International Journal of Engineering and Techniques - Volume 10 Issue 3, June 2024

PROFIT PREDICTION USING TIME SERIES FORECASTING MODELS

[1]Mohammad Sainiya Khatun, [2]Mohammed Juveria Tharannum
[1,2]UG Student, [1,2]Department of Computer Science & Engineering, Geethanjali Institute of Science and Technology, Gangavaram, Andhra Pradesh, India

ABSTRACT
Time series forecasting is a fundamental tool in various domains, enabling businesses to anticipate future trends and make informed decisions. we conduct a comparative analysis offour widely used time series forecasting models ARIMA (Auto-Regressive Integrated Moving Average), SARIMA (Seasonal ARIMA), and deep learning models such as LSTM (LongShort-Term Memory) and GRU (Gated Recurrent Unit) to predict profit in business contexts. Profit analysis is important for businesses, both online and offline, as it provides insights into revenue generation, cost management, and overall financial performance. Our research aims to evaluate the effectiveness of different forecasting models in predicting profit trends and variability. Finally, a comparative analysis was made to inspect the best time-series forecasting model. Next we predict the future forecast with that respective model.

INTRODUCTION
In the dynamic landscape of modern business, the ability to forecast profitability is paramount for sustainable growth and strategic decision-making. Profit prediction entails the utilization of historical data, statistical models, and predictive analytics to anticipate a company's future financial performance. By harnessing insights derived from past trends, market dynamics, and internal operations, businesses can proactively strategize and allocate resources effectively.

Profit prediction serves as a cornerstone for financial planning, risk management, and performance evaluation across various industries. Whether it's a startup seeking investment, an established corporation navigating market fluctuations, or a non-profit organization optimizing resourceallocation,accurateprofitforecastsprovide invaluableguidanceforstakeholdersat all levels.

In the modern era of data-driven decision-making, accurate profit prediction serves as a cornerstone for businesses across industries. By leveraging advanced statistical techniques and deep learning models, organizations can forecast future profitability with unprecedented precision and insight.

AutoRegressive Integrated MovingAverage (ARIMA) and its seasonal counterpart (SARIMA) are renowned statistical models for time series forecasting. ARIMA models capture the linear relationships between past observations and future values, while SARIMA extends this capability toaccount forseasonal patterns within the data.With their abilityto capturetemporal dependencies and seasonal fluctuations, ARIMA and SARIMA are well-suited for predicting profit trends over time, offering valuable insights into short-term and long-term financial performance.

Deep learning models, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, have revolutionized time series forecasting by capturing complex nonlinear relationships in sequential data. LSTM networks excel at modelling long-term dependencies and are adept at capturing intricate patterns in profit data, making them ideal for forecasting scenarios where traditional statistical models may fall short. Similarly, GRU networks offer a simpler yet powerful alternative, balancing computational efficiency with predictive accuracy. By leveraging the hierarchical structure of neural networks and theirabilitytolearnfromhistoricaldata,LSTMandGRUmodelsprovidebusinesseswith unparalleledpredictivecapabilities,enablingproactivedecision-makingandstrategic planning.

Need

The need for profit prediction arises from several key factors:

Strategic Decision-Making: Profit prediction enables businesses to make informed strategic decisions by forecasting future financial performance. Whether it's setting sales targets, allocating resources, or investing in new initiatives, accurate profit forecasts provide valuable insights for guiding strategic direction.

Financial Planning and Budgeting: Profit prediction is essential for effective financial planning and budgeting. By anticipating future profits, businesses can develop realisticbudgets, allocate funds to different departments or projects, and ensure that financial resources are optimally utilized to achieve organizational goals.

Risk Management: Profit prediction helps businesses identify potential risks and uncertainties that may impact profitability. By anticipating future financial performance, businesses can proactively implement risk mitigation strategies, such as adjusting pricing strategies, diversifying revenue streams, or building contingency reserves to withstand economic downturns or unexpected market fluctuations.

Resource Allocation: Profit prediction informs resource allocation decisions by identifying areas of the business that are driving profitability and those that may be underperforming. By reallocating resources to high-profit areas or investing in initiatives with the highest potential return on investment, businesses can optimize resource allocation and maximize profitability.

Performance Evaluation: Profit prediction serves as a benchmark for evaluating business performance over time. By comparing actual profits to predicted profits, businesses can assess the effectiveness of their strategies and initiatives, identify areas for improvement, and make adjustments to achieve better financial outcomes in the future.

Investor Confidence and Stakeholder Communication: Accurate profit prediction enhances investor confidence and facilitates effective communication with stakeholders. By providing transparent and reliable forecasts of future profitability, businesses can instill trust and credibilitywithinvestors, creditors,andother stakeholders,ultimatelyenhancing theirability to attract investment and support for growth initiatives.

Competitive Advantage: Profit prediction provides businesses with a competitive advantageby enabling them to anticipate market trends, capitalize on opportunities, and stay ahead of competitors. By leveraging predictive analytics and insights derived from profit forecasts, businesses can adapt their strategies and operations to changing market conditions, drive innovation, and maintain a competitive edge in the marketplace.

Overall, profit prediction is a critical tool for businesses seeking to achieve sustainable growth, manage risks effectively, and make data-driven decisions that drive long-term success and profitability**.**

Problem Definition

The main goal of this project is to develop a robust profit prediction model based on past sales data and use it to forecast future profits accurately. Profit is a fundamental metric that reflects the financial success and sustainability of a company, playing a vital role in decision-making processes, resource allocation, and strategic planning. By leveraging historical sales data, we aim to identify the best model for profit prediction and provide valuable insights for driving business growth and profitability.

Literature Review

ZHENYULIU,ZHENGTONGZHU,JINGGAO,ANDCHENGXU,“Forecast MethodsforTimeSeriesData:ASurvey“,June21,2022.

