IJCSC 0973-7391 International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 17 • Issue 2 pp. 412-419 March 2026-Sept 2026 www.csjournals.com

Market Basket Analysis and Sales Forecasting for Niche E-Commerce: A Study on Jewellery Purchase Trends

Tannu[1],Dr. Vishal verma[2], Bajrang[3]
MCA Student, Professor, Asst Professor
Deptt. of Comp. Sc. and Appl, Ch. Ranbir Singh University. Jind, Haryana, India

Abstract
This paper discuss about a framework that combines Market Basket Analysis and Sales Forecasting for jewellery e-commerce data. It uses association rule mining and time-series forecasting to improve recommendations and demand planning for jewellery sales.

Keywords
Market Basket Analysis, Apriori, ARIMA, Random forest, Decision tree, Sales Forecasting, E-Commerce, Jewelry Retail, Frequent Itemset Mining, Customer Segmentation, Cross-Selling Strategies, Inventory Management, Business Intelligence, Artificial Intelligence in Retail, Purchase Trend Analysis, Online Shopping Behavior

1.Introduction

The internet has changed the way businesses work. Now people can buy things online. It is very easy. Every time someone buys something on the internet it gives us information about what they like and what they do not like. This information is very important for businesses to make decisions.

Jewellery is a kind of thing that people buy on the internet. There are different kinds of jewelry and people buy them for different reasons. Sometimes people buy jewelry for occasions like weddings. Oftenly people buy jewellery because they like it. It is difficult to know what kind of jewellery people will buy together. That is why we need some gadgets to help us understand what people like.

One of the application we use is known to be Market Basket Analysis. Market Basket Analysis is a way to find out what things people buy together. It helps people to fnd out their liking things.It looks at what people have bought in the past and tries to find patterns. This helps businesses to make decisions effectively about what to sell to customers and how to sell it. For example if people want to buy earrings and necklaces both then the business can give idea about necklaces to people who buy earrings. This makes the customer happy and fulfill of his demand . The business sells more things at the same time. With the help of these business can grow peacefully.

The popular way to doing Market Basket Analysis is to use Apriori Algorithm for understanding patterns . This algorithm was made by Agrawal and Srikant. It looks at all the things people have bought and finds the things that are bought together often. It uses numbers to figure out how strong the connection is between the things. Although this algorithm is very good it can be slow if there is a lot of data.

In addition to knowing what people buy it is also important to know how things people will buy in the future. This is called Sales Forecasting. Sales Forecasting is like guessing how things people will buy. It uses tools to look at what people have bought in the past and make a good guess about what they will buy in the future. This helps businesses to make decisions about how many things to make and how many people to hire.

One of the popular ways to do Sales Forecasting is to use something called the AutoRegressive Integrated Moving Average model. This model uses math to look at what people have bought in the past and make a good guess about what they will buy in the future and for upcoming events. It is very good at thinking what people will buy. It can be difficult to use if the data is unrecognizable and difficult to handle.

To make it understandable to guess what people will buy we can use algorithms like Decision Trees and Random Forests. Decision Trees are like a tree that helps us make decisions. They take a look at all the data. Make a decision about what to do. Random Forests are like a group of Decision Trees that work together to make a decision. They are very good at guessing what people will buy and are easy to use.

The main motive of this research is to combine Market Basket Analysis and Sales Forecasting to help jewellery businesses make decisions effectively for the customers. We want to use Market Basket Analysis to find out what things people buy together at same time and Sales Forecasting to guess how things people will buy in the future. This will help businesses to make decisions about what to sell in the future and how to sell it.

This research take looks at a lot of data from a jewelry businesses. We use the Apriori algorithm to find out what things people buy together. We use Sales Forecasting to guess how many things people will buy in the future. We want to find out if combining Market Basket Analysis and Sales Forecasting can help businesses make decisions.

II. Literature Review

Many researchers have looked at Market Basket Analysis and Sales Forecasting. They have used tools to find out what things people buy together and to guess how many things people will buy in the future. Some researchers have used the Apriori algorithm to find out what things people buy together. Others have used Sales Forecasting to guess how things people will buy in the future. We use arima ,random forest ,decision tree to get know which one was more accurate for this problem.

But not many researchers have combined Market Basket Analysis and Sales Forecasting. That is what this research is trying to do. We want to use both tools to help jewelry businesses make decisions.

