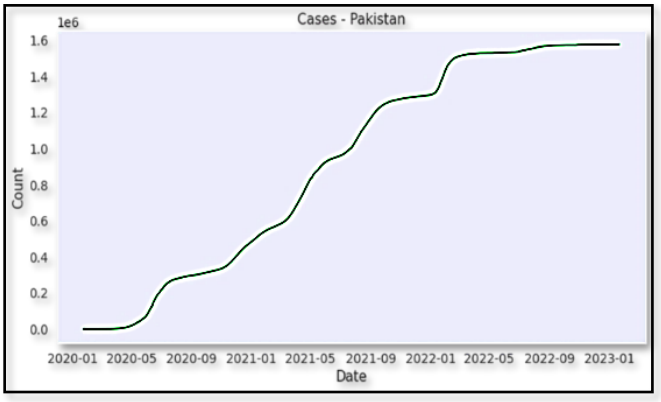

Figure 2: Frequency Plot of Confirmed COVID-19 cases per day

C. Quantities Variables

The frequency plot of confirmed cases and deaths in Pakistan since 2020 is broken down below into before and after cleaning statistics. Pakistan's COVID-19 confirmed cases curve shows the trend of the number of confirmed cases over time. Typically, the curve rises slowly from a low point to a peak before leveling off or falling. The speed of virus propagation, the success of containment efforts, and the accessibility of testing are some variables that affect the curve's shape. The trend of the number of deaths attributable to the virus over time is depicted by the COVID-19 death curve for Pakistan, as shown in Figure 2.

D. Calculation Of Quantitative Variables

The quantitative numerical features that have a significant role in data exploration and analysis of the pattern of confirmed cases and deaths in time series are statistically represented. The plot shows the skewness vs. density of calculated numerical features according to the presence of outliers and the distribution for the study of the time series dataset. The following variables are calculated. These metrics are essential in the analysis of COVID-19 because they provide important insights into the current state and future trajectory of the pandemic.

E. Incremental Cases And Deaths

Incremental cases in the context of COVID-19 refer to the number of new cases reported in a given period. This value can be calculated by subtracting the number of cases reported in the previous period from those registered in the current period. For example, if there were 100 cases reported on Monday and 120 cases reported on Tuesday, then the incremental cases for Tuesday would be 20 (120 - 100). This value represents the increase in the number of cases from one day to the next. It is used to forecast future trends in the spread of the disease. In Figure 3, we can observe the Incremental cases in Pakistan throughout the pandemic.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

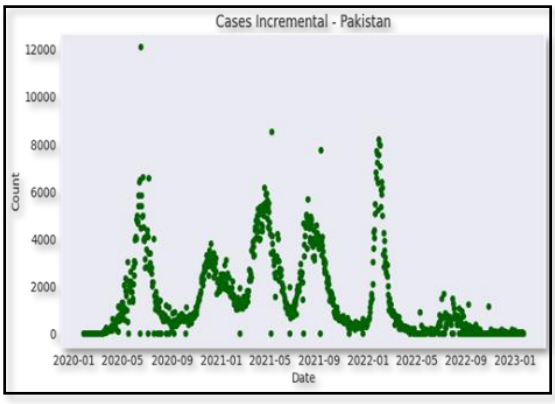

Figure 3: Frequency Plot of C19 Incremental cases per day

Incremental deaths of COVID-19 refer to the number of new casualties caused by the COVID-19 virus that have occurred in a given period. This value can be calculated by subtracting the number of COVID-19 deaths reported in the previous time period from the number of COVID-19 deaths reported in the current time period. This value represents the increase in the number of deaths caused by the COVID-19 virus from one day to the next. It is used to forecast future trends in the spread of the disease.