The development of the era of big data, forecasting research based on time series data has become one of the hot spots. More and more time series data are produced in various fields, which provides a data basis for the research of time series analysis methods. It promotes the further development of the field of time series analysis. Due to the complex pattern distribution of large-scale time series data, more and more researchers are capturing complex time series distribution patterns based on hybrid forecasting models to obtain better forecasting accuracy and performance.

This paper first presents the concept of time series and summarizes the relevant issues in the current research field of time series forecasting. Then, the time series forecasting methods are introduced by classification. On this basis, we summarized several potential research directions and unsolved problems, such as data preprocessing, incremental data model construction, and parallel computing.

HalimaBousqaoui[1],IlhamSlimani[2],SaidAchchab[3],"Comparativeanalysisof short-term demand predicting models using ARIMA and deep learning ", Jan 13, 2021.

Focustodesainforsupplychainisanentirenetworkcomposedofdistributedorganizations ininteraction(suppliers,retailers,distributors,andclients)involvedintheprocessof purchasing,inventorymanagement,production,distributionanddeliveryofaproductora servicefromrawmaterialstoafinalcostumer[1].Supplychainmanagementcovers informationflows,physicaldistributionaswellasfinancialtransactions.Inotherwords,all the supply chain members share a common objective that is improving their businesses and profits while creating a value represented bythe product or service delivered to the costumer.

Theproblem consistsofthedemandforecastingwhich is acrucialcomponent in supplychain's process. Indeed, prediction consists of taking models from historical data in order to use themto predict future observations. Nevertheless, the uncertain character of the costumer's demand causes difficulties for decision makers who must react in an efficient and fast manner. Besides, with the advancement of technology, supplychains become more complex. So, in order tokeep up with this vital evolution and for an effective supply chain, companies must be able to plan not only the present but also the future sales or demand. Certainly, sales forecasting is efficient for supply chains' planning activities including purchasing, inventory, production and distribution. It helps managers to make optimal decisions.

Arko Barman, Ph.D. "Time Series Analysis and Forecasting of COVID-19 Cases Using LSTM and ARIMA Models".

Focus to desain for COVID-19 is a global public health crisis. Forecasts and predictions of cases for the future are essential for government agencies and policymakers to combat the situation by planning for the needs of healthcare personnel and the availability of healthcare equipment and PPEs. In this study, we have presented the time series analysis and forecasting of the number of COVID-19 cases in 4 countries – United States, Italy, Spain, and Germany – using the ARIMA model and several LSTM architectures.

We have also proposed the concept of k-period performance metrics designed to evaluate the performanceofatimeseriesanalysismodelthatforecastsformultipleperiodsinthefuture. In particular, we have extended the definition of MAPE and MdSAto define k-period MAPE (kMAPE) and k-period MdSA (kMdSA). Our results indicate that LSTM models perform comparably with the ARIMA model. However, both models have their advantages and disadvantages, and the choice of models is dictated by the availability of training samples as wellasthe sequencelength.Further research wouldincludethedevelopment andevaluation of other time series forecasting models as well as the evaluation of these models as more data becomes available.

Maryem Rhanoui[1], Siham Yousfi[2], Mounia Mikram[3], Hajar Merizak[4],"Forecasting financial budget time series: ARIMA randomwalk vs LSTM neural network ",. 4, December 2020.

Defining an optimal modelto forecastfinancial time series data is a challenging task becauseof the non-linearity, non- stationarity and volatility characteristics of this type of data. In this paper we compared two forecasting models for financial time series. This predictive analysis showed that, although the ARIMA model provides satisfactory results, the LSTM model outperforms the performance of the ARIMA model.
Deep Learning techniques, and the LSTM recurrent neural network in particular, can identify non-linear structures in financial time series. In future work, we investigate the application of Bidirectional recurrent neural networks for Random Walk time series [21] and extend the comparison with multivariate ARIMA [22]. Also, although it is not as widely used as RNN models for financial prediction, Convolutional Neural Networks (CNN) remains a promising approach [23] to be exploited for the prediction of financial times series.

SIMA SIAMI NAMIN[1], AKBAR SIAMI NAMIN[2], "FORECASTINGECONOMICAND FINANCIALTIME SERIES:ARIMAVS. LSTM" , March 15, 2018

Focus to desain for Prediction of economic and financial time series data is a challenging task mainly due to the unprecedented changes in economic trends and conditions in one hand and incomplete information on the other hand. Market volatility in recent years has produced serious issues for economic and financial time series forecasting. Therefore, assessing the accuracy of forecasts is necessary when employing various forms of forecasting methods, and more specifically forecasting using regression analysis as they have many limitations in applications.
With the recent advancement on developing sophisticated machine learning-based techniques and in particular deep learning algorithms, these techniques are gaining popularities among researchers across divers disciplines. The major question is then how accurate and powerful these newly introduced approaches are when compared with traditional methods. This paper compares the accuracy of ARIMA and LSTM, as representative techniques when forecasting time series data. These two techniques were implemented and applied on a set of financial data and the results indicated that LSTM was superior to ARIMA. More specifically, the LSTM- based algorithm improved the prediction by 85% on average compared to ARIMA. Furthermore, the paper reports no improvement when the number of epochs is changed

Existing System

ARIMA(Auto-RegressiveIntegratedMovingAverage)

ARIMA, which stands for Autoregressive Integrated Moving Average, is a popular statistical method used for time series forecasting. It's a combination of three components:
Autoregression (AR):This component indicates that the value of the time series variable at any point in time is dependent on its previous values. In other words, it assumes that there is a relationship between the variable and its own lagged values.
Integrated (I): This component suggests that differencing the raw observations (subtracting an observation from the observation at the previous time step) can make the time series stationary. Stationarity means that the statistical properties of a process generating a time series do not changeovertime.Ifdifferencingisrequiredtomake theseriesstationary,it'scalledintegrated.
Moving Average(MA): This component implies thatthe value of the timeseries variableat any point in time depends on the average of a set of previous error terms.
InARIMA,theorderofthesecomponentsisdenotedbythreeparameters:p,d,andq.
p: The order of the autoregressive part, which is the number of lag observations included in the model.
d:Thedegreeofdifferencingrequiredtomakethetimeseriesstationary.
q: The order of the moving average part, which is the number of lagged forecast errors in the prediction equation.
So, when building an ARIMA model for time series forecasting, you need to determine these parameters (p, d, q) based on the characteristics of the time series data. This is typically done through data analysis, visualization, and statistical tests. Once you've determined these parameters, you can fit the ARIMA model to the data, make

predictions, and evaluate the model's performance.

