Model Comparison Report

1. Introduction

In this section, we evaluate the performance of five time series forecasting models: ARIMA, Auto Regressive (AR), SARIMA, Holt-Winters, and Facebook Prophet. These models were trained on COVID-19 data from India and the United States, and their effectiveness in predicting future trends was assessed using Mean Squared Error (MSE). By evaluating these models, we aim to determine which model provides the most accurate forecasts for COVID-19 cases in these two countries.

2. Model Overview

The following models were considered for this evaluation:

ARIMA (Auto Regressive Integrated Moving Average)

A foundational model in time series forecasting that captures trends in non-seasonal data. ARIMA works well when the data does not exhibit significant seasonal patterns but requires the data to be stationary.

Auto Regressive (AR)

A model based purely on the past values of a time series. AR models predict future values by using the weighted sum of previous values. It’s particularly effective for short-term forecasting but doesn't handle seasonality or irregularities like other models.

SARIMA (Seasonal ARIMA)

An extension of ARIMA that includes seasonality, making it useful for predicting periodic events such as COVID-19 waves. SARIMA models both trend and seasonality, which is especially useful for pandemics that experience repeated surges.

Holt-Winters Exponential Smoothing

This model captures both trend and seasonality through exponential smoothing. It is particularly effective for modelling time series with clear seasonal effects, such as COVID-19 waves that follow seasonal patterns.

Facebook Prophet

A modern forecasting tool that excels in handling missing data, large datasets, and multiple seasonality effects. Prophet is particularly robust for forecasting irregular and noisy data and has been shown to work well for public health data like COVID-19.

3. Evaluation Metrics

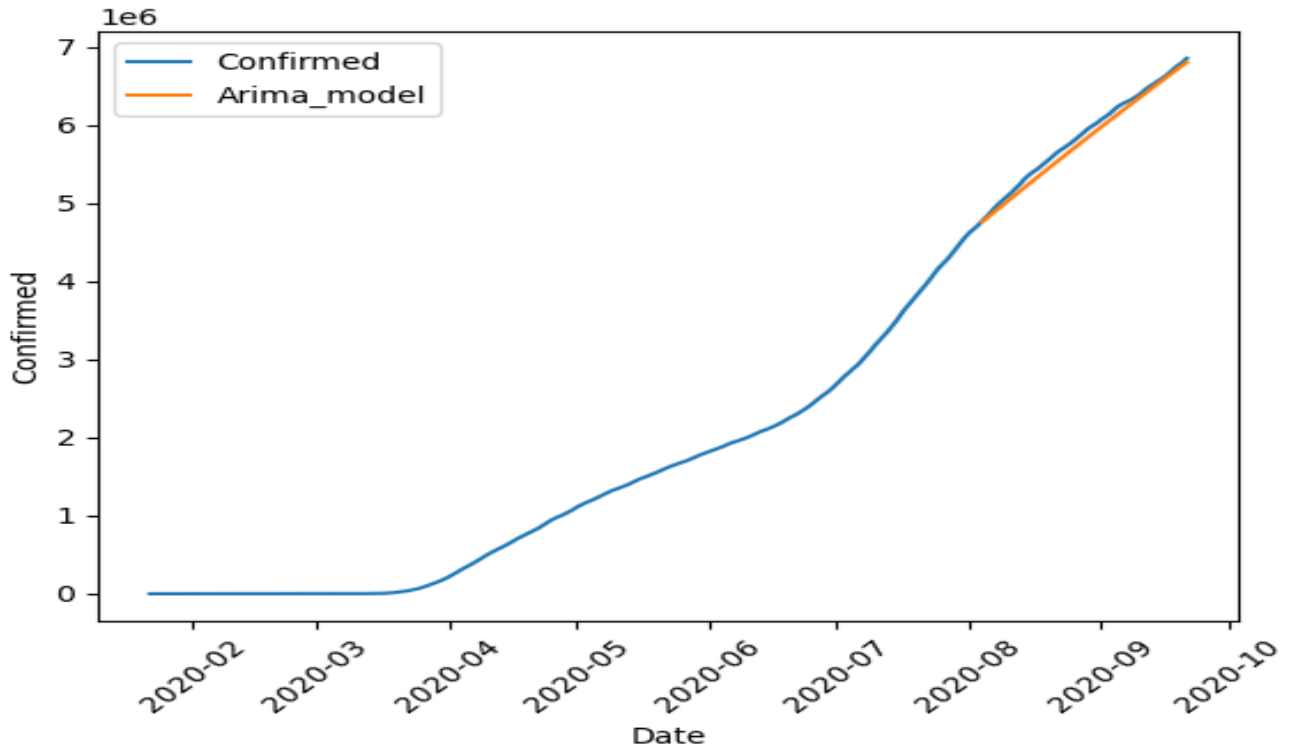
The evaluation metric used to assess the performance of each model was Mean Squared Error (MSE). The MSE measures the average squared difference between the actual and predicted values, with a lower MSE indicating better model performance. This metric is particularly useful in comparing the overall accuracy of the models in predicting future trends.

4. Model Performance Evaluation

1) ARIMA

ARIMA was effective in capturing overall trends in the data but struggled to predict seasonal fluctuations and large surges. Its MSE was relatively high compared to the other models due to the lack of seasonal components and its limited capacity for handling irregularities.