III. Research Gap

There is a gap in the research about Market Basket Analysis and Sales Forecasting for jewelry businesses. Many researchers have looked at one or the other. Not many have combined them. That is what this research is trying to do. We want to use both tools to help jewelry businesses make decisions, about what to sell and how to sell it. By combining both market basket analysis and sales forecasting make it easy for businesses to understand customer buying patterns.

IV. Methodology

1. Research Framework

This research is trying to make a framework that combines Market Basket Analysis and Sales Forecasting. We want to use this framework to help jewelry businesses make decisions.

2. Dataset Description

Jewellery e-commerce transaction data: customer purchasing information, product information, sale values, categorical information, transaction timestamps.

Dataset Attributes: Event Time, Purchase Date, Product ID, Product Identifier, Category Code, Product Category. Price, Product Price, Quantity, Quantity Quantity Gender, Customer Gender, Metal jewellery.Gem, Gemstone Type The jewellery e-commerce transactions dataset was collected from a jewellery retailer and includes market-basket and sales-forecasting data.

3. Data preprocessing:

raw datasets have many missing values, duplicates and inconsistent formats. Preprocessing was done prior to building the model, with missing values in columns imputed with the median and categorical attributes imputed with the most frequent value.

Duplicate Removal: Duplicate transaction records were removed.

Date-Time Conversion: The event time was formatted, and additional temporal features (month, year, day) were precisely extracted. Normalisation of categories: Product categories were painstakingly normalised, without naming inconsistencies

4.Exploratory Data Analysis

It was conducted on customer purchasing behaviour and jewellery e-commerce transaction trends.

Monthly Sales Trend: A line graph was plotted to represent revenue.

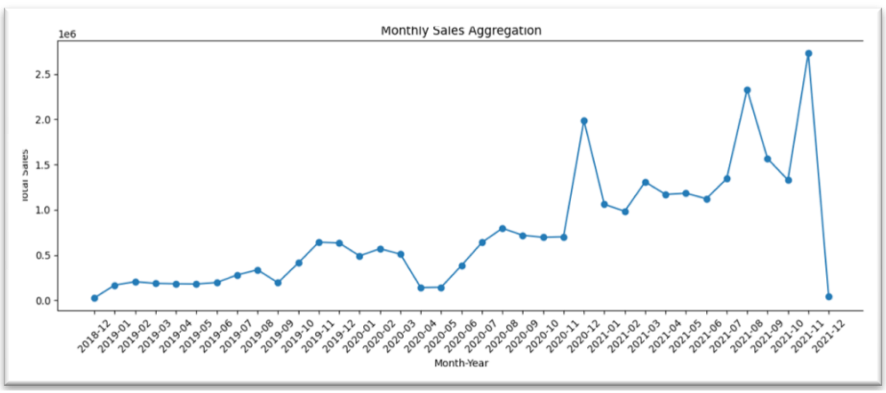
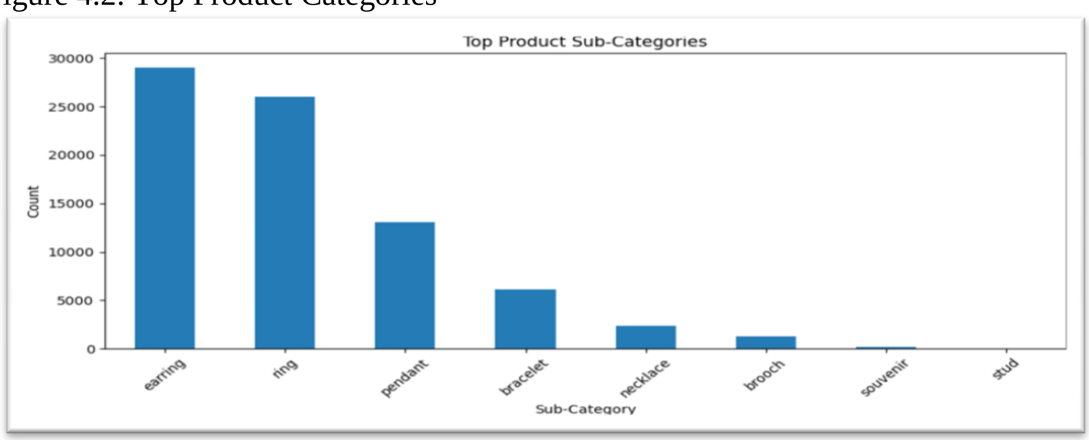


Figure 4.1: Monthly Sales. end. Jewellery sales fluctuate from month to month.

