

AI-Driven Business Analytics System for Sales, Customer Insights, and Forecasting

Author Details:

Barath L¹, Pooja D¹, Suganthavanathan M¹, Dr. Perumal S.⁴


^{1 2 3} UG Scholar, Department of Computer Science and Information Technology, Vels Institute of Science, Technology & Advanced Studies VISTAS, Chennai, India [barath272005@gmail.com]¹, [poojachoyal20@gmail.com]², [kstm107@gmail.com]³

⁴ Professor, Department of Computer Science and Information Technology, Vels Institute of Science, Technology & Advanced- Studies VISTAS, Chennai, India



<https://doi.org/10.55041/ijst.v2i4.528>

Cite this Article: L, B., M, S. & D, P. (2026). AI-Driven Business Analytics System for Sales, Customer Insights, and Forecasting. International Journal of Science, Strategic Management and Technology, 02(04). <https://doi.org/10.55041/ijst.v2i4.528>

License:  This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

I. Abstract—

In the current digital era, organizations continuously produce large amounts of data through transactions, customer engagements, and online platforms. Transforming this raw data into meaningful information is essential for gaining business value. Business analytics enables this transformation by identifying underlying patterns, trends, and relationships within the data. Traditional reporting methods often struggle with data inconsistency, increasing volumes, and limited analytical capabilities.

Artificial Intelligence has become a key enabler in improving business decision-making processes. It allows systems to efficiently process large datasets and identify complex patterns that might be missed by conventional tools. This project introduces a comprehensive, AI-driven business analytics solution that combines sales prediction, customer behavior analysis, and forecasting into a single unified platform. The primary objectives are to apply data preprocessing techniques, implement customer segmentation using clustering, develop forecasting models

using regression, and integrate AI-based decision support through an automated reporting dashboard and chatbot.

II. LITERATURE REVIEW

Extensive research has explored the integration of AI within business intelligence. Radhakrishnan [1] presented an AI-driven framework aimed at forecasting market trends and understanding consumer behavior. While effective for predictive modeling, it lacked efficient real-time visualization tools. Goswami [2] emphasized the application of clustering and classification algorithms for customer segmentation and personalized marketing; however, it lacked a predictive component for forecasting sales trends. Zamil [3] evaluated evaluating various decision support models for financial forecasting but highlighted challenges regarding data integration complexity and the absence of unified frameworks combining financial results with customer insights.

From the existing literature, a significant gap exists regarding a holistic system that processes preprocessing,

segmentation, sales forecasting, and interactive AI querying all on a single scalable platform. The proposed system fills this gap by implementing an integrated application that handles end-to-end data pipelines for enterprise analytics.

Keywords—Business Analytics, Customer Segmentation, Data Mining, K-Means Clustering, Linear Regression, Streamlit.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed AI-Driven Business Analytics System follows a modular and layered architecture. It consists of multiple functional components that execute specialized analytical tasks. User input is provided as structured datasets (CSV or Excel). The data follows a pipeline: it is ingested, cleaned, and transformed by the Data Preprocessing Module. The cleaned datasets are passed into distinct analytical modules.

The architecture incorporates a Sales Analytics Module for descriptive analysis, a Customer Analytics Module for RFM clustering, a Product Performance Analyzer, and a Sales Forecasting Module. All results are seamlessly visualized through a central dashboard powered by Streamlit and Plotly, alongside an AI Chatbot and a PDF Report Generator for business recommendations.

IV. METHODOLOGY

A. Data Preprocessing

Data accuracy relies heavily on robust preprocessing. The system computes descriptive statistics, manages missing values (mean/median/mode imputation), removes duplicates, standardizes date formats, and rectifies outliers using the Interquartile Range (IQR) method before the dataset is forwarded for model training.

B. Customer Segmentation (K-Means Clustering)

To group customers according to their behaviors and revenue contributions, a K-Means clustering algorithm was applied. Customers were first mapped to Recency, Frequency, and Monetary (RFM) profiles based on their transaction history. Due to scale variance, the data was normalized using StandardScaler. Users can dynamically visualize different cluster counts to identify high-value, regular, and low-value customers.

C. Sales Forecasting (Linear Regression)

A Linear Regression model was developed to predict future sales trends. Historical sales data aggregated monthly served as the input features. The dependent variable was the sales value while the time indices acted as predictors. The model generates short-to-medium term forecasts, complete with a $\pm 10\%$ confidence interval to highlight variance, and is evaluated using metrics like Mean Absolute Error (MAE) and R^2 Scores.

V. RESULTS AND DISCUSSION

The system provides detailed interactive results and metrics for users. The Dashboard delivers immediate insights into Key Performance Indicators (KPIs), average sales totals, and region-wise distributions. The Sales Analytics module visually outputs product performance benchmarks, monthly revenue trends, and growth rates using Plotly graphics.

The Customer Analytics module successfully segmented user profiles into distinct purchasing classes, revealing potential avenues for personalized marketing. Furthermore, the Forecast Module accurately plotted predicted revenues against historical baselines, yielding actionable planning resources.

In conjunction with predictive metrics, the AI Decision Support Module dynamically processes outputs to suggest strategies, and the integrated AI Chatbot allows non-technical management staff to execute complex queries and export standard results using the Report Generator.

VI. FUTURE ENHANCEMENTS

Future development will prioritize connecting the platform to real-time external APIs, databases, and enterprise data streams without needing manual CSV uploads. More sophisticated time-series models (e.g., ARIMA, LSTM) will be implemented to augment the linear regression approach for complex seasonality. Cloud deployment on AWS/Azure will allow scalable organizational access, alongside stricter role-based user protocols.

VII. CONCLUSION

The Integrated Business Analytics System serves as an innovative integration of data mining models and interactive visualization libraries to extract actionable



business insights from raw transactions. Through algorithms like K-Means clustering and Linear Regression, the system effectively automates customer segmentation and trend forecasting. The unified Streamlit-driven application ultimately empowers real-time data interpretation, thereby eliminating the inefficiencies of traditional manual reporting and facilitating competitive, data-driven enterprise governance.

REFERENCES

- [1] Radhakrishnan. AI-driven business intelligence framework for market trend forecasting.
- [2] Goswami. AI-driven business intelligence systems for customer data insights.
- [3] Zamil. AI-driven business analytics models for financial forecasting.
- [4] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann, 2011.
- [5] P. N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining, 2nd ed. Pearson Education, 2019.
- [6] A. Geron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd ed. O'Reilly Media, 2022.
- [7] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.