Deep Learning-Based Market Basket Analysis Using Association Rules

Hamid Ghous[1],   Mubasher H. Malik[2],   Iqra Rehman[3]

Abstract

Market Basket Analysis (MBA) is a data mining technique assisting retailers in determining the customer's buying habits while making new marketing decisions as the buyer's desire frequently changes with expanding needs; therefore, transactional data is getting large every day. There is a demand to implement Deep Learning (DL) methods to manipulate this rapidly growing data. In previous research, many authors conducted MBA applying DL and association rules (AR) on retail datasets. AR identifies the association between items to find in which order the customer place items in the basket. AR is only used in mining frequently purchased items from retail datasets. There is a gap in classifying these rules and predicting the next basket item using DL on the transactional dataset. This work proposes a framework using AR as a feature selection while applying DL methods for classification and prediction. The experiments were conducted on two datasets, InstaCart and real-life data from Bites Bakers, which operates as a growing store with three branches and 2233 products. The AR classified at 80,20 and 70,30 splits using CNNN, Bi-LSTM, and CNN-BiLSTM. The results considering simulation at both splits show that Bi-LSTM performs with high accuracy, around 0.92 on the InstaCart dataset. In contrast, CNN-BiLSTM performs best at an accuracy of around 0.77 on Bites Bakers dataset.

Keywords: Market Basket Analysis, Deep Learning, Data Mining, Association Rules, Frequency Pattern Mining.

1.   Introduction

To meet consumers' expanding needs and compete in the market, retailers are interested in customer purchasing patterns [1]. The ability to choose different versions of the same product based on consumer preferences has been made possible by technological advancements, which have increased the frequency of product updates [2]. Due to the growing customer demand as more products are introduced to meet changing needs, the dataset is growing in size daily [3]. Hence, customers buy different items in a single store visit. Due to the large volume of products and customers, the traditional method takes more time to find purchasing behavior. So data mining methods must extract useful knowledge from data [4], [5]. Knowledge Data Discovery (KDD) is the series of iterative steps to discover beneficial, salient, and comprehend patterns from massive datasets [6]. Data mining is the crux of KDD, including techniques to understand and explore data,

[1]*Department of Computer Science, Institute of Southern Punjab, Multan, Pakistan (hamidghous@isp.edu.pk)*
[2]*Department of Computer Science, Institute of Southern Punjab, Multan, Pakistan (mubasher@isp.edu.pk)*
[3]*Department of Computer Science, Institute of Southern Punjab, Multan, Pakistan (iqrarehman@isp.edu.pk)*

augmenting meticulous models, and finding crucial patterns in extracting knowledge [7]. Knowledge discovered from the KDD process helps identify frequent patterns in the database. Frequent pattern mining is applied to relational, transactional, and time series databases [8]. Frequent pattern mining is active in market basket analysis, bioinformatics, crime prevention, educational mining, and forensic analysis. Frequent pattern mining involves association, clustering, classification, neural networks, and other data facts [9]. The retail industry is always interested in customers' buying behavior to increase profit and plan new marketing strategies—mining of frequent items in retail known as affinity analysis or market basket analysis [10]. MBA is an effective technique to find the most frequent patterns for retailers. MBA identifies purchasing behavior using a transactional dataset [11].

Transactional data is analyzed using different data mining methods for MBA; One data mining technique to identify recurring retailer patterns is AR [12]. Transaction ids and a list of items that have been purchased make up the transactional dataset. One item's association with another item is found in the transactional dataset [13]. Data exploration is done, and preprocessing methods are used to clean and normalize the data. After cleaning, an AR mining algorithm was used to produce frequently used itemsets, and AR was created to discover consumer purchasing habits [14], [15].

Due to the massive data size, researchers are focusing on DL methods to propose new online and offline shopping experiences as retail grows rapidly [16]. Using offline and online retail datasets, DL applications in MBA include next-basket recommendations, churn prediction, customer segmentation, and sales prediction. DL methods, known as artificial neural networks, are based on the structure of the human brain [17], [18]. These neural networks don't need to interact with people to learn from unlabeled data. The neural network is made up of layers, a learning algorithm, and an activation function. Data only moves in one direction in feed-forward neural networks, from the input node to the hidden node and output node. In contrast, the Recurrent neural network operates in both directions [19].

The transactional data is getting large as customer demand is increasing. Finding customer purchase patterns and predicting the next item is a crucial research problem so retailers can upgrade their business strategies accordingly [20]. There is a need to blend AR with DL methods to make predictions and handle rapidly growing transactional datasets. This paper addresses this gap by using AR as input in the DL method to predict the next product category. This paper proposes a novel market basket analysis framework to predict the next product category using AR as a feature selection. The objective of this paper is to 1) Propose a novel model to predict the next product category on two datasets, InstaCart and Bites Bakers, 2) Extraction of AR using Eclat and FP-Growth on InstaCart and Bites Bakers datasets, 3) Classification and prediction using the extracted AR by implementing CNN, Bi-LSTM and CNN-BiLSTM on InstaCart and real-life datasets. In the first phase of the Model, AR is used to find frequent items, while DL methods are used to predict the next product category purchased by the customer.

2. Literature Review

In the past, many researchers conducted MBA using various data mining techniques. A detailed review is published related to the research problem discussed in this paper (Rehman and Ghous, 2021). A brief review of previous research related to MBA using AR and deep learning methods on offline and online retail datasets delineated as follows:

In 2023, Apriori and FP-Growth algorithms were used for market basket analysis with real-world data. The performance of these algorithms was assessed using multilevel association rules mining. According to numerical research results, FP-Growth performs better than Apriori at all levels of product groups in terms of run time and memory [21]. In 2022, a novel next-group recommendation approach based on sequential market basket information was proposed. The proposed method was based on the upper similarity approximation clustering, Borda majority count, and PrefixSpan algorithms. With the proposed approach, customers may have more opportunities to increase their purchases by receiving recommendations for the next group rather than the next item [22]. Apriori algorithms were used in 2022 to gather data on the relationships between sales patterns from records of customer shopping cart transactions, specifically sales of vehicle parts and transactions for vehicle repair services. Interviews and documentation were used to collect data. The criteria used in this study were a minimum of 20 transactions with a frequent item set of 1.7%, a confidence value of 40%, and a lift ratio value greater than 1. The study's findings produced nine sales pattern relationships with 100% certainty [23].

By establishing association rules in 2022, the Apriori algorithm was used. It entails identifying groups of items in a market basket whose combination could result in greater economic benefits for businesses. This study aimed to analyze historical sales data from product groups to identify relationships that would allow companies in the sector to generate patterns to propose portfolio expansion based on the products with the highest purchasing trends [24]. In 2022, a method was proposed to analyze buyer behavior patterns when purchasing an item simultaneously. The FP-Growth algorithm was chosen because it is more efficient and faster. The result of the research is a system that can quantify the value of product associations in transactional data. Five hundred and seventy-one transaction data items were tested, and four rules with lift ratio test values greater than one were discovered [25]. A study was conducted in Pithoragarh (Uttarakhand) city in 2022 in various retail stores such as Bachat Store, Vishal Mart, Buy Chance, Pithoragarh Army Canteen, and so on. he transaction information was gathered from 45 of these stores' customers. RStudio was used to implement the Apriori algorithm. Based on the literature, the support and confidence values were set at 20% and 80%, respectively. The analysis yielded three association rules. All three association rules showing a positive relationship between antecedents and consequents had lift values greater than 1, which was discovered [26].

MBA using Association Rules was held in 2021. The study relied on sales data from any Vancouver Island University website supermarket. A data set containing 225 different products was used

to analyze the Weka program data. Apriori and FP Growth were chosen for assessment. Because the data set is definite, the Apriori algorithm produced no results. As a result, the FP Growth algorithm was used, and the top ten rules were assigned based on the conviction value [27]. In 2020, an Apriori algorithm was used for a frequent item set of sales transactional data of Maharani Supermarket using temporal AR. The items purchased at certain intervals during the fasting month, Christmas, and new year are recommended for product layout at the supermarket. The best rule is milk with a snack for 12, 6, and 3 months at a minimum support of 0.07 and confidence of 0.3. Seven rules were recommended as product layouts appeared at Christmas 2017 and 2018 for next new year's upcoming events [28]. In 2020, the Eclat algorithm was used on a transactional dataset of 212 Mart to find the most frequently bought products in different time intervals. The dataset is divided into quarters in Quarter 1 from February 21, 2018, to May 2018, with 13261 transactions; the bestselling item is Indomie goreng 86gr with 391 pieces. In Quarter 2, from May 22, 2018, to August 2018, with 11978 transactions, the bestselling item is alpha one 600ml with a quantity of 355 pieces. In Quarter 3, from August 22, 2018, to November 2018, the best-selling item is a 60ml milo stick with a quantity of 268 pieces [29].

After reviewing previous research, many authors conducted an MBA using AR to find customer purchase habits. The MBA helps the retailer to find keystone products that recognize in the market and could harm the business if these products are not available or costly. But as data is increasing continuously due to the growing customer demands, it is difficult to find customer purchase behavior by only extracting AR. So it is obvious that AR and DL methods need to be combined to predict and classify these rules, as AR is only used to find customer purchase behavior (Basysyar et al., 2021). There is a lack of using AR as a feature selection with DL methods.

3.   Methodology

Analysis of customer purchase behavior is essential to acknowledge the shoppers about the intention of buyers. It helps retailers design business strategies while knowing consumer purchase behavior and success for new product launches [30]. The change in consumer buying habits will help to predict the market trend. Prediction of the next item to buy is an important research problem of an MBA [31]. In the literature review, this problem is solved using different DL methods. The AR is only used to find buying patterns of customers. There is a gap in using AR as a feature selection to predict the next product category. This paper presents a model using AR as a feature selection and DL to predict the next product category on grocery datasets—the framework implemented on two datasets, InstaCart, and real-life data obtained from a grocery store. Nowadays, DL has emerged as a prominent and emerging research area in a variety of fields. In contrast to conventional ML, DL enables multi-layer computation models to learn representations of data by processing them in their raw form [32]. DL Methods were adopted for this research.

3.1   Dataset

InstaCart is an online grocery store dataset [33] available on kaggle.com. This dataset contains information on aisles, departments, products, and order products. It contains 131209 transactions and 39123 different items. Bites Bakers (https://www.facebook.com/bitesbakersvhr/) is an offline grocery store. It has three branches in Hasilpur, Vehari, and Bahawalpur, Punjab. The sales receipts were collected from Bites Bakers in Jalandhar Colony Hasilpur, Punjab. The receipt dataset of November 2020 to February 2021 consists of 2233 products and 32665 transactions. From these 140 products are the Bites special products, as shown in Figure 1. The proposed model diagram showed in Figure 2.
The methods used to implement the proposed framework are described:

3.2   Association Rules:

Market basket analysis is applied to find customer purchase behavior in retail. AR is commonly used to analyze transactional datasets to find the relationship between items purchased by customers.

Figure 1: Bites Special Products

Market basket analysis is applied to find customer purchase behavior in retail. Association rules are commonly used to analyze transactional datasets to find relationships between items purchased by customers. A transactional dataset is given in the following table. TID is the transaction id of each item. The item's column contains a list of goods purchased by the customer.

Table 1: Association Rules Formation

TIDY	Items
1	Bread, Jam
2	Bread, Egg, Jam, Diaper
3	Bread, Jam, Diaper
4	Bread, Milk, Jam, Coke, Diaper
5	Bread, Milk, Coke, Diaper

Examples of Association rules are:

{Bread} ⟶ {Jam}

The left-hand side item is known as antecedent and the right-hand item is consequent.

Customers who buy Bread also buy jam.

{Bread, Jam} ⟶ {Diaper}

customers who buy bread and jam also buy Diapers. Different dimensions of Association rules are Single Dimension Association Rules which consist of one antecedent and consequent. While Multidimensional Association Rules consist of more than antecedent and consequent.

{Bread} ⟶ {Jam}

{Bread, Milk} ⟶ {Jam, Diaper}

Association rules are generated using different algorithms while doing MBA. These algorithms include FP-Growth, Apriori, Eclat, and Enhanced Apriori with MapReduce Rapid Association Rule Mining, CLOSET, CHARM, CARMA, Sequence, and Apriori with Hashing. In this work, multi-dimension association rules were identified using Eclat and FP Growth algorithms on InstaCart and Bites Bakers datasets respectively.

3.3   FP-Growth:

FP-Growth is an AR mining algorithm for creating frequent item sets in the transactional dataset [34]. It works in two steps

- Construct a data structure called FP-Tree
- Generate frequent itemsets from the FP-Tree
-

It scans the dataset in two phases: In Phase 1
- Scan the dataset and calculate support for each item
- Drop infrequent items
- Arrange frequent items based on support. While in Phase 2
- Build the FP-Tree by scanning the transactions

The FP-Growth is a fast and memory-efficient algorithm because it generates frequent itemsets in just two data scans [35]. Therefore, it works fast than the Apriori Algorithm.

3.4   Eclat:

Eclat (Equivalence class transformation) is a method to find frequent item sets in the transactional dataset [36]. Zaki suggested it in 2001 [37]. It generates frequent itemsets using the depth-first approach. It works on vertical data layout using intersection to calculate the support of itemset [38].

- Generate transaction id (tid) list of each item
- The intersection of an item tid list with a tidy list of all other subsets of that item
- Repeat this process for all items

Eclat is fast and consumes less memory due to the depth-first approach [39]. It scans data only once compared to the Apriori algorithm, which scans original data repeatedly.

3.5   Convolutional Neural Networks (CNN):

CNN is a deep artificial neural network to solve image and text classification research problems. A simple CNN structure consists of convolution, max pooling, and fully connected layers [40].
The work of the convolution layer is to fetch the features of data. In the convolution procedure, the kernel slides on input data, so the kernel weights are multiplied by the corresponding vector values. The next pooling layer extracts the original features decreasing the training parameters and ratio of overfitting. Pooling operations are divided into two types: max