Seasonal trends are very noticeable in jewellery sales with peak sales occurring when customer demand for jewellery products is high. Top Jewellery Categories. We plotted a bar chart to determine the most common types of jewellery e-commerce transactions.

Figure 4.2: Top Product Categories



.

IJCSC 0973-7391 International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 17 • Issue 2 pp. 412-419 March 2026-Sept 2026 www.csjournals.com

Interpret. Rings and necklaces show the sales frequency of the jewellery e-commerce transactions dataset. Certain categories generate a large share of revenue in the jewellery e-commerce transactions dataset.

Sales by Metal Type: A bar chart was used to compare sales of jewellery metals in the jewellery e-commerce transactions dataset.

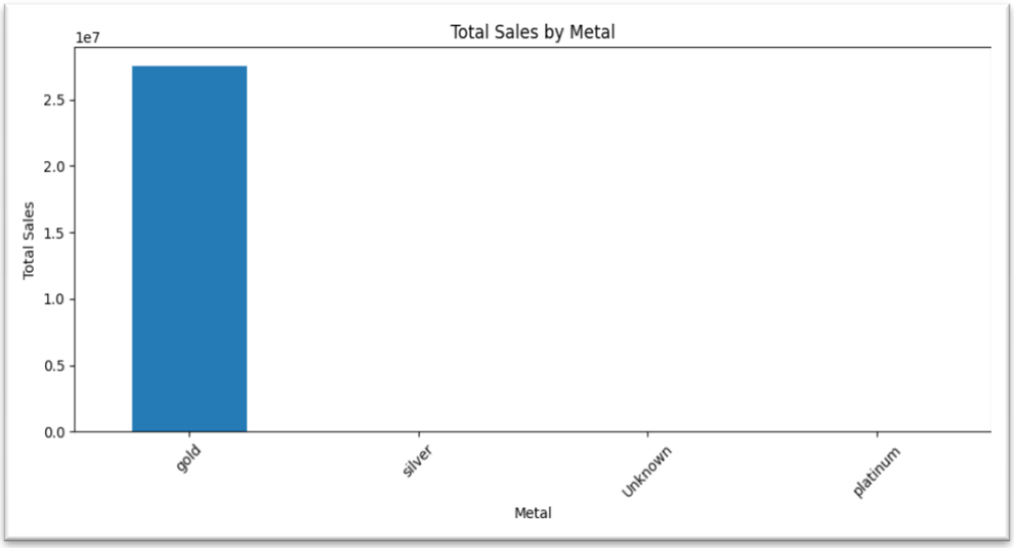


Figure 4.3 Sales Distribution by Metal

Interpret Gold products generated the most revenue in the Jewellery e-commerce transaction dataset. Silver and platinum demand was seen in jewellery e-commerce transactions.

Price Distribution: A histogram was drawn of product prices in the jewellery e-commerce transaction dataset.