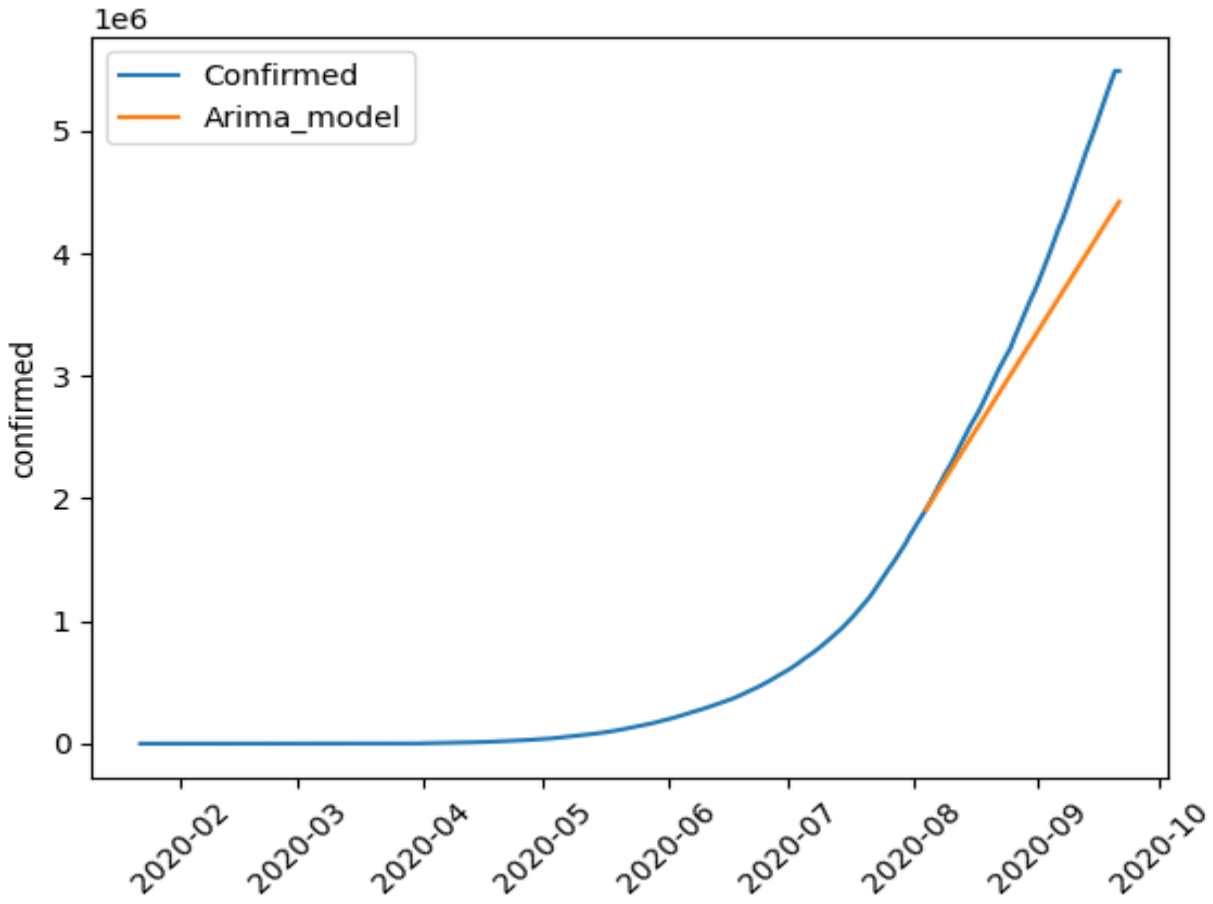
Line Plot Comparing Actual vs ARIMA Model Predictions for COVID-19 Cases for US region



Insights

- The ARIMA model closely follows the actual confirmed cases, indicating strong short-term predictive capability.
- It effectively captures the overall trend without significant deviation, which suggests that past values are well-utilized in forecasting.
- The model assumes a mostly linear or slightly exponential growth pattern.
- Since ARIMA is a statistical model relying on past values, it might not fully capture sudden shifts (e.g., lockdown effects, vaccination rollouts).

Line Plot Comparing Actual vs ARIMA Model Predictions for COVID-19 Cases in India