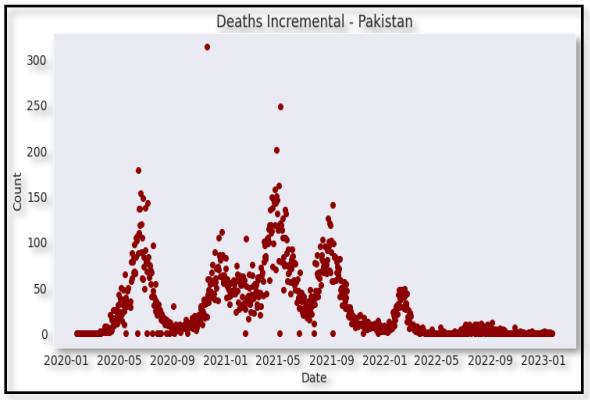

Figure 4: Frequency Plot of Incremental C19 deaths per day

In Figure 4, we can observe the Incremental deaths in Pakistan throughout the pandemic. The incremental deaths of COVID-19 in Pakistan refer to the number of new deaths reported in specifically monitoring the impact of mitigation measures and guiding public health interventions.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

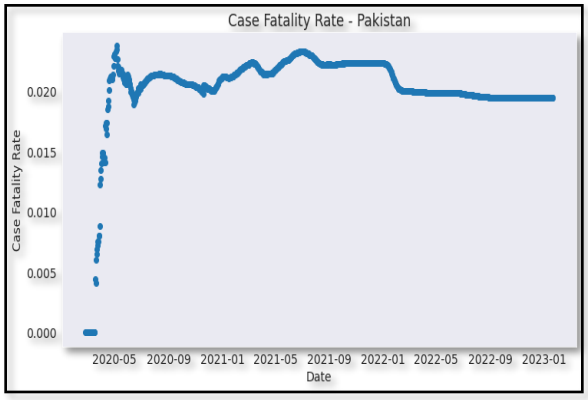

Figure 5: Frequency Plot of Case Fatality Rate per day in Pakistan

F. Case Fatality Rate

The number of COVID-19-related deaths divided by the total number of confirmed cases yields the Case Fatality Rate (CFR), which indicates the severity of the illness. The CFR is expressed as a percentage and indicates the risk of death for someone contracting the disease. A lower CFR means that the disease is less severe, while a higher CFR implies that the disease is more severe. It is important to note that the CFR can be influenced by many factors, such as the accuracy of case reporting, the timing and quality of medical intervention, and the age and underlying health conditions of those infected. In the context of COVID-19, the CFR has significantly varied depending on the location and the stage of the pandemic. In Figure 5 and Figure 6, we can observe the Case Fatality Rate in Pakistan throughout the pandemic.

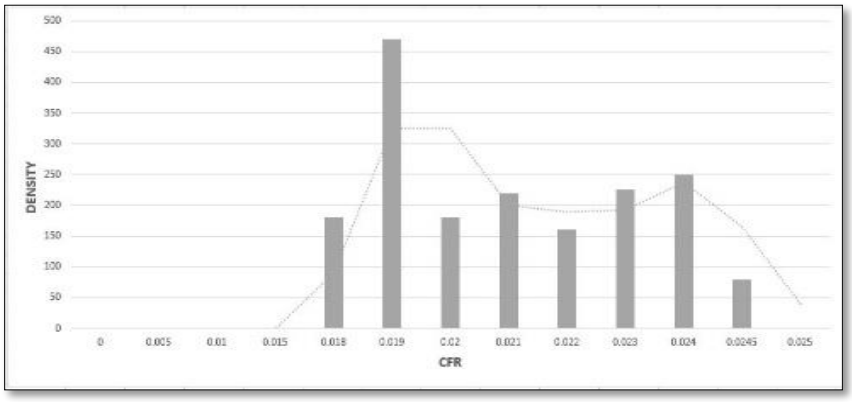

Figure 6: Distribution Plot of Case Fatality Rate per day in Pakistan

The number of COVID-19 deaths in Pakistan divided by the total number of confirmed cases is known as the COVID-19 Case Fatality Rate, or CFR. Pakistan's CFR is estimated to be 1.5% as of January 2023. This indicates that 1.5% of the population in the nation dies from COVID-19 for every 100 confirmed cases of the virus.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

