**LITERATURE REVIEW**

Enhancing E-Commerce Decision-Making with an Online Retail Prediction Model

1. Introduction

E-commerce businesses generate vast amounts of transaction data daily, creating opportunities to leverage machine learning and analytics for strategic decision-making. Numerous studies and projects have explored predictive modeling in e-commerce, particularly in forecasting sales, analyzing customer behavior, and optimizing inventory. This literature review examines existing works on online retail prediction models and highlights how this project differs by focusing on unique methodologies, feature engineering techniques, and business-driven insights.

2. Existing Work on E-Commerce Prediction Models

Several machine learning models have been applied in past studies to predict key e-commerce metrics. A commonly referenced project is available on Kaggle, which utilizes the Online Retail Dataset to perform customer segmentation and sales trend analysis. The Kaggle project primarily focuses on Recency, Frequency, and Monetary (RFM) analysis to categorize customers and employs clustering algorithms like K-Means to segment buyers.

Another study explored time-series forecasting using ARIMA and LSTM models to predict sales patterns over time. Additionally, researchers have implemented association rule mining (e.g., Apriori Algorithm) to uncover frequently bought-together products, helping businesses optimize their marketing strategies.

3. Differentiation from Existing Studies

While previous projects have successfully analyzed customer segmentation and sales trends, my project extends beyond conventional approaches in the following ways:

Predictive Modeling for Business Decision-Making: Unlike many existing studies that focus on descriptive analytics, my model is designed to provide actionable forecasts that can guide decision-making. I aim to predict future sales, demand fluctuations, and customer purchasing behavior using machine learning models such as Random Forest, Gradient Boosting, and XGBoost.

Feature Engineering and Data Enrichment: I am incorporating additional derived features, including time-based aggregations (weekly/monthly trends), customer lifetime value (CLV), and product popularity metrics. This approach provides a deeper understanding of customer purchasing patterns beyond traditional RFM analysis.

Handling Data Challenges Effectively: Many previous studies do not adequately address real-world data issues such as missing values, duplicate transactions, and outliers. My project places emphasis on thorough data cleaning and preprocessing to improve model reliability and accuracy.

Comparative Model Analysis: Unlike many Kaggle projects that focus on a single modeling technique, I aim to compare multiple machine learning approaches, evaluating their effectiveness based on metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

Application-Oriented Approach: My study is designed with practical business implications in mind, helping e-commerce businesses make informed inventory management and marketing decisions. The end goal is not just to analyze past transactions but to provide real-time predictive insights that can drive future strategies.

4. Conclusion

This project builds upon previous studies by incorporating advanced machine learning models, extensive feature engineering, and a focus on real-world business applications. Unlike prior work, which primarily focuses on exploratory data analysis and customer segmentation, my approach leverages predictive modeling to generate actionable insights for e-commerce decision-makers. By addressing key data challenges and evaluating multiple predictive techniques, this study aims to develop a more robust and practical online retail prediction model that goes beyond existing implementations.