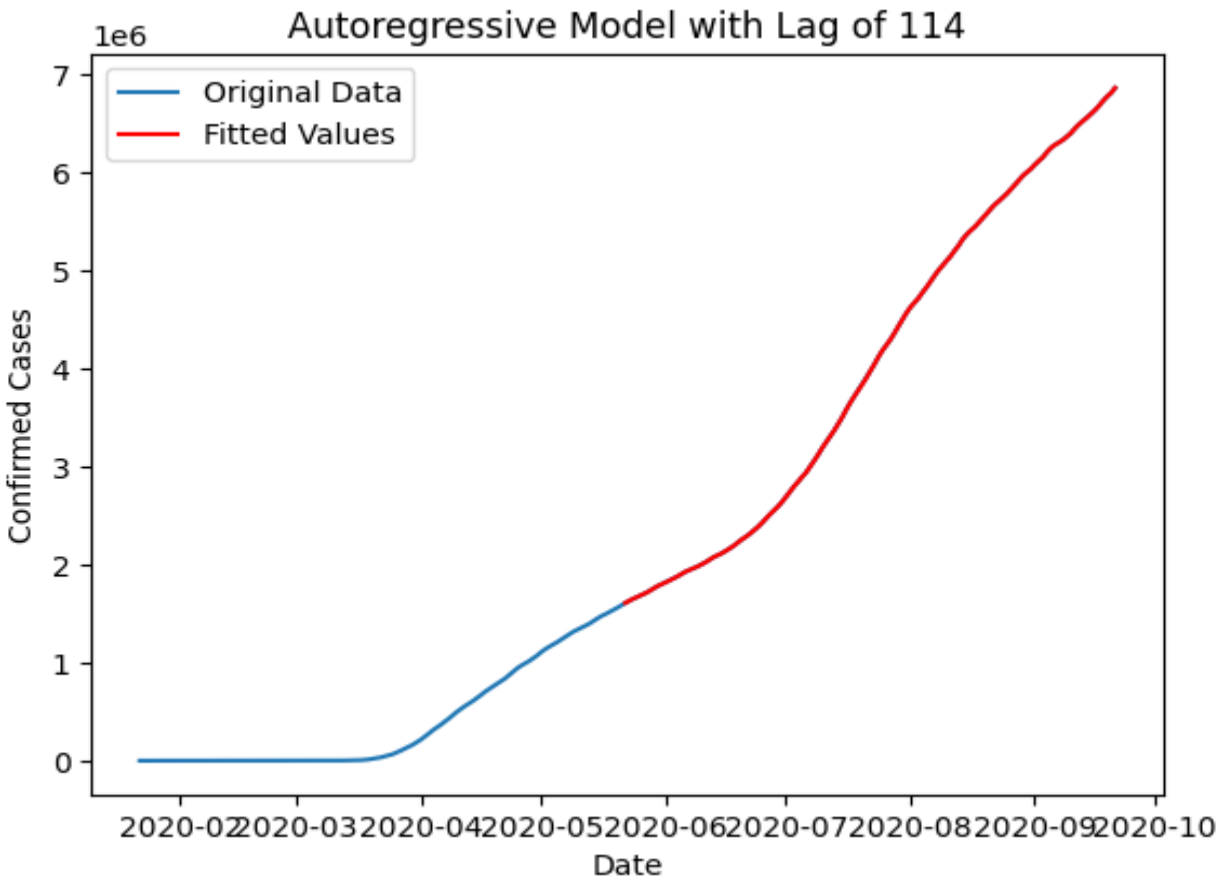
Insights

- The ARIMA model (orange line) underestimates the rapid increase in confirmed cases.
- The actual cases (blue line) show an exponential growth trend, while the model predicts a more linear increase.
- Initially, the ARIMA model aligns with real data but starts deviating as cases rise sharply.
- This suggests that the model struggles to adapt to sudden acceleration in case numbers.

2) Auto Regressive (AR)

The AR model was useful for short-term forecasting and did well with data that did not exhibit significant seasonal patterns. However, it struggled to capture large fluctuations and seasonality in COVID-19 data. The AR model's performance showed that while it was good for trend-following, it was less robust than models like SARIMA and Holt-Winters, which could better handle irregular data.

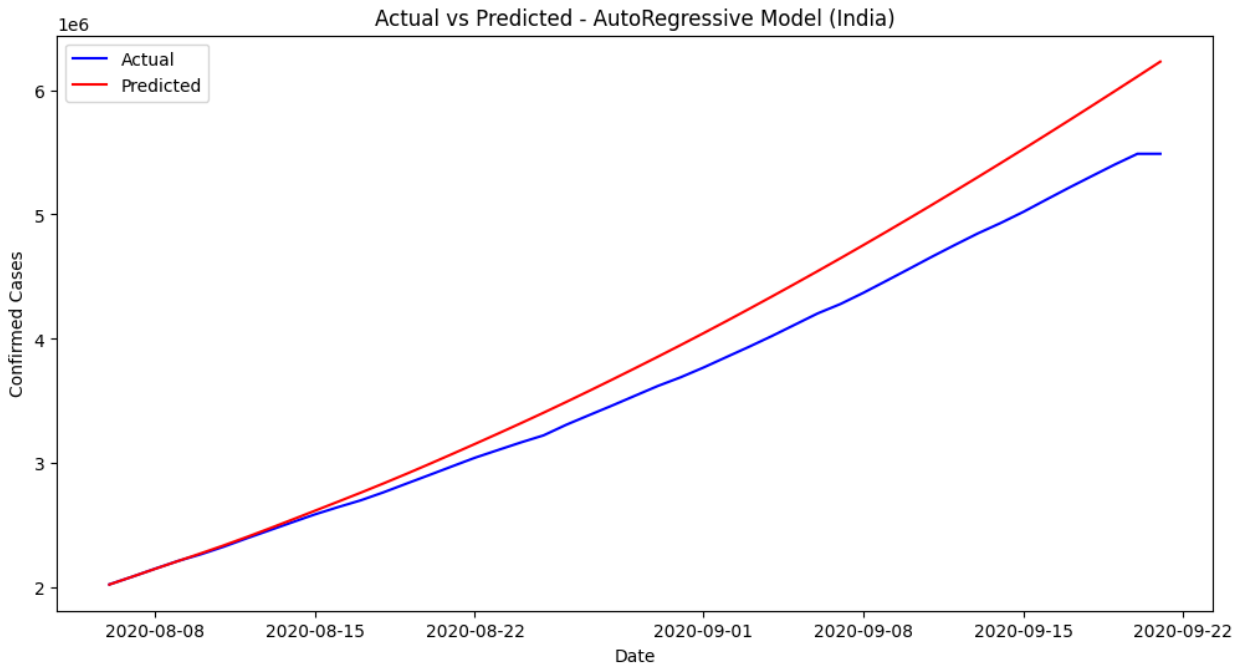
Autoregressive Model with Lag of 114: Fitted vs Actual Data for US region



Insights

- The model's fitted values (red line) closely follow the original data (blue line), indicating that the autoregressive model effectively captures past trends.
- A lag of 114 days suggests the model relies heavily on past values for predictions.
- While this improves short-term accuracy, it may reduce generalizability for long-term forecasting.
- The fitted curve matches the sharp rise in cases, indicating the autoregressive model is better suited for capturing exponential trends compared to simpler models like ARIMA.

