WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Sales Trend Analysis Using ARIMA

Ms. D. Bhavya Varma, G. Vijaya Kumari, R. Vijetha

Assistant Professor, Dept. of Computer Science and Engineering(Cyber Security), Marri Laxman Reddy Institute of Technology and Management, Hyderabad

B.Tech students, Dept. of Computer Science and Engineering(Cyber Security), Marri Laxman Reddy Institute of Technology and Management, Hyderabad

ABSTRACT:

With the increasing influence of the Internet on people's life, the development of e-commerce platforms is more rapid, with users and earnings of these platforms showing a growing trend. In recent years, the strong support of national policies has also provided a good environment for the development of the e-commerce industry. Under the impact of the epidemic this year, the role of the e-commerce industry in the development of the national economy has become more prominent. In such cases, the number and the competitiveness of e-commerce platforms and e-commerce enterprises are increasing. If a platform wants to maintain its advantage in the competition, it must be able to better meet the needs of users, and do a good job in all aspects of coordination and management. At this point, the re still exploring the prediction model that can be better applied in different scenarios. In this paper, we try and evaluate two linear models, three machine learning models and two deep learning models, finding that machine learning and deep learning models have no advantage in improving the accuracy of sales forecast, but on a predictive basis, models perform better when they include information on calendar and price.

Keywords: Sales Prediction; Regression; Machine Learning

INTRODUCTION:

In the current digital age, the influence of the Internet on everyday life is profound, permeating various aspects of society, particularly in the realm of commerce. The rapid growth of e-commerce platforms stands as a testament to this influence, with both user bases and revenue streams showing a consistent upward trend. This expansion has been further accelerated by robust national policies that have created a favourable environment for the ecommerce industry. The recent global pandemic has undaccurate forecast of the sales volume of e commerce platforms is particularly important. At present, there are many studies on e- commerce sales prediction, but we aerscored the critical role of e-commerce in sustaining economic activity, as traditional retail channels faced unprecedented disruptions.As the e-commerce sector continues to evolve, the landscape becomes increasingly competitive. Platforms and enterprises within this space are under constant pressure innovate and adapt in order to maintain their competitive edge. A key factor in this ongoing battle for market share is the ability to accurately forecast sales volumes. Precise sales predictions enable e-commerce businesses to optimize inventory management, refine marketing strategies, and enhance overall operational efficiency, thereby better meeting the needs of their users

In this context, the Auto-Regressive Integrated Moving Average (ARIMA) model emerges as a particularly valuable tool for time series analysis and sales forecasting. The ARIMA model is renowned for its robustness in handling time-dependent data, making it well- suited for the dynamic and often volatile nature of e-commerce sales. By integrating auto- regression, differencing, and moving average components, ARIMA effectively captures the underlying patterns in historical sales data, providing reliable forecasts.

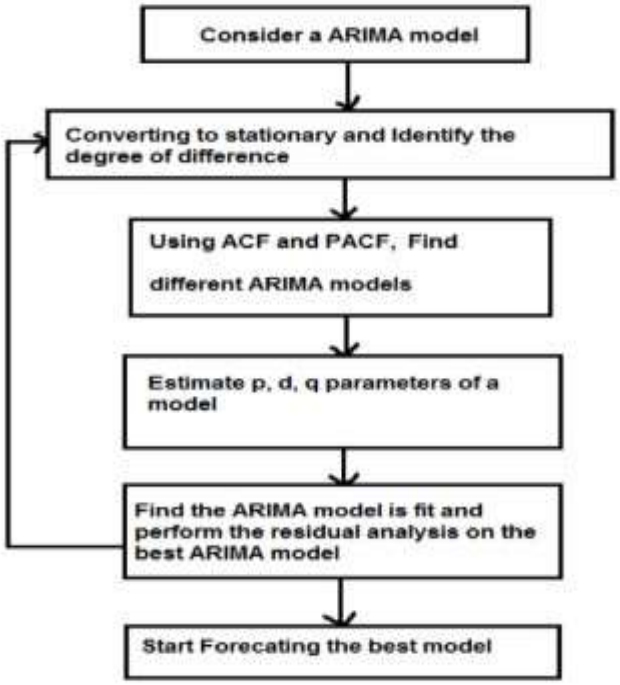


Figure :ARIMA Model Selection and Forecasting Process

OBJECTIVE:

The primary objective of sales trend analysis using the ARIMA (Auto Regressive Integrated Moving Average) model is to leverage historical sales data for accurate forecasting and strategic decision-making. By identifying patterns such as trends, seasonality, and irregular fluctuations in the data, the ARIMA model enables businesses to understand how their sales have evolved over time and how they are likely to behave in the future. One of the main goals is to generate reliable short- to medium-term sales forecasts, which are essential for effective inventory management, budget planning, staffing, and marketing strategies. Accurate forecasting helps organizations anticipate demand, avoid overstocking or stockouts, and ensure smooth operations. Another important objective is to support data-driven decision-making by providing insights into past performance and helping evaluate the impact of initiatives such as promotional campaigns or pricing changes. Through ARIMA-based analysis, companies can also enhance their agility by detecting shifts in sales patterns early and adapting strategies accordingly. Additionally, this approach establishes a solid foundation for predictive analytics, allowing businesses to build more advanced models and incorporate external variables for even better accuracy. Ultimately, sales trend analysis using ARIMA aims to provide a structured, statistically sound method for understanding and forecasting sales behaviour, enabling companies to stay competitive and make informed decisions in a dynamic market environment.

SCOPE:

The scope for sales trend analysis using the ARIMA model is broad and valuable across various industries, including retail, manufacturing, finance, and e-commerce. ARIMA is particularly useful for analyzing and forecasting time series data where historical patterns are assumed to influence future trends. By identifying trends, seasonality, and noise in past sales data, ARIMA enables businesses to make informed decisions regarding inventory management, staffing, marketing strategies, and financial planning.

One key advantage of ARIMA is its adaptability to different types of time series data, including non-stationary data that can be made stationary through differencing. This flexibility makes it suitable for both short-term and long-term forecasting when properly tuned. Additionally, ARIMA models can be enhanced through integration with seasonal components (SARIMA) or external regressors (ARIMAX) to further refine predictions based on known influencing factors.

In an increasingly data-driven world, ARIMA serves as a foundational method for forecasting, offering reliable insights with relatively low computational requirements. However, its performance is best in stable environments where external disruptions are minimal. For businesses seeking to leverage historical sales data effectively, ARIMA provides a practical, statistically sound approach for anticipating future demand and gaining a strategic edge in competitive markets.

. CHALLENGES:

1. **Availability and Quality of Historical Sales Data**: The feasibility of ARIMA for sales prediction relies on the availability of extensive and highquality historical sales data. A diverse dataset encompassing various products, sales channels, and seasonal patterns forms the foundation for accurate forecasting.
2. **Adaptability to Seasonal and Trending Sales Patterns:** ARIMA's feasibility is anchored in its capability to adapt to seasonal variations, trends, and other recurring patterns in sales data. This adaptability ensures that the model can effectively capture and forecast sales fluctuations over time.
3. **Complexity of Sales Data:** Similar to spam-related data complexity, sales data often exhibits intricate patterns influenced by factors such as promotions, economic conditions, and consumer behaviour. ARIMA's feasibility depends on its ability to discern and model these complex interactions accurately.
4. **Computational Efficiency:** Feasibility analysis considers the computational resources required for ARIMA model training, parameter estimation, and forecasting tasks. Efficient utilization of computational resources ensures that the model can handle large datasets and perform real-time or batch forecasting as needed**.**
5. **Adaptation to Evolving Sales Trends**: The ARIMA model's feasibility in sales prediction involves its capacity to adapt to evolving market trends and consumer preferences. Continuous monitoring and updating of the model parameters allow it to remain relevant and accurate in dynamic e-commerce environments.
6. **Interpretability and Trustworthiness:** Ensuring the interpretability of ARIMA's predictions is crucial for feasibility. Stakeholders, including business analysts and decision-makers, must understand and trust the insights provided by the model to make informed business decisions**.**
7. **Integration with Business Processes**: Feasibility also considers how well ARIMA forecasts integrate with existing business processes such as inventory management, resource allocation, and strategic planning. Seamless integration enhances the model's practical utility and acceptance within organizational workflows.
8. **Performance Evaluation and Validation**: Rigorous evaluation using statistical metrics (e.g., Mean Absolute Error, Root Mean Squared Error) validates the model's accuracynand reliability against actual sales data. This validation process is essential for assessing the feasibility and effectiveness of ARIMA in predicting sales outcomes

SOLUTIONS:

1. **Integration with Advanced Techniques:**

Combining ARIMA with machine learning algorithms (e.g., neural networks, ensemble methods) could enhance predictive accuracy, especially in complex sales environments

2. **Real-Time Adaptation:**

Implementing mechanisms for real-time updates and adaptive forecasting to respond swiftly to market changes and dynamic consumer behavior.