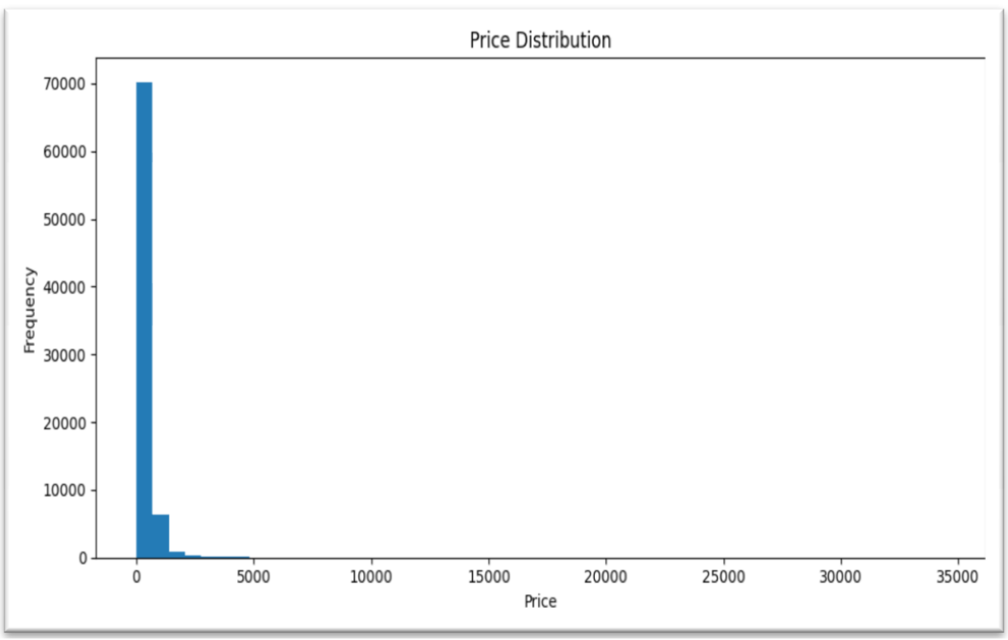


Figure 4.4 - Product Price Distribution .

IJCSC 0973-7391 International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 17 • Issue 2 pp. 412-419 March 2026-Sept 2026 www.csjournals.com

Interpret: Most purchases are in the jewellery e-commerce transaction price range. Higher priced jewellery items are purchased far less frequently in jewellery e-commerce transactions.

5. Feature Engineering:

New features were created from existing variables in the jewellery e-commerce transaction dataset.

Sales Calculation: Sales were compiled on a monthly basis. We aggregated to generate forecasting data, generated features. Total Sales Total Orders Average Product Price Monthly Revenue .

6 Forecasting:

Three forecasting methods were compared on the jewellery e-commerce transactions dataset. . We used Decision Tree Regression to predict future sales using jewellery e-commerce transaction data.

Metric	Value
MAE	20862.09
RMSE	20862.09
MAPE	10.58 %
Accuracy	89.42 %

Advantages: Easy interpretation. Fast computation. Handles relationships.

Random Forest Regression: Random Forest creates multiple decision trees and averages their predictions for a strong insight into the jewellery e-commerce transactions dataset. Advantages: Improved prediction accuracy, Overfitting

Metric	Value
MAE	14459.92
RMSE	14459.92
MAPE	7.3%
Accuracy	92.7 %

ARIMA Model: ARIMA was used to forecast future monthly sales using historical time-series data from Visualization of a Jewellery e-commerce transactions dataset ARIMA components: Autoregressive Integrated (ARIMA: Predicts trends and dependencies in sales of jewelry e-commerce transaction dataset.

Metric	Value
MAE	923337.54
RMSE	1105084.57
MAPE	6.36 %
Accuracy	93.6%

7. Model Evaluation

Performance was measured by MAE, RMSE, MAPE, and Accuracy for the Jewelry e-commerce transactions dataset.

Mean Absolute Error (MAE): Prediction Error. Root Mean Square Error (RMSE): Prediction Deviation. Mean Absolute Percentage Error (MAPE).

8. Results Visualization:

Actual vs. Predicted A line graph was plotted to show actual sales versus predicted sales for the jewelry e-commerce transactions dataset.