Auto Regressive Model (Lag=7) for COVID-19 Cases in India: Actual vs Predicted



Insights

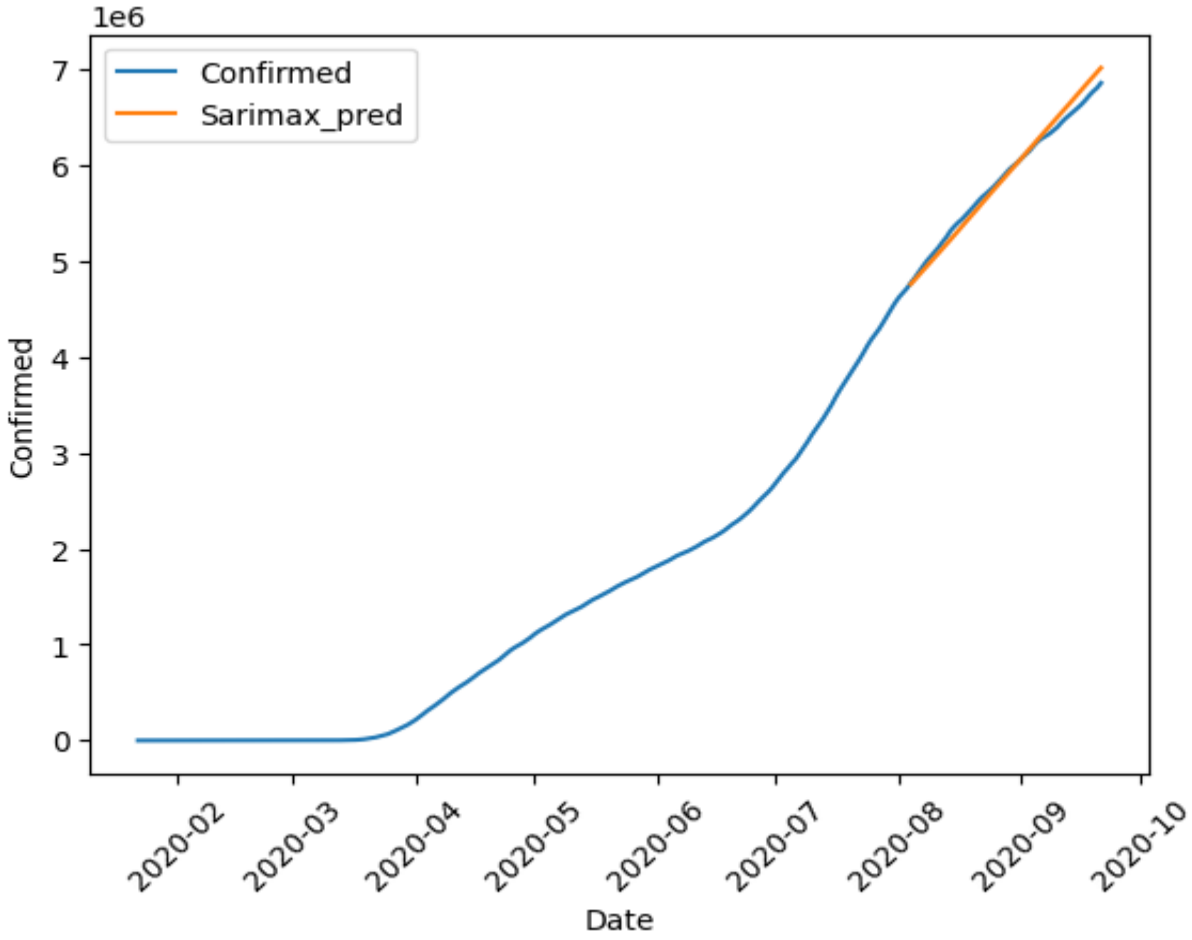
- The predicted values (red line) consistently overestimate the actual values (blue line), especially after early September.

- This suggests that the model assumes a higher growth rate than what was observed in reality.
- In the early stages (early August), the predicted and actual values closely align, indicating the model effectively captures short-term trends.
- As time progresses, the gap between predicted and actual cases widens, showing the model's limitations in handling trend changes.
- The pandemic dynamics changed (e.g., lockdowns, public health measures).
- The model assumes a consistent growth pattern, which did not hold true.

3) SARIMA

SARIMA outperformed ARIMA, as it accounted for seasonal fluctuations, leading to a lower MSE. It captured seasonal peaks in COVID-19 cases more accurately, especially during waves of infections. For India, the SARIMAX model was particularly suitable, accounting for seasonal changes in the data.

Line Plot Comparing Actual vs SARIMAX Model Predictions for COVID-19 Cases in the US

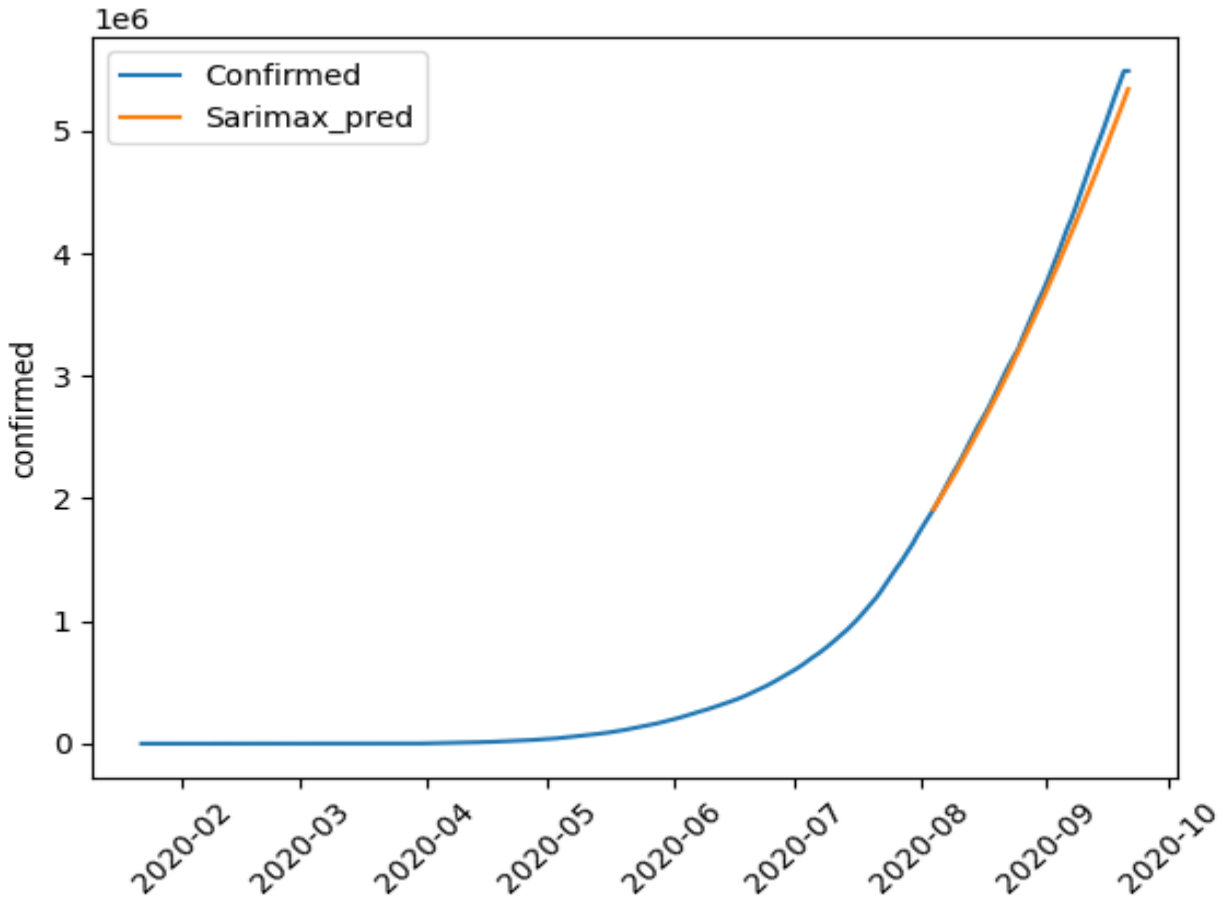


Insights

- The SARIMAX predictions (orange line) closely follow the actual confirmed cases (blue line), indicating that the model effectively captures the overall trend.
- Unlike the previous Auto Regressive model, SARIMAX does not significantly overpredict or underpredict.
- The gap between actual and predicted values is small, making this model a reliable short-term predictor.
- The SARIMAX model assumes a consistent trend with a slight smoothing effect, meaning it avoids sharp fluctuations that may exist in real-world data.