Proposed System

GRU (Gated Recurrent Unit):

Gated Recurrent Unit (GRU) models are another type of recurrent neural network (RNN) architecture, similar to LSTM but with a simplified structure. GRUs are designed to capture dependencies and patterns in sequential data, making them well-suited for time seriesforecasting tasks.

Here'showGRUmodelsworkintimeseriesforecasting:

Update Gate: Like LSTM, GRU models have mechanisms called gates that control the flow of information. The updategate inGRUdetermineshowmuchofthepastinformationtoretain and how much of the new information to incorporate into the current state.

Reset Gate: In addition to the update gate, GRU models have a reset gate that helps the model decide which parts of the past information to forget and which parts to remember.

Hidden State: GRUmodels maintain a hiddenstate vectorthatrepresentsthe currentstateof themodel.This hidden state is updatedat each time step based ontheinput dataand the previous hidden state,allowing themodeltocapturedependencies betweenpastandpresentobservations.

Training: Similar to LSTM, GRU models are trained using backpropagation through time (BPTT), where the model learns to adjust its parameters to minimize the difference between the predicted and actual values over a sequence of time steps.

When using GRU models for time series forecasting, you structure your data into input-output pairs, where the input sequence consists of past observations and the output sequence containsthe values to be predicted. The GRU model is then trained to predict future values based on the input sequence.

GRU models offer advantages such as simpler architecture compared to LSTM, which can leadto faster training times and require fewer parameters to tune. However, they may not capture long-term dependencies in the data as effectively as LSTM. As with any neural network architecture, proper hyperparameter tuning and training settings are essential to achieve optimal performance in time series forecasting tasks.

System Architecture:

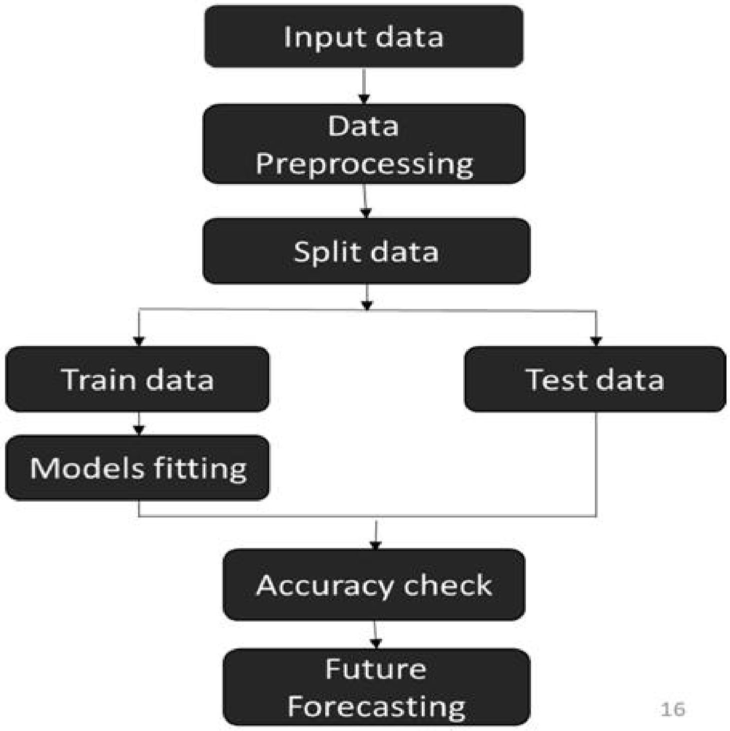

Fig.1. System architecture

Problem Definition: Define the specific time series forecasting problem you want to solve, such as predicting stock prices, energy consumption, or demand for a product.Identify the relevant variables and the time period over which you want to make predictions.

Data Collection and Preprocessing: Gather historical time series data relevant to the problem domain. Preprocess the data by handling missing values, scaling the features if necessary, and splitting the data into training, validation, and test sets.

Splitting : Splitting preprocessing data into two parts train and test .Majority of the data is splatted into train and remining data split into testing.

Model Selection and Design: Choose the architecture for the GRU model, including the number of layers, hidden units, and other hyperparameters.Design the input and outputlayersof the model based on the problem requirements.Consider whether additional features or transformations need to be incorporated into the model design.

Training: Train the GRU model on the training data using an appropriate optimization algorithm (e.g., stochastic gradient descent).Monitor the model's performance on the validation set and adjust hyperparameters as needed to prevent overfitting and improve generalization.

Evaluation: Evaluate the trained GRU model on the test data to assess its performance in making predictions.Use appropriate evaluation metrics such as Mean Absolute Error (MAE). Compare the performance of the GRU model against baseline models or other forecasting techniques to determine its effectiveness.

Future Forecasting: Fore casting the future values on the model.