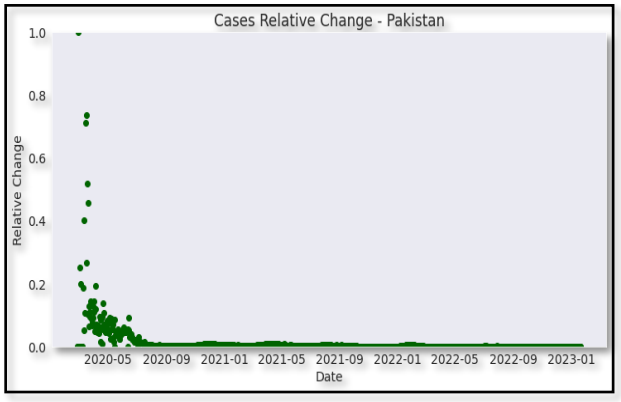

Figure 7: Frequency Plot of the day-to-day relative change in cases in Pakistan

G. Day-To-Day Relative Change In Cases And Deaths

Day-to-day relative change in COVID-19 cases refers to the percentage increase or decrease in the number of confirmed COVID-19 cases from one day to the next. It is calculated as the difference between the number of cases on two consecutive days divided by the number of cases on the previous day.

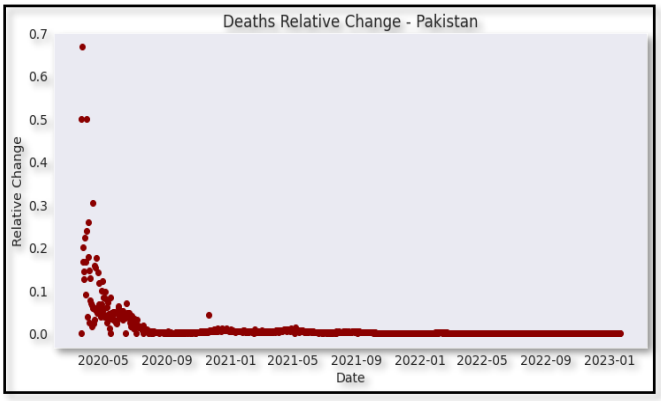

Figure 8: Frequency Plot of the day-to-day relative change in deaths in Pakistan

A0 higher number of cases is indicated by a positive value, whereas a lower number of cases is indicated by a negative value. This metric provides insight into the trend of the spread of the disease Day-to-day relative change in COVID-19 deaths refers to the percentage increase or decrease in the number of COVID-19 deaths from one day to the next. A positive value in Figures 7 and 8 denotes an increase in deaths, whereas a negative value denotes a decrease in deaths.

H, Comparison of cases vs. Deaths curve

The comparison between the COVID-19 cases and deaths curve refers to the graphical representation of the number of confirmed COVID-19 cases and the number of COVID-19-related deaths over time. This comparison provides a visual representation of the spread of the disease and the impact of the disease on human health, as shown in Figure 9.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

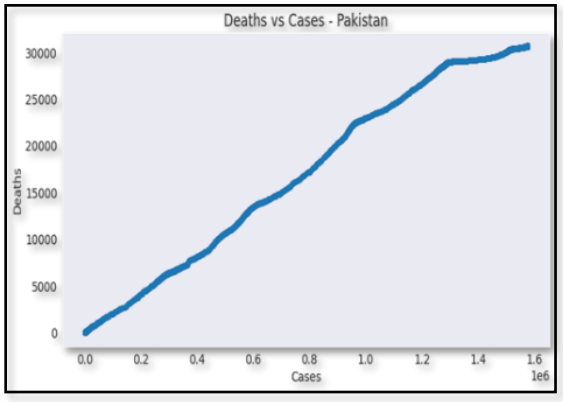

Figure 9: Frequency and Distribution Plot of day-to-day relative change in deaths in Pakistan

I. Correlation Analysis

Correlation analysis is a statistical method for determining whether and how strongly two variables or datasets are related. In market research, correlation analysis is a tool used to look for any significant patterns, trends, or relationships between quantitative data collected through methods such as surveys and polls. Finding trends in datasets is the main application of correlation analysis. Spearman's correlation is used when the linear relationship between two continuous variables is unknown, and the Pearson correlation coefficient evaluates the linear association between the two variables.