- This suggests that the model relies on historical trends and seasonal patterns rather than sudden changes.

Line Plot Comparing Actual vs SARIMAX Model Predictions for COVID-19 Cases in India



Insights

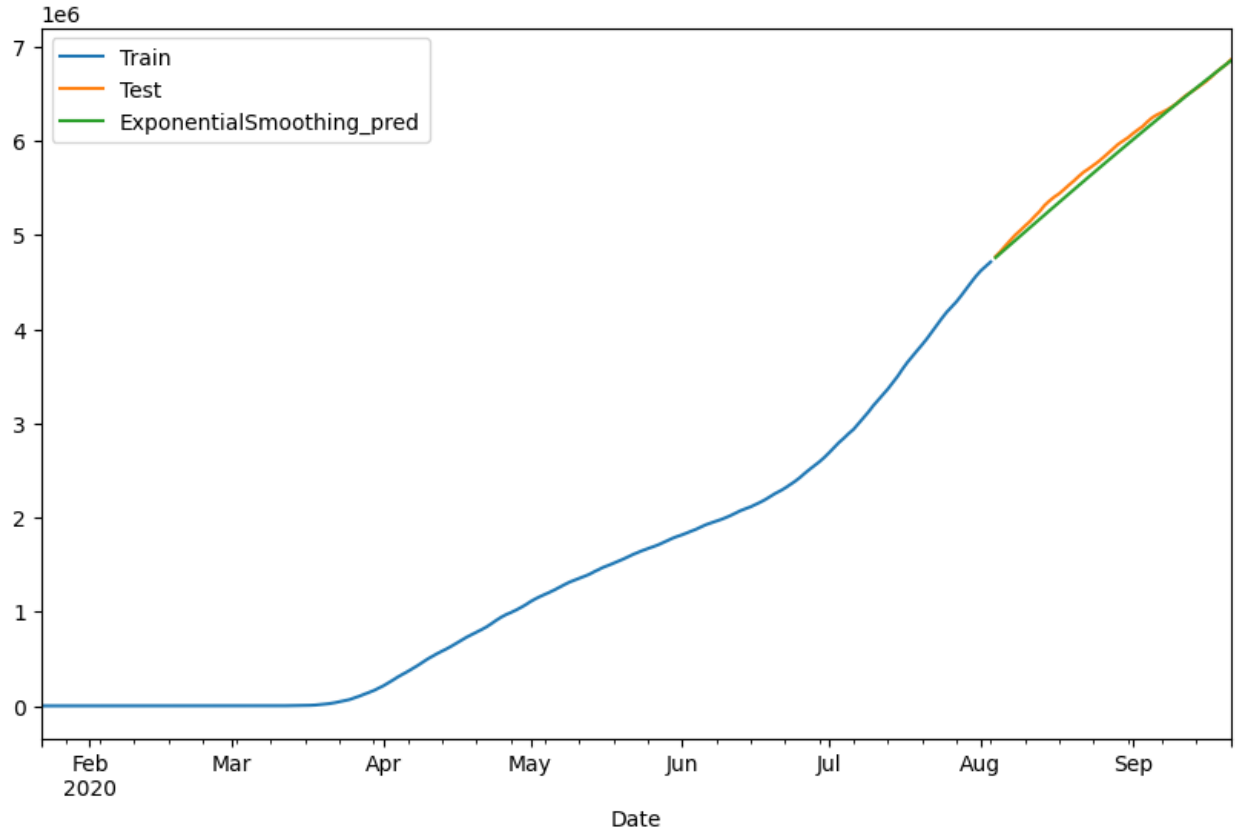
- The predicted values (orange line) closely match the actual confirmed cases (blue line), indicating that SARIMAX effectively captures the trend and seasonality in COVID-19 case growth.

- Initially, the model starts diverging slightly, but as more data becomes available (closer to September 2020), the predictions align almost perfectly with actual cases, suggesting that SARIMAX benefits from a larger training window.

4) Holt-Winters

The Holt-Winters model performed well in capturing both trends and seasonality, especially for countries with clear seasonal patterns like the US. However, it struggled with irregular surges, leading to a slightly higher MSE compared to SARIMA.