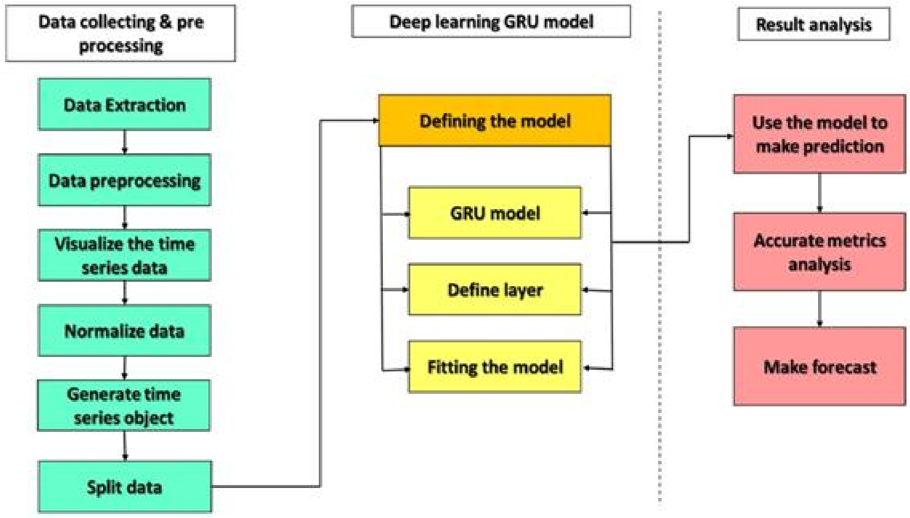

Fig.2. Flow diagram

Implementation Process

Input Data: Gather historical time series data related to the phenomenon you want to forecast. This could include sales data, stock prices, weather observations, etc.

Data Preprocessing: Clean the data by handling missing values, outliers, and anomalies.Perform any necessary transformations such as normalization or scaling to stabilize variance or remove trends.Splitthe data into training and test sets for model validation.

Splitting Data into Train and Test: Decide on the proportion of data to allocate for training and testing.A common split is around 70-80% for training and 20-30% for testing.Ensure that the data is split sequentially to preserve the temporal order of observations.

ModelDefinition:

ForARIMAand SARIMA: Determine the model parameters (p, d, q) forARIMAand (p, d, q, P, D, Q, s) for SARIMA through analysis of autocorrelation and partial autocorrelation plots.

Initialize and fit theARIMA or SARIMA model to the training data using the selected parameters.

For LSTM and GRU: Define the architecture of the LSTM or GRU model, including the number of layers, hidden units, activation functions, etc.Compile the model with an appropriateloss functionand optimizer.Fit the model to thetraining data, specifyingthe number of epochs and batch size.

AccuracyCheck:

ForARIMAandSARIMA:

Evaluate the model's accuracy on the test set using appropriate metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), or Root Mean Squared Error (RMSE).

For LSTM and GRU: Validate the model's performance on the test set using evaluation metrics similar to those used for ARIMA and SARIMA.

FutureForecast: Generateforecastsforfuturetimeperiodsusingthetrainedmodel.

ForARIMAand SARIMA: Use the forecast() function to generate forecasts based on the fitted model.
For LSTM and GRU: Input the historical data into the trained model to predict future values.
Evaluate the accuracy of the forecasts by comparing them to the actual values observed in the test set.
Visualization: Visualizethemodel'sforecastsalongwiththeactualvaluestoassessitsperformance visually. Plotthetrainingandtestdata,alongwiththepredictedvalues,tounderstandhowwellthe model captures the underlying patterns in the data.