J. Correlation between C19 Cases and Deaths in Pakistan and the US

The mathematics of Spearman's correlation (also called Spearman's rank-order correlation coefficient) can be calculated as follows: Convert the data into ranks: Convert the original data into rankings, with the lowest value getting a rank of 1, and so on.

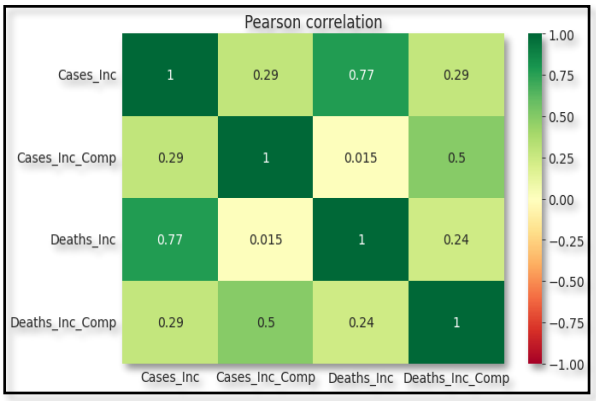

Figure 10: Pearson Correlation of Pak vs. USA C19 Data

Figures 10 and 11 show that it is impossible to state the correlation between COVID-19 cases, deaths, incremental cases, and incremental deaths in Pakistan and the US without thoroughly analyzing the available data.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

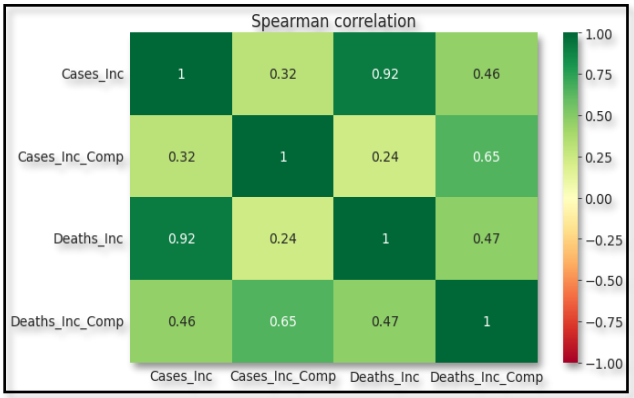

Figure 11: Spearman Correlation of Pak vs. USA C19 Data

K. Data Transformation and Variable Selection

ARIMA models use a combination of three components to model time series data:

1. Auto regression (AR): This element models the relationship between current and past observations.
2. Integration (I): This element models the effect of non-stationarity in the data, such as trends and seasonality.
3. Moving average (MA): This element models the residuals or errors in the time series data.

It is represented by the following equation (1):

$$y_t = c + \phi_1 * y_{\{t-1\}} + \phi_2 * y_{\{t-2\}} + \ldots + \phi_p * y_{\{t-p\}} + e_t \quad (1)$$

Moving average (MA): Moving average models the residuals or errors in the time series data. It is represented by the following equation (2):

$$e_t = \mu + \theta_1 * e_{\{t-1\}} + \theta_2 * e_{\{t-2\}} + \ldots + \theta_q * e_{\{t-q\}} + \varepsilon_t \quad (2)$$

Finally, ARIMA models combine the AR, I, and MA components to form a single equation that models the time series data. The ARIMA model's order is represented by the notation (p, d, q), where p represents the AR component, d denotes the integration order, and q represents the MA component.

L. Implementation of ARIMA

- Data Preprocessing: The first stage in implementing the ARIMA model is to preprocess the time series data. This involves removing any outliers, checking for stationarity, and transforming the data if necessary.
- Model Identification: The next stage is to identify the appropriate ARIMA model for the time series data.
- Model Fitting: Once the appropriate ARIMA model has been identified, the model is fitted to the time series data using a maximum likelihood estimation technique.
- Model Evaluation: The final stage is to evaluate the performance of the ARIMA model by comparing its forecast with the actual values.