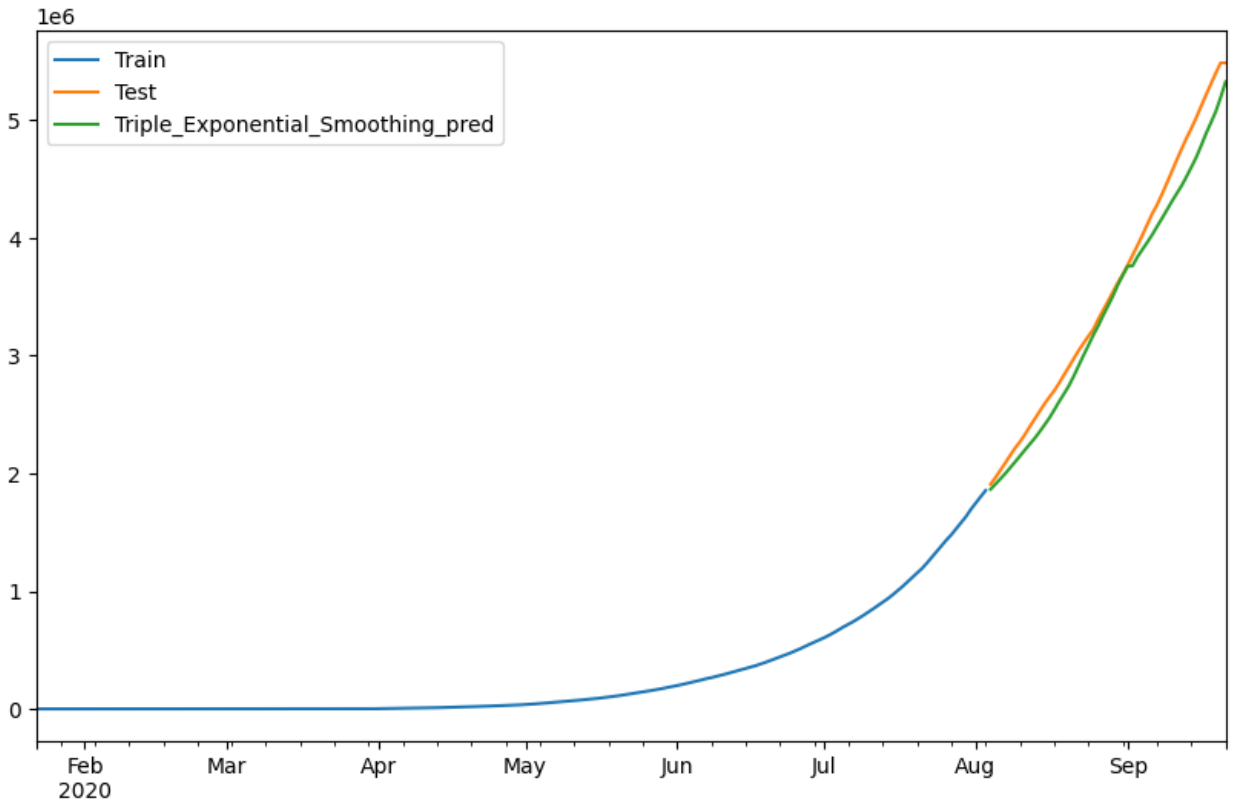
Triple Exponential Smoothing/ Exponential Smoothing Model / Holt-Winters Model for COVID-19 Cases: Train vs Test vs Predicted for US region



Insights

- The green line (Exponential Smoothing prediction) follows the test data (orange line) very closely, indicating that the model successfully captures the overall increasing trend in confirmed cases.
- While the predictions align well with the short-term test data, Exponential Smoothing may struggle with sudden trend changes or seasonal variations, as it assumes a consistent pattern in the data.

Triple Exponential Smoothing Model / Holt-Winters Model for COVID-19 Cases in India: Train vs Test vs Predicted



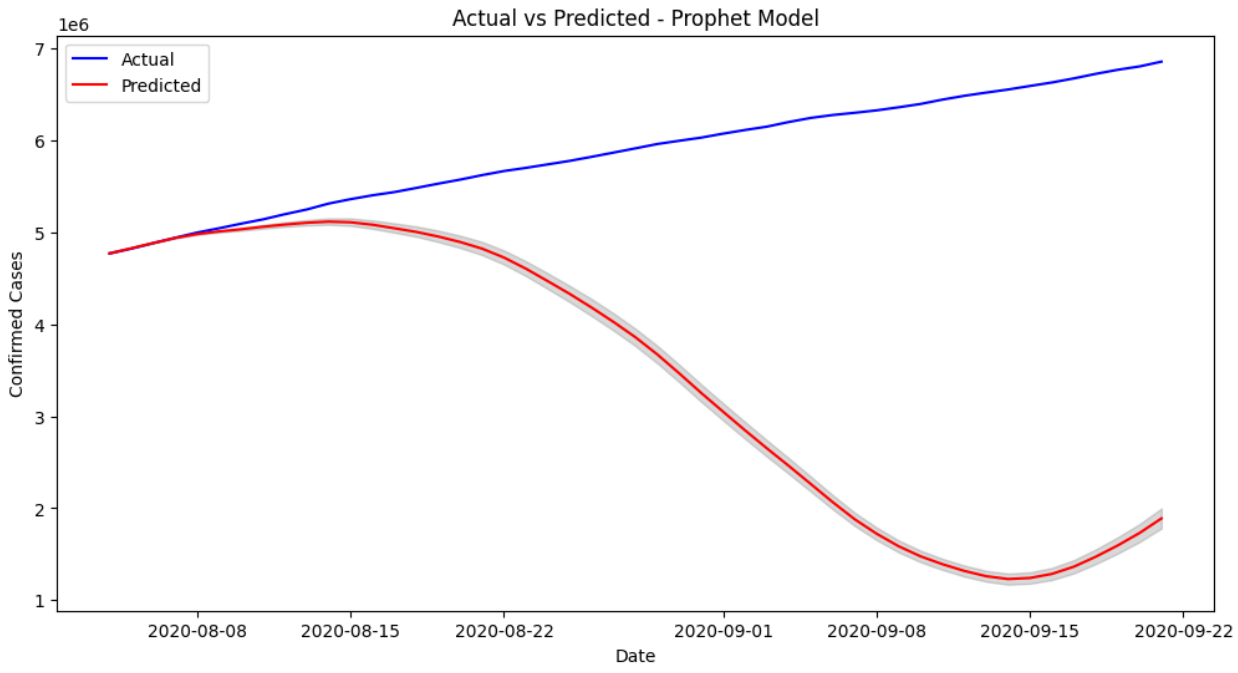
Insights

- Compared to simple Exponential Smoothing, this model (green line) captures the rapid increase in cases more effectively, aligning well with the test data (orange line). This suggests that the model accounts for trend acceleration.
- As cases surge sharply towards the end, the model slightly lags behind the actual test values, indicating that while it adapts well, it may still struggle with highly non-linear growth.

5) Facebook Prophet

Facebook Prophet is a modern time series forecasting tool that excels in handling irregularities in data, missing values, and fluctuations. While it performed well across both countries, SARIMAX (for India) and Holt-Winters (for the United States) were specifically chosen as the best models due to their ability to better capture the seasonal fluctuations and trends specific to each country's COVID-19 data.