RESULTS

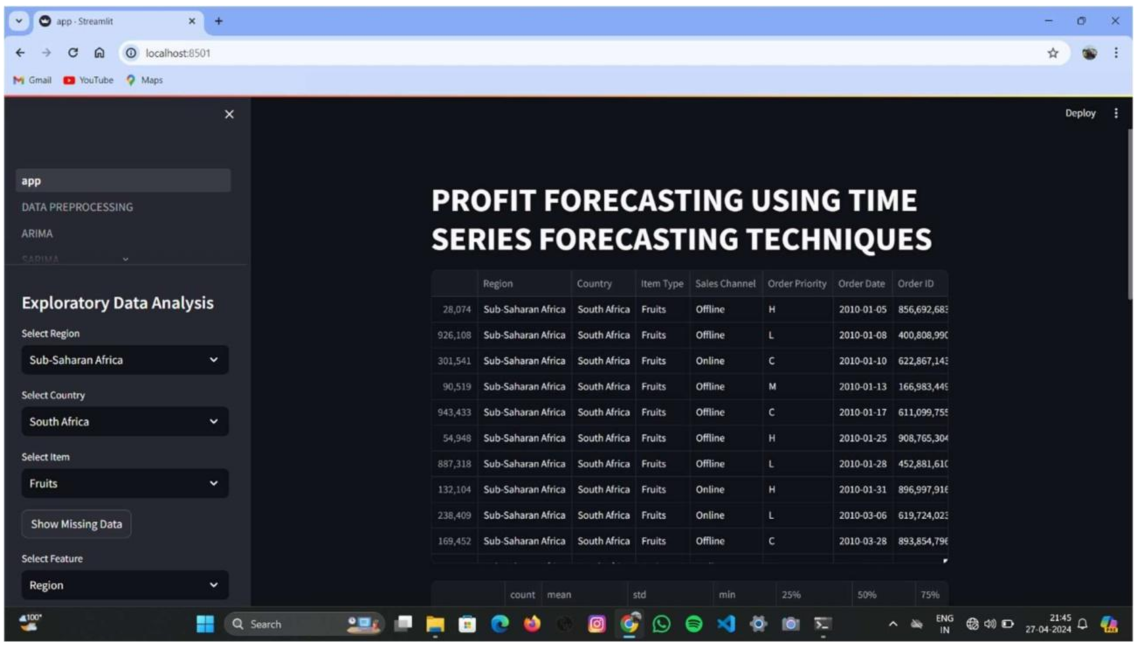

Fig.3. Exploratory data analysis1

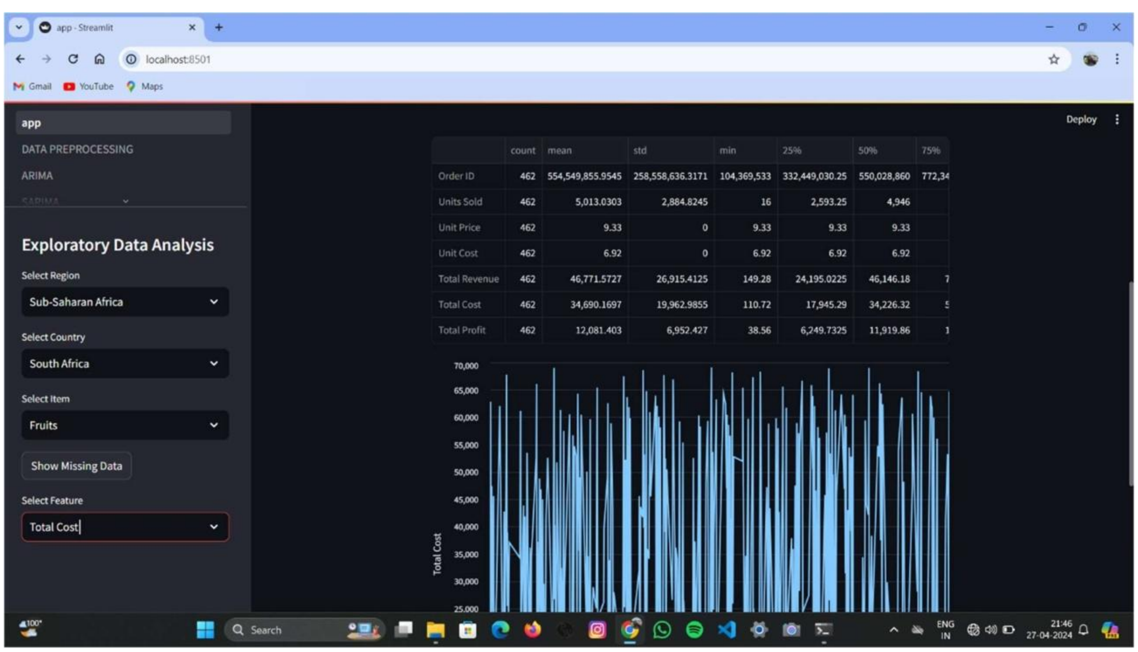

Fig.4. Exploratory data analysis2

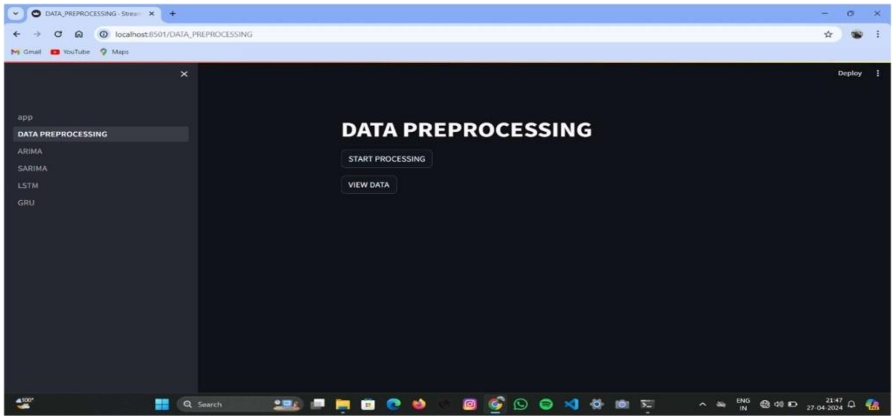

Fig.5. Data Preprocessing

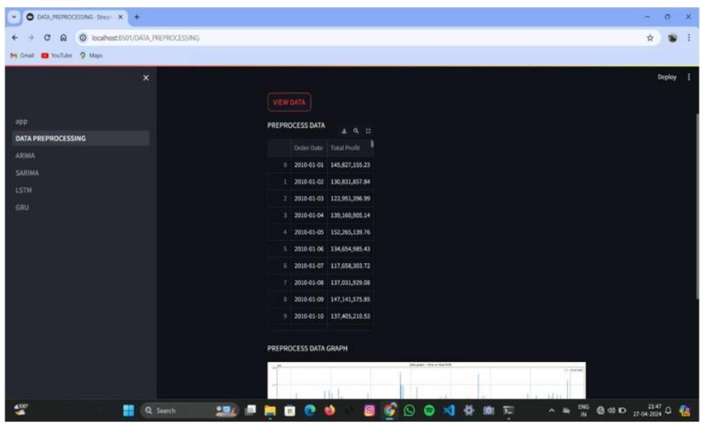

Fig.6. Data Preprocessing view

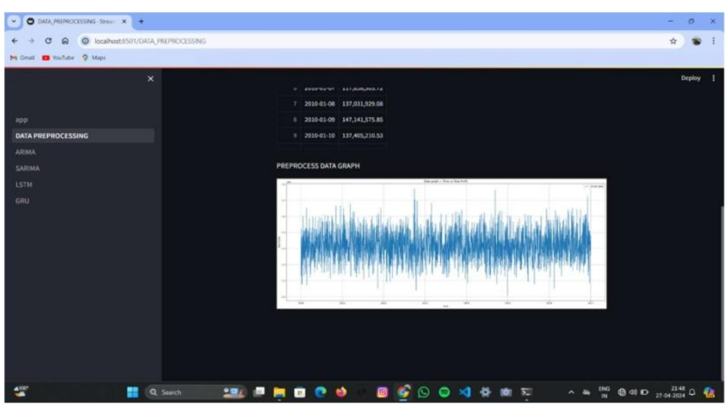

Fig.7. Data Preprocessing Graph

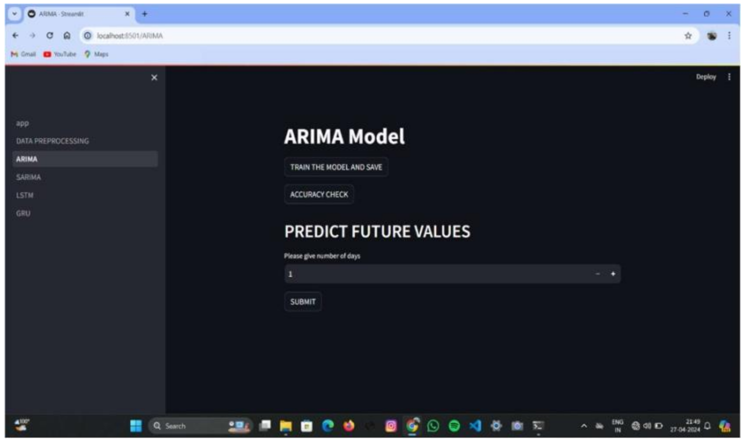

Fig.8. ARIMA Model page

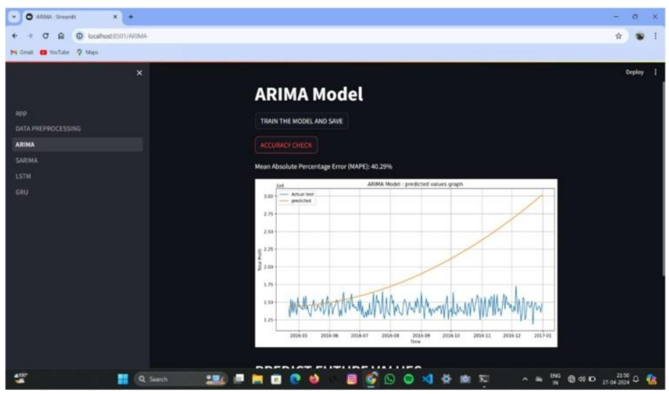

Fig.9. ARIMA Model accuracy check