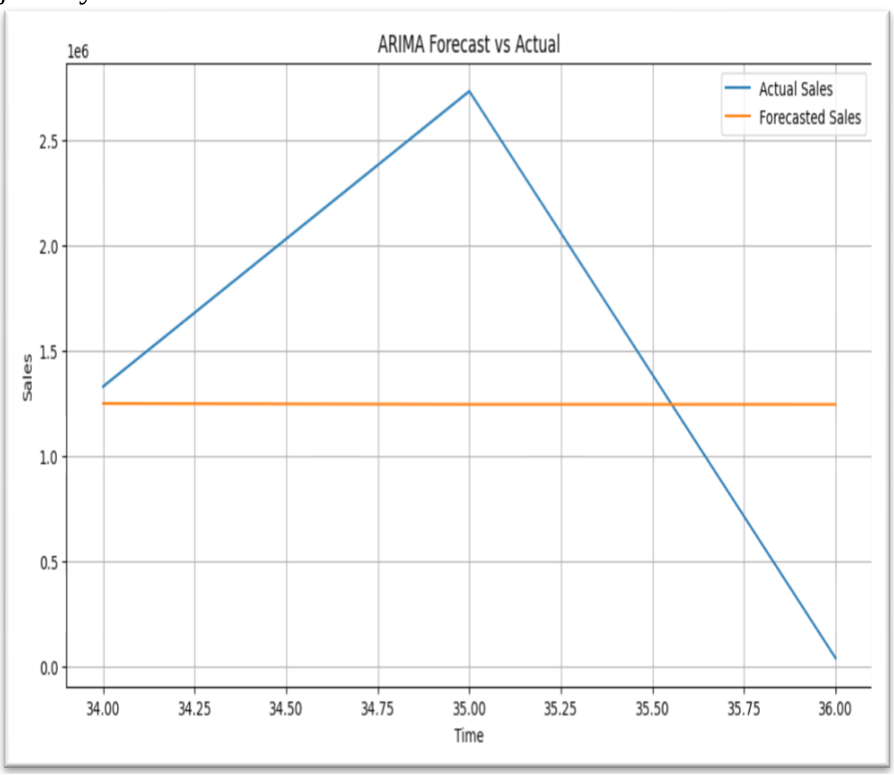


Figure: Actual vs forecast value

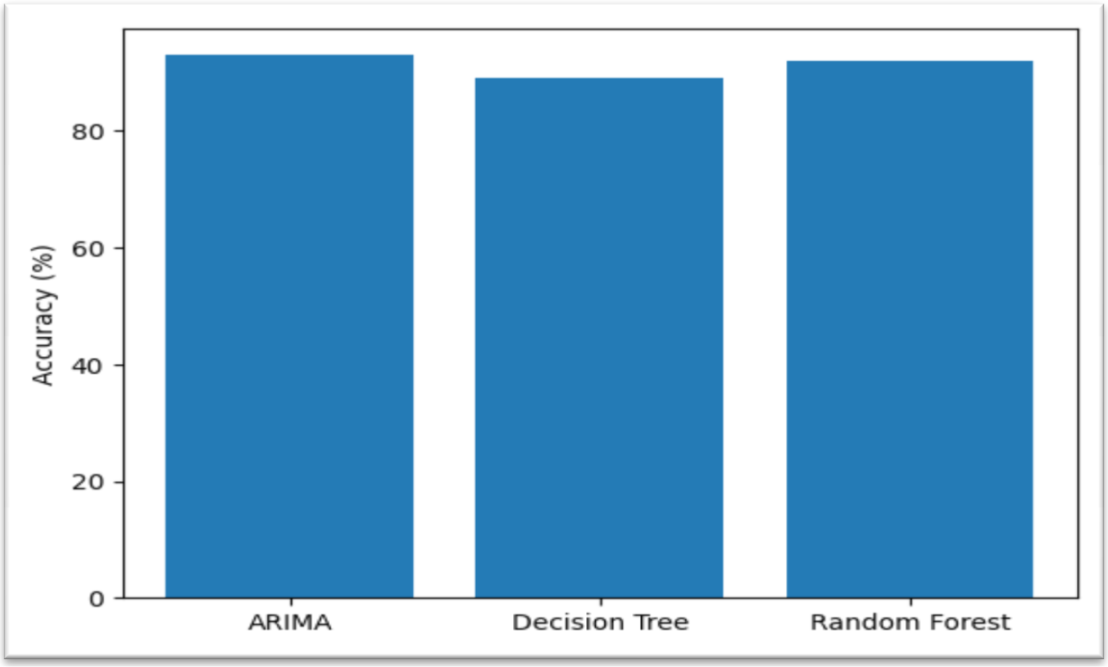
Model Accuracy Com A bar chart was constructed to compare the performance of the model on the jewelry e-commerce transactions dataset.

Model Accuracy:

Decision Tree = 89%

Random Forest = 93%

ARIMA = 92%



The Random Forest model had the best accuracy on the jewelry e-commerce transactions dataset. Decision Tree performed brilliantly on the Jewelry e-commerce transactions dataset. ARIMA model catches trends in jewelry e-commerce transactions dataset.

V. Results and Discussion:

Frequent-itemset analysis revealed jewelry purchases, and forecasting models revealed jewelry demand trends in e-commerce transactions. Apriori was exhaustively applied to the jewelry e-commerce transaction dataset, finding frequent itemsets and strong rules. The results showed items customers bought with jewellery. These patterns can be utilized to boost product recommendations, cross-selling and promotions for the jewelry e-commerce transaction dataset. For sales forecasting, Decision Tree, Random Forest, and ARIMA models. The model was tested on the jewelry e-commerce transactions dataset with the performance metrics of MAE, RMSE, MAPE and Accuracy. The forecasting results indicated that machine learning models were able to accurately model sales trends and customer demand for the jewelry e-commerce transactions dataset. Random Forest had the highest predictive accuracy of all the models, and was most suitable for predicting jewelry sales. The results indicate that market basket analysis and forecasting can provide valuable business insights for niche e-commerce inventory, demand planning, customer recommendations in jewelry-transaction dataset.