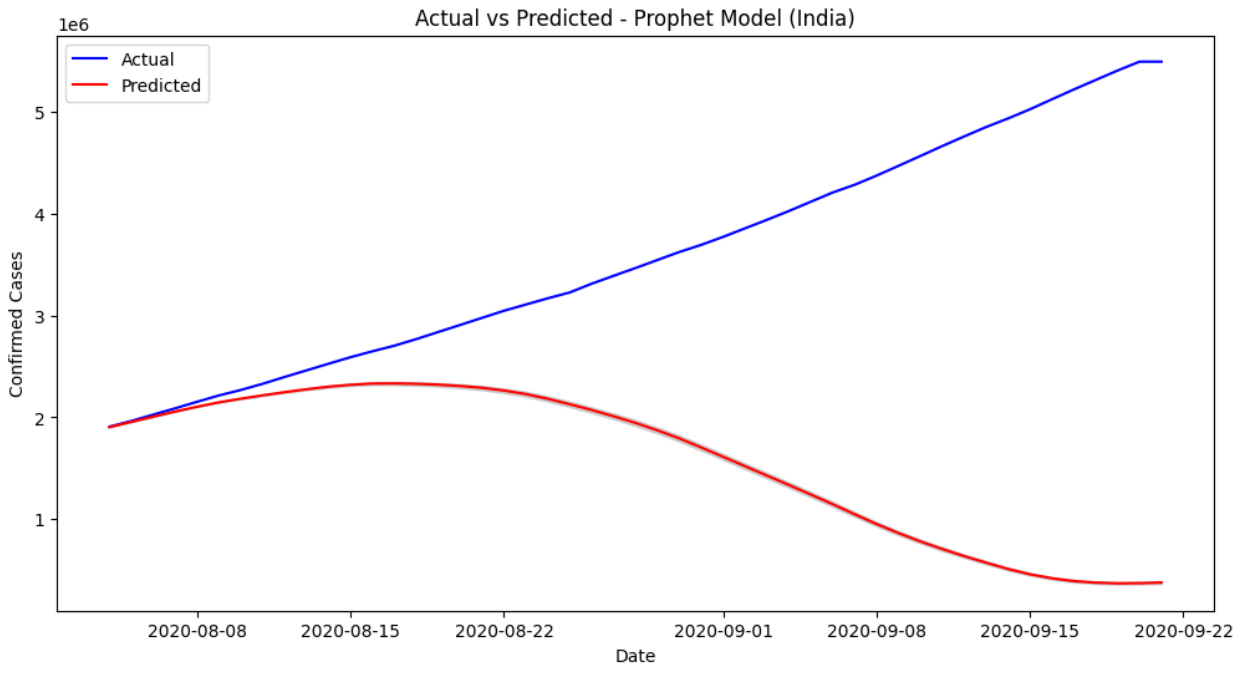
Forecasting COVID-19 Cases Using Facebook Prophet Model: Actual vs Predicted for US region



Insights

- The Prophet model (red line) fails to capture the actual trend (blue line) beyond mid-August. It predicts a sharp decline in confirmed cases, while the actual trend continues to rise. This suggests that the model may not be well-suited for capturing exponential growth or lacks critical trend components.
- The grey shaded region around the prediction indicates a high degree of uncertainty in the forecast. This further reinforces that the model struggles with the underlying pattern in the data, possibly due to improper seasonality assumptions or lack of recent trend adaptation.

Forecasting COVID-19 Cases in India Using Facebook Prophet Model: Actual vs Predicted for India



Insights

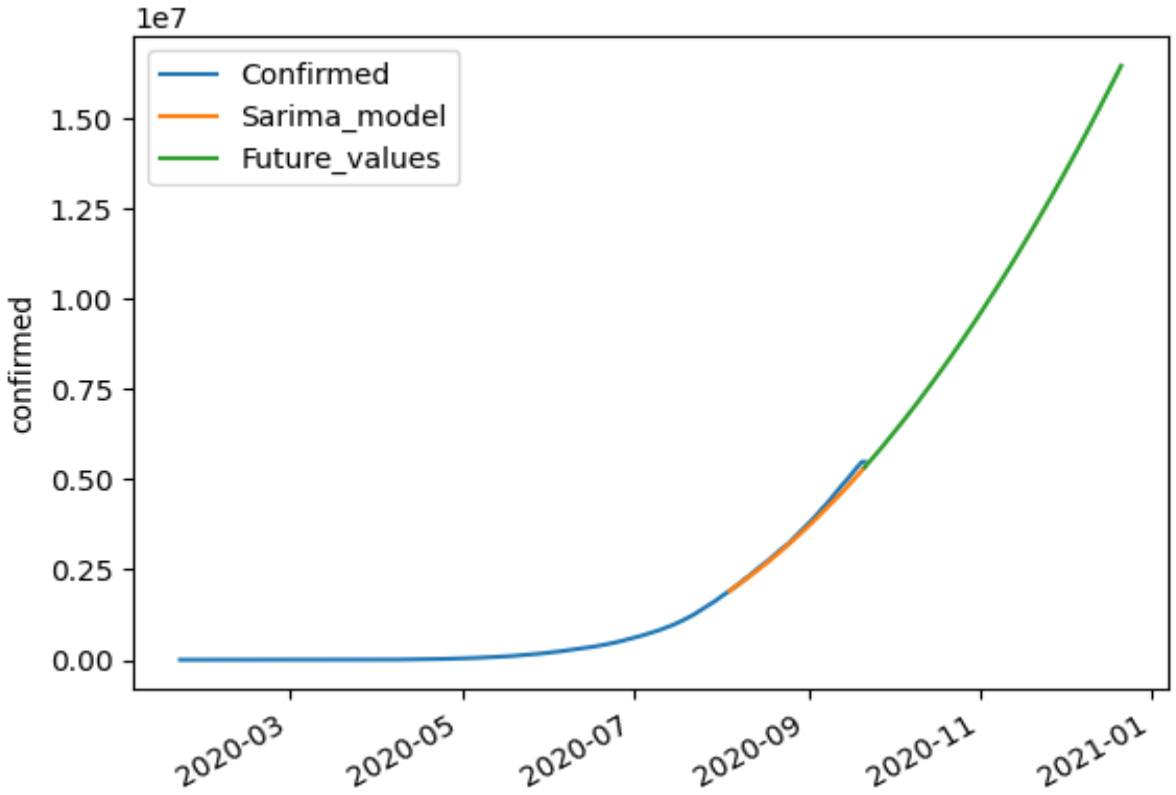
- The Prophet model (red line) fails to capture the actual trend (blue line) and incorrectly predicts a decline in confirmed cases after mid-August, whereas the actual cases continue to rise steadily. This suggests the model may not be properly accounting for the growth trend or recent patterns in the data.
- The Prophet model might be assuming a mean-reverting trend or seasonality that does not align with the real-world scenario. This could be due to inadequate tuning of hyperparameters or failure to incorporate external factors influencing the spread of cases.