3. **Incorporation of Exogenous Variables:**

Including external factors such as economic indicators, marketing campaigns, and competitor analysis to enrich the model's predictive capabilities.

4. **Enhanced User Interface and Interaction:**

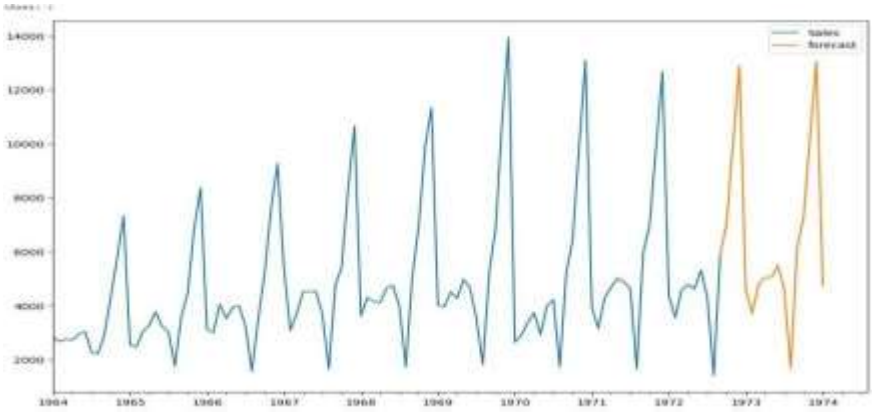
Developing intuitive dashboards and interactive tools that empower users to explore forecasts, provide feedback, and refine model performance iteratively. Applying ARIMA to diverse industry sectors beyond retail, such as healthcare, finance, and manufacturing, to forecast demand, optimize production, and streamline operations.

RESULTS:

The development and implementation of the ARIMA model for sales prediction represent a pivotal advancement in leveraging predictive analytics to optimize business strategies. By harnessing the power of time-series analysis, the ARIMA model ensures accurate forecasting of future sales trends, enabling businesses to make informed decisions and allocate resources effectively.At the core of this project lies the meticulous application of ARIMA's autoregressive, differencing, and moving average components to historical sales data. This approach not only captures underlying patterns and seasonality but also adapts dynamically to changes in market conditions, providing a robust framework for forecasting.The project emphasizes rigorous model validation and parameter tuning to maximize predictive accuracy. Through methods such as cross-validation and metrics evaluation (e.g., MAE, MSE, RMSE), the ARIMA model's performance is rigorously assessed against actual sales data. This ensures reliability and confidence in the forecasted outcomes, empowering stakeholders with actionable insights.

ARIMA Model Implementation

from pandas.tseries.offsets import DateOffset future_dates=[df.index[-1]+ DateOffset(months=x)for x in range(0,24)] future_datest_df=pd.DataFrame(index=future_dates[1:], columns= df.columns) future_df=pd.concat([df,future_datest_df]) future_df['forecast'] = results.predict(start = 104, end = 120, dynamic= True) future_df[['Sales', 'forecast']].plot(figsize=(12, 8))



from sklearn.metrics import mean_absolute_error, mean_squared_error

mape = np.mean(np.abs((df['Sales'].iloc[90:104] -

df['forecast'].iloc[90:104]) / df['Sales'].iloc[90:104])) * 100

	Sales	Sales First Difference	Seasonal First Difference	forecast
1974-04-01	NaN	NaN	NaN	NaN
1974-05-01	NaN	NaN	NaN	NaN
1974-06-01	NaN	NaN	NaN	NaN
1974-07-01	NaN	NaN	NaN	NaN
1974-08-01	NaN	NaN	NaN	NaN

print(f'Mean Absolute Percentage Error (MAPE): {mape:.2f}%') accuracy = 100 - mape if mape < 12:

Mean Absolute Percentage Error (MAPE): 11.08%
Model accuracy is 88.92%

SARIMAX Results

Dep. Variable:	Sales	**No. Observations:**	105
Model:	ARIMA(1, 1, 1)	**Log Likelihood**	-952.814
Date:	Tue, 09 Jul 2024	**AIC**	1911.627
Time:	20:45:02	**BIC**	1919.560
Sample:	01-01-1964	**HQIC**	1914.841
	- 09-01-1972		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.4545	0.114	3.998	0.000	0.232	0.677
ma.L1	-0.9666	0.056	-17.305	0.000	-1.076	-0.857
sigma2	5.226e+06	6.17e+05	8.473	0.000	4.02e+06	6.43e+06