M. Vector Auto Regression Var

Vector Auto regression is a statistical model used for multivariate time series forecasting. It models the linear interdependence between multiple time series by estimating the relationships between the variables. The model assumes that each variable is dependent on a set of lagged values of itself and the other variables in the model. It is important to note that VAR modeling is just one of the many time series modeling techniques available. It might only be suitable for some types of data and problem scenarios.

IV. RESULTS AND DISCUSSIONS

Although common sense suggests that the fundamental test of Forecasting is to provide forecasts closer to the actual outcomes; however, how well they can forecast is more complex than it may seem. Suppose a model predicts more accurately than the other does by comparing forecasting with outcomes. In that case, it is simple to compare the two

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

models of interest in this paper. Despite differing qualities, comparisons are feasible, e.g., between expected and actual test values. Anticipated values, generated via machine learning from the same data, allow assessment using metrics like MAE, MSE, and Root Square Score.

Table: 1 Performance of ARIMA

	ARIMA		VAR	
Model	**Cases**	**Deaths**	**Cases**	**Deaths**
MSE	2.71E+10	16207273	1.79E+11	15494119
MSA	156372.8	3914.842	366399.8	3080.086
Variance	-0.08478	-4.17949	-6.15193	-3.95158

The values of each feature's correlation coefficient show a unit change in the dependent component correlated with a unit change in the independent feature. There is a high correlation or strong Forecasting power between the estimate and the response features, and the estimate or coefficient is not zero. Table 1 presents the performance of the ARIMA and VAR, which is performing better than FB Prophet and LSTM, but ARIMA, as well as VAR for Cases and in death forecasting VAR performed well.

The novelty lies in its in-depth analysis and comparison of different time series forecasting models applied to COVID-19 data. It provides insights into the strengths and weaknesses of each model, helping readers make informed decisions when choosing a forecasting approach for similar datasets and scenarios.

V. CONCLUSION

In this paper, we examined COVID-19 data samples, specifically focusing on two variables: confirmed cases and deaths. Our analysis involved preprocessing and exploratory data analysis techniques. We calculated and visualized multiple metrics, including incremental cases and deaths, Case Fatality Rate, daily and 7-day changes, and week-to-week differences. Our objective was to identify the most active time and days for COVID-19 cases, as well as analyze the relationship between cases and deaths through the cases vs. death curve for COVID-19 in Pakistan. These metrics are essential in the analysis of COVID-19 because they provide important insights into the current state and future trajectory of the pandemic. We also plotted comparison graphs of deaths vs. cases compared to Pakistan cases to another country to observe the effects. It is shown that the data analysis process we applied to analyze COVID-19 cases and deaths greatly influences understanding the COVID-19 effects concerning this dataset and is most crucial to add to results. In terms of performance, in the COVID-19 Cases of Pakistan dataset, ARIMA outperforms Deep Learning and other Time series models. It is lower in MAE and MSE, indicating that a model is better and has a lower error rate than other models when the error is smaller. It is -0.08478, which is also better in terms of variance/Root Square Score, meaning that the correlation increases with model quality. The ARIMA outperforms the other models because it performs better in all two metrics. The correlation between the actual and projected variables is quite good with the best-fitted model, and the COVID-19 Cases of the Pakistan dataset using the ARIMA Model performed significantly connected to a better model. Given that it performed better in each of the two metrics, VAR surpasses the other models in forecasting deaths. Forecasting algorithms can be valuable tools in the fight against COVID-19 in Pakistan, as they can help provide projections of the future spread of the virus and inform public health decision-making. In conclusion, forecasting algorithms can be helpful in the fight against COVID-19 in Pakistan. Still, it is vital to use them in conjunction with other data and information and to continuously update and refine the models based on new data and information as they become available.

In summary, the novelty of this research lies in its synthesis of the study's key findings, the exploration of analysis metrics and temporal factors in COVID-19 data, the evaluation of forecasting models, and the discussion of practical implications and future research possibilities in the fight against Covid-19 and other diseases.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 74-84, November 2023
https://doi.org/10.22555/pjets.v11i1.1029
E-ISSN 2224-2333
P-ISSN 2222-9930

REFERENCES

[1] D. J. Cennimo, et al., "Coronavirus Disease 2019 (COVID-19)," Coronavirus Disease 2019 (COVID-19) Management, Nov 10, 2022.
[2] A. HAYES, "What Is a Time Series and How Is It Used to Analyze Data?," Data Discovery Analysis on Complex Time Series Data, pp. 50-51, Jun 12, 2022.
[3] R. Somyanonthanakul, K. Warin, W. Amasiri, et al., "Forecasting COVID-19 cases using time series modeling and association rule mining," BMC Medical Research Methodology, 22, article 281, 2022.
[4] S. Abolmaali, S. Shirzaei, "Forecasting COVID-19 Number of Cases by Implementing ARIMA and SARIMA with Grid Search in United States", MedRxiv, 2021.
[5] H. Alabdulrazzaq, M. N. Alenezi, Y. Rawajfih, B. A. Alghannam, A. A. Al-Hassan, F. S. Al-Anzi, "On the accuracy of ARIMA based prediction of COVID-19 spread," Results in Physics, vol. 27, Aug, 2021.
[6] F.Mlawa, E. Makoba, N.Maduma, "A Machine Learning Model for detecting Covid-19 Misinformation in Swahili Language", Engineering, Technology and Applied Science Research, Vol. 13, No. 3, pp. 10856-10860, 2023.
[7] S.G. Paul et al,"Combating Covid-19 using machine learning and deep learning: Applications, challenges, and future perspectives." Array (New York, N.Y.) vol. 17 , pp.1-19,2023.
[8] C. Shorten, T. M. Khoshgoftaar , B. Furht, "Deep Learning applications for COVID 19",Journal of Big Data, vol.8, no.18, pp.1-54, 2021.
[9] K. V. Cortés-Martínez, H. Estrada-Esquivel, A. Martínez-Rebollar, Y. Hernández-Pérez, and J. Ortiz-Hernández, "The State of the Art of Data Mining Algorithms for Predicting the COVID-19 Pandemic," Axioms, vol. 11, no. 5, p. 242, May 2022.
[10] R. Pal, A.A. Sekh , S.Kar, D.K.Prasad, " Neural Network Based Country Wise Risk Prediction of COVID-19", Appl. Sci. ,vol.10, pp.6448,2020.
[11] M.Pourhomayoun ,M. Shakibi . "Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making", Smart Health, vol.20, 2021.
[12] S. Roy, G. S. Bhunia, P. K. Shit, "Spatial prediction of COVID-19 epidemic using ARIMA techniques in India," Modeling Earth Systems and Environment, 7, 2021.
[13] M. Taimoor, S. Ali, I. Shah, F. R. Muwanika, "COVID-19 Pandemic Data Modeling in Pakistan Using Time-Series SIR," Computational and mathematical methods in medicine, Jun 28, 2022.
[14] H. Alalawi, M. Alsuwat, and H. Alhakami, "A Survey of the Application of Artifical Intellegence on COVID-19 Diagnosis and Prediction", Eng. Technol. Appl. Sci. Res., vol. 11, no. 6, pp. 7824–7835, Dec. 2021.
[15] A. Qaiser, S. Hina, A. K. Kazi, S. Ahmed and R. Asif, "Fake News Encoder Classifier (FNEC) for Online Published News Related to COVID-19 Vaccines," Intelligent Automation & Soft Computing, vol. 37, no. 1, pp. 1-14, Apr. 2023, doi: [10.32604/iasc.2023.017978].
[16] M. Mahyoob, J. Algaraady, M. Alrahiali, and A. Alblwi, "Sentiment Analysis of Public Tweets Towards the Emergence of SARS-CoV-2 Omicron Variant: A Social Media Analytics Framework", Eng. Technol. Appl. Sci. Res., vol. 12, no. 3, pp. 8525–8531, Jun. 2022.