5. Conclusion

Based on the MSE evaluation

1) For India, SARIMAX was the best model for COVID-19 prediction. This model effectively captured seasonal fluctuations, leading to the lowest MSE for India's data. SARIMAX's ability to handle seasonality made it particularly suitable for India's COVID-19 data, where periodic surges were observed. So, SARIMAX Model was used for future prediction for India region.

COVID-19 Cases in India: SARIMA Model Predictions with Future Forecast

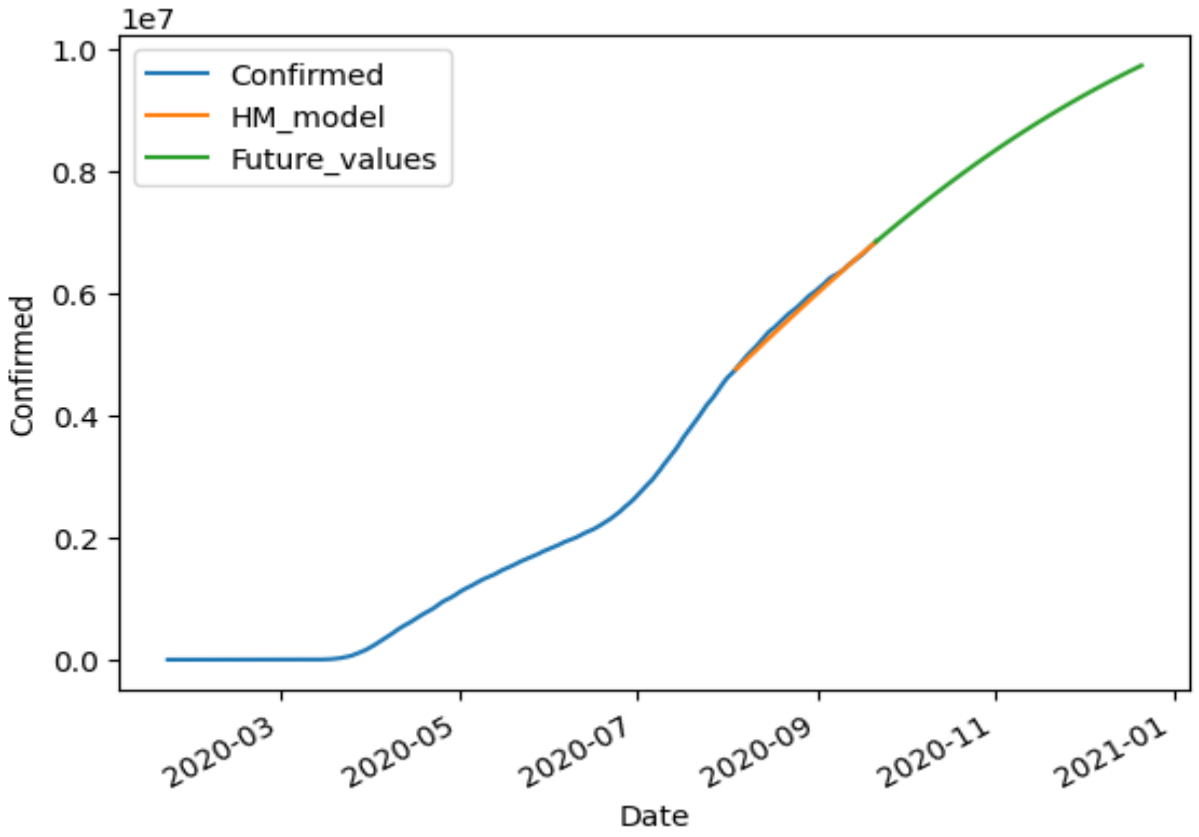


Insights

- The orange line (SARIMA model) closely follows the blue line (actual confirmed cases) before the forecast period, indicating that the model captures the trend effectively.
- A slight deviation might suggest room for fine-tuning seasonal and trend components.
- The green line (Future_values) extends the SARIMA predictions into the future, showing an exponential increase in confirmed cases.
- This suggests that without intervention (e.g., lockdowns, vaccinations), cases would continue rising at a rapid rate.
- The model assumes historical trends persist, but real-world changes (policy shifts, virus mutations) could alter outcomes.

2) For the United States, Holt-Winters (Triple Exponential Smoothing) was the best choice. The model accurately modelled the seasonal variations and trends in the US, even though it struggled with larger surges in cases during unusual periods, which led to a slightly higher MSE compared to SARIMAX. So, Holt-Winters Model was considered as the best model for future prediction for US region.

COVID-19 Cases in US region: Holt-Winters Model Predictions with Future Forecast



Insights

- The HM model (orange line) follows the actual confirmed cases (blue line) closely, indicating a well-fitted model that captures the historical trend effectively.

- The Future values (green line) suggest a continuous upward trend in confirmed cases, but at a slightly decelerating rate. This could indicate the model expects interventions, immunity buildup, or external factors to slow down the growth.

3) Facebook Prophet, while robust, did not outperform SARIMAX for India or Holt-Winters for the United States when evaluated on MSE. Prophet is still a great model, particularly in handling irregular data and missing values, but it was outperformed by the country-specific models in this analysis.

6. Summary

In summary, Facebook Prophet proved to be a good model overall but did not outperform SARIMAX for India and Holt-Winters for the United States when evaluated on MSE (Mean Square Error). SARIMAX and Holt-Winters were better suited for these specific countries due to their ability to handle seasonality and trends in COVID-19 data more effectively.