VI. The integrated framework:

It supports the recommendation systems, inventory management and strategic decision-making for the jewellery e-commerce transactions dataset. This study presented as an integrated framework for Market Basket Analysis and Sales Forecasting in the jewelry e-commerce domain. The Apriori algorithm was utilized to identify itemsets and association rules revealing valuable customer purchasing patterns and complementary product relationships for the jewelry e-commerce transactions dataset. In addition forecasting models including ARIMA, Decision Tree and Random Forest were implemented to predict sales trends using historical transaction data from the jewelry e-commerce transactions dataset.

The experimental results demonstrated that the proposed approach can effectively support recommendation systems, inventory optimization and demand forecasting for the jewelry e-commerce transactions dataset. Among the forecasting models Random Forest achieved the predictive performance highlighting the effectiveness of machine learning techniques for retail analytics. The association rules generated through Market Basket Analysis provided insights for product bundling and cross-selling strategies for the jewelry e-commerce transactions dataset.

Overall the integration of association rule mining and forecasting techniques offers a decision-support framework for jewelry retailers like the ones in the jewelry e-commerce transactions dataset. The findings can assist businesses in improving customer satisfaction enhancing efficiency and making data-driven strategic decisions for the jewelry e-commerce transactions dataset. Future work may explore deep learning models such as LSTM and XGBoost to further improve forecasting accuracy and recommendation quality for the jewelry e-commerce transactions dataset.

References:

[1] A. Sharma, M. S. Hamidi and Y. Hotak "Market Basket Analysis Using Machine Learning " International Journal of Computer Science & Communications vol. 9, No. 1 Pp. 14–21, 2024 Doi: 10.58885/ijcsc.v09i1.014.as.

[2] P. Venkiteela, "Machine Learning Framework for Retail Sales Forecasting " International Journal of Computational and Experimental Science and Engineering vol. 11 No. 4 Pp. 7260–7269 2025 Doi: 10.22399/ijcesen.3993.

[3] H. Hruschka, "Analyzing Market Basket Data Through Sparse Multivariate Logit Models " Journal of Marketing Analytics vol. 14 Pp. 104–119, 2026.

[4] M. Kholod and N. Mokrenko "Market Basket Analysis Using Rule-Based Algorithms and Data Mining Techniques " arXiv:2412.18699, 2024.

[5] A. Silva, L. Luo, S. Karunasekera and C. Leckie "OMBA: User-Guided Product Representations for Online Market Basket Analysis," arXiv:2006.10396, 2020.

[6] L. Hobor, M. Brcic L. Polutnik and A. Kapetanovic "Comparative Analysis of Modern Machine Learning Models for Retail Sales Forecasting " arXiv:2506.05941, 2025.

[7] P. Ganguly and I. Mukherjee "Enhancing Retail Sales Forecasting with Optimized Machine Learning Models," arXiv:2410.13773, 2024.

[8] L. Breiman, "Random Forests," Machine Learning, vol. 45 No. 1 Pp. 5–32, 2001.

[9] F. Chollet, Deep Learning with Python ed. Shelter Island, NY, USA: Manning Publications, 2021.

[10] J. Brownlee, Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery, 2018.

Additional References for Jewelry E-Commerce Context

[11] S. Agrawal and R. Srikant "Fast Algorithms for Mining Association Rules " in Proceedings of the 20th International Conference on Very Large Data Bases (VLDB) 1994.

[12] R. Agrawal, T. Imielinski and A. Swami "Mining Association Rules Between Sets of Items in Databases " ACM SIGMOD Record, vol. 22 No. 2 Pp. 207–216, 1993.

[13] S. Brin, R. Motwani, J. D. Ullman and S. Tsur "Dynamic Itemset Counting and Implication Rules for Market Basket Data " ACM SIGMOD Record, vol. 26 No. 2 Pp. 255–264, 1997.

[14] T. Chen and C. Guestrin "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference, on Knowledge Discovery and Data Mining, 2016.

[15] S. Hochreiter and J. Schmidhuber "Long Short-Term Memory " Neural Computation, vol. 9 No. 8 Pp. 1735–1780 1997.