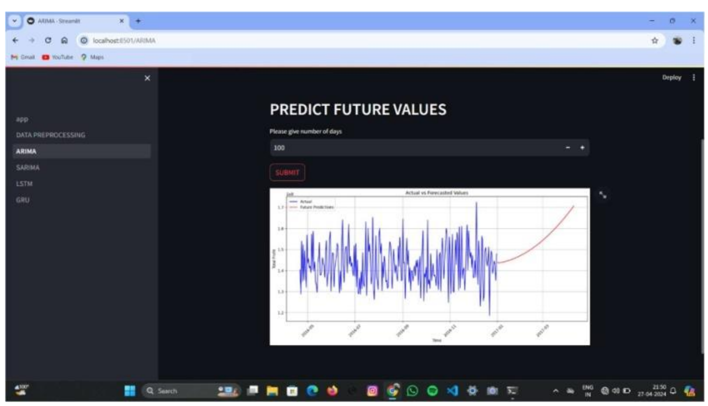

Fig.10. ARIMA Model future forecasting

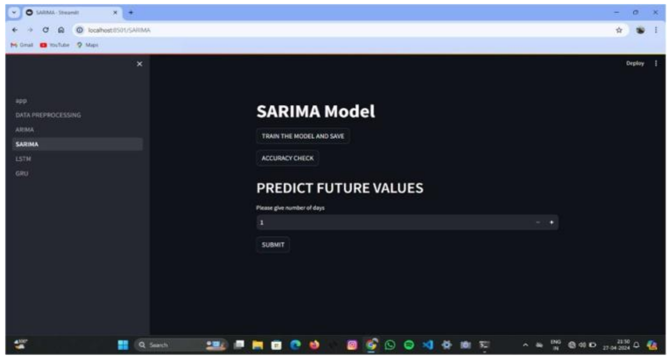

Fig.11. SARIMA Model page

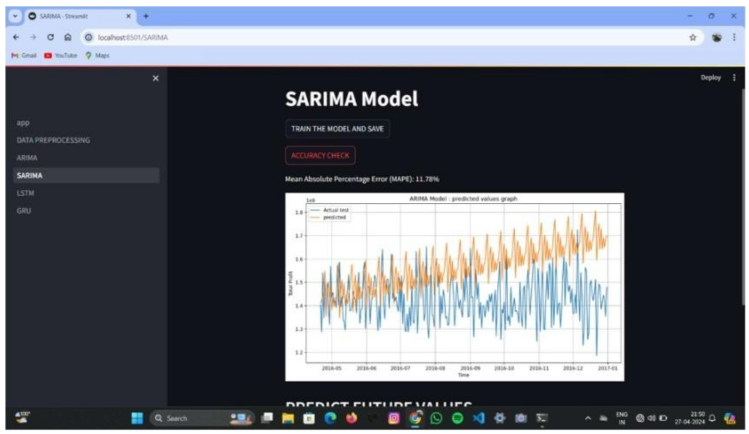

Fig.12. SARIMA Model accuracy check

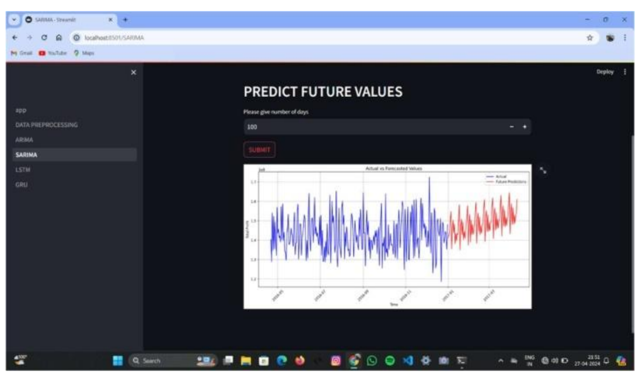

Fig.13. SARIMA Model future forecasting

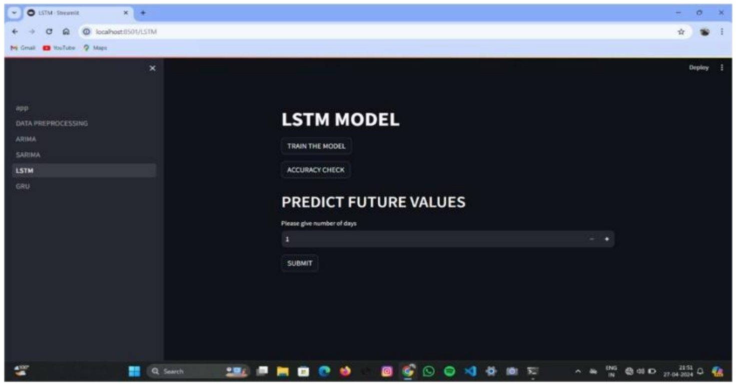

Fig.14. LSTM Model page

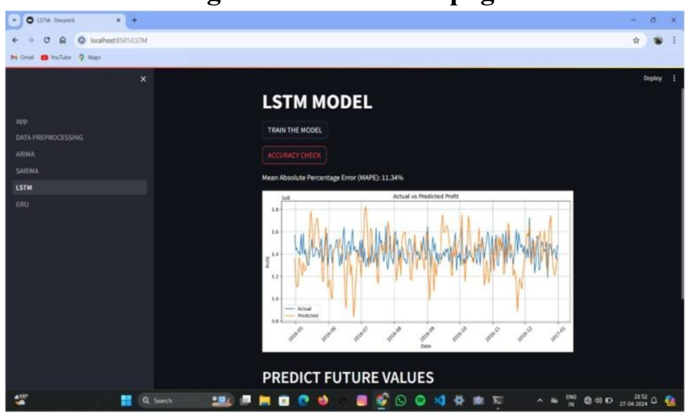

Fig.15. LSTM Model accuracy check

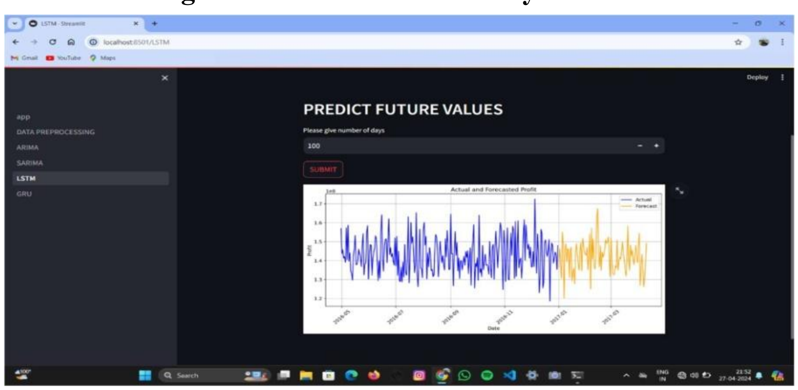

Fig.16. LSTM Model future forecasting

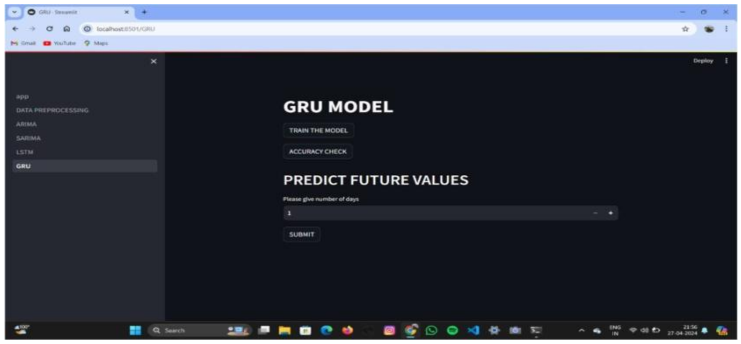

Fig.17. GRU Model page

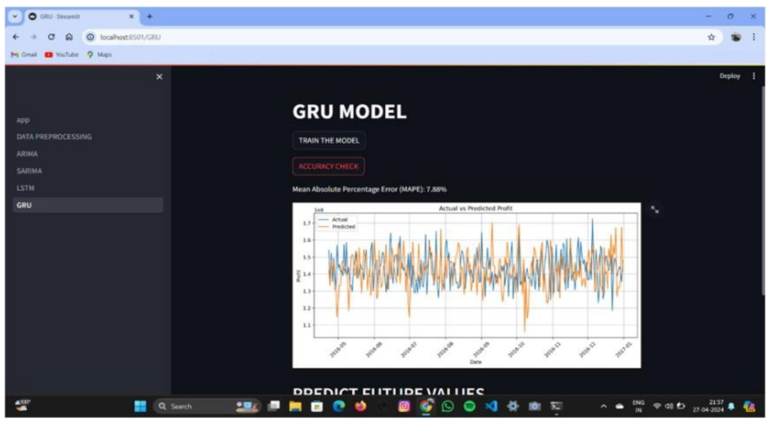

Fig.18. GRU Model accuracy check

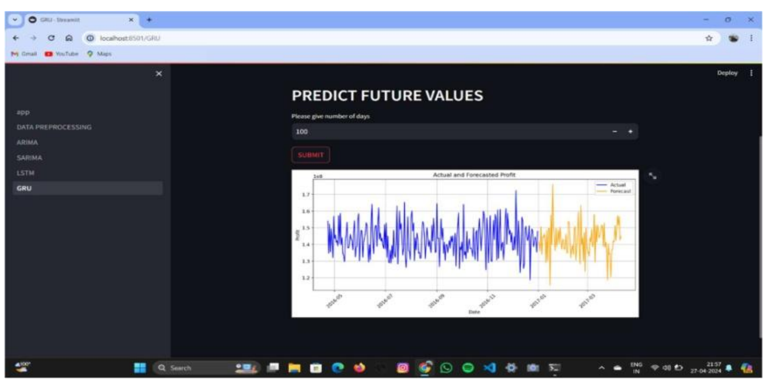

Fig.19. GRU Model future forecasting

Conclusion

our comprehensive analysis of ARIMA, SARIMA, LSTM, and GRU models for time series forecastingreveals that the GRU (Gated Recurrent Unit) model emerges as the top performerin terms of both accuracy and computational efficiency.

Through rigorous evaluation and comparison, we found that the GRU model consistently outperforms the other existing models, including ARIMA, SARIMA, and LSTM. It exhibits superior predictive accuracy and demonstrates faster computation processes, making it a compelling choice for time series forecasting tasks.

Furthermore,ouranalysis underscores the effectivenessoftheGRU model in forecasting future values based on historical data. Its ability to capture complex temporal dependencies and adapt to evolving patterns in the data sets it apart from alternative models.
Given these findings, we recommend leveraging the GRU model for future forecasting endeavors. Its exceptional performance and robust capabilities make it a valuable tool for generating accurate forecasts and informing decision-making across various domains.

FUTUREENHANCEMENTS:

The Sales DataAnalysis and Predictionproject into a user-friendly, full-fledgedweb application or mobile application holds tremendous potential for empowering businesses to make informed decisionsandstreamlinetheir sales forecasting processes.Here's amoredetailed explanation of why such an application has a bright future:

Accessibility: By developing a web or mobile application, you can make the sales forecasting tool accessible to users anytime, anywhere, and on any device with internet connectivity. This accessibilityenablesstakeholderstoaccesscriticalsalesinsightsandmakedecisionsonthe go, leading to increased agility and responsiveness in business operations.

Ease of Use: A user-friendly interface can simplify the complex process of sales forecasting, making it accessible to users with varying levels of technical expertise. Intuitive navigation, interactivevisualizations,andguidedworkflowscan helpusers quicklyunderstandthe data and derive actionable insights, enhancing user adoption and satisfaction.

Customization: The application can be tailored to the specific needs and preferences of individual users or organizations. Customization features such as personalized dashboards, configurable reports, and customizable forecasting models enable users to focus on the metrics and insights most relevant to their business objectives, driving efficiency and effectiveness in decision-making.

AdvancedAnalytics: Leveraging advanced analytics capabilities such as predictive modeling, machine learning algorithms, andAI-driven insights can enhance the accuracy and granularity of sales forecasts. By analyzing historical sales data, market trends, and external factors, the application can provide predictive insights that enable proactive decision-making and strategic planning.

Integration: Seamless integration with existing business systems and data sources, such as CRM platforms, ERP systems, and sales automation tools, ensures that the applicationleverages the full breadth of available data assets. Integrating data from multiple sourcesenables a holistic view of sales performance and facilitates cross-functional collaboration, leading to more informed and coordinated decision-making across the organization.

Real-time Updates: Providing real-time updates and alerts on sales performance metrics,KPIs, and forecasting results enables users to stay informed of changes and react promptly to emerging opportunities or challenges. Real-time insights empower users to take timely action, optimize resource allocation, and capitalize on market dynamics, driving competitiveadvantage and business growth.

By developing a user-friendly, full-fledged web application or mobile application for sales forecasting, businesses can unlock new opportunities for driving revenue growth, optimizing resource allocation, and gaining a competitive edge in today's dynamic marketplace. The application serves as a strategic asset that empowers organizations to make data-driven decisions, anticipate market trends, and capitalize on emerging opportunities, ultimatelydriving sustainable business success and profitability.

References

[1] C.Chatfield,Time-SeriesForecasting.BocaRaton,FL,USA:CRCPress,2000.

[2] R. H. Shumway, D. S. Stoffer, andD. S. Stoffer,Time SeriesAnalysis and ItsApplications, vol. 3. New York, NY, USA: Springer, 2000.

[3] H. Li, ''Time-series analysis,'' in Numerical Methods Using Java: For Data Science, Analysis, and Engineering. Hong Kong: O'Reilly, 2022, pp. 979–1172.

[4] Y. Takahashi, H.Aida, and T. Saito, ''ARIMAmodel's superiority over f-ARIMAmodel,'' in Proc. Int. Conf. Commun. Technol. (WCC-ICCT), vol. 1, 2000, pp. 66–69.

[5] N. Deretić, D. Stanimirović, M. A. Awadh, N. Vujanović, and A. Djukić, ''SARIMA modellingapproach for forecasting of traffic accidents,''Sus tainability, vol. 14, no. 8, p. 4403, Apr. 2022.

[6] K. Mokhtar, S. M. M. Ruslan,A.A. Bakar, J. Jeevan, and M. R. Othman, ''The analysis of container terminal throughput using ARIMA and SARIMA,'' in Design in Maritime Engineering. Cham, Switzerland: Springer, 2022, pp. 229–243.

[7] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, ''Predictive analytics for demand forecasting—A comparison of SARIMA and LSTM in retail SCM,'' Proc. Comput. Sci., vol. 200, pp. 993–1003, Jan. 2022. [Online]. Available: https://www.sciencedirect. com/science/article/pii/S1877050922003076

[8] N. Mishra and A. Jain, ''Time series data analysis for forecasting—A literature review,''Int. J. Mod. Eng. Res., vol. 4, no. 7, pp. 1–5, 2014.

[9] C. Luo, J.-G. Lou, Q. Lin, Q. Fu, R. Ding, D. Zhang, and Z. Wang, ''Correlating eventswith time series for incident diagnosis,''in Proc. KDD. New York, NY, USA:Association for Computing Machinery, Aug. 2014, pp. 1583–1592.

[10] C. C. Ihueze and U. O. Onwurah, ''Road traffic accidents prediction modelling: An analysis of Anambra State, Nigeria,'' Accident Anal. Prevention, vol. 112, pp. 21–29, Mar. 2018.

[11] M. Sangare, S. Gupta, S. Bouzefrane, S. Banerjee, and P. Mühlethaler, ''Exploring the forecasting approach for road accidents:Analytical mea sures with hybrid machine learning,'' Expert Syst. Appl., vol.