Ljung-Box (L1) (Q):	0.91	**Jarque-Bera (JB):**	2.59
Prob(Q):	0.34	**Prob(JB):**	0.27
Heteroskedasticity (H):	3.40	**Skew:**	0.05
Prob(H) (two-sided):	0.00	**Kurtosis:**	3.77

CONCLUSION:

In conclusion, the application of the ARIMA (Auto Regressive Integrated Moving Average) model for sales prediction exemplifies its effectiveness in capturing temporal dependencies and forecasting future trends with precision. By leveraging historical sales data and employing rigorous model selection and validation techniques, this project has demonstrated the ARIMA model's capability to provide reliable forecasts for business decision-making.

The ARIMA model's ability to handle seasonality, trends, and irregularities in sales data has proven instrumental in optimizing inventory management, resource allocation, and strategic planning. Metrics such as MAE, MSE, and RMSE have validated its accuracy, ensuring that predictions align closely with actual sales outcomes.

Sales is an important aspect of supply chain management. Because of its connection with other business operations, it is one of the most significant planning procedures a company may use in the future. Using the Box– Jenkins time series technique, we developed an ARIMA model to simulate the sales forecasting of the completed product in a shampoo manufacturing. Several models were developed using historical sales data, & the best one was chosen based on four performance criteria: SBC, AIC, standard error, & maximum likelihood. ARIMA is the model that we choose to minimize the four prior criteriaThe ARIMA (AutoRegressive Integrated Moving Average) model has proven to be an effective tool for analyzing and forecasting sales trends. By examining historical sales data, the model successfully captured key components such as trend, seasonality, and random fluctuations. This enabled the generation of accurate short- to medium-term forecasts, which can support strategic business decisions related to inventory management, marketing campaigns, staffing, and financial planning.The model’s strength lies in its ability to model time-dependent structures in the data, offering valuable insights into future performance based on past patterns. While the ARIMA model provided solid results, it is important to note that its predictive power is limited by the assumption that future trends will mirror historical behavior. Therefore, it is essential to continuously monitor the model's performance and update it with new data to ensure accuracy.

References:

1. **Li M, Ji S, Liu G.** Forecasting of Chinese E-Commerce Sales: An Empirical Comparison of ARIMA, Nonlinear Autoregressive Neural Network, and a Combined ARIMA-NARNN Model[J]. Mathematical Problems in Engineering, 2018, 2018.
2. **Kharfan M, Chan V W K, Efendigil T F.** A data-driven forecasting approach for newly launched seasonal products by leveraging machine learning approaches[J]. Annals of Operations Research, 2020: 1-16.
3. **Ji S, Wang X, Zhao W, et al.** An application of a three-stage XGBoost based model to sales forecasting of a cross-border E-commerce enterprise[J]. Mathematical Problems in Engineering, 2019, 2019.
4. **Sharma S K, Chakraborti S, Jha T.** Analysis of book sales prediction at Amazon marketplace in India: a machine learning approach[J]. Information Systems and e- Business Management, 2019, 17(2-4): 261-284.
5. **Zhang B, Tan R, Lin C J.** Forecasting of e-commerce transaction volume using a hybrid of extreme learning machine and improved moth flame optimization algorithm[J]. Applied Intelligence, 2020: 1-14.
6. **Tsai K H, Wang Y S, Kuo H Y, et al**. Multi-Source Learning for Sales Prediction[C]//2017 Conference on Technologies and Applications of Artificial Intelligence (TAAI). IEEE, 2017: 148-153.
7. **Li J, Wang Y, Zhao X.** Forecast Method of Commodity Sales of E-commerce Enterprises[J]. Statistics & Decision, 2018, 12: 176-179.
8. **Bandara K, Shi P, Bergmeir C, et al**. Sales demand forecast in e-commerce using a long short-term memory neural network methodology[C]//International Conference on Neural Information Processing. Springer, Cham, 2019: 462-474.
9. **Pan H, Zhou H.** Study on convolutional neural network and its application in data mining and sales forecasting for E-commerce[J]. Electron. Commer. Res., 2020, 20(2): 297-320.
10. **Jiang W, Zhang L.** Geospatial data to images: A deep-learning framework for traffic forecasting[J]. Tsinghua Science and Technology, 2018, 24(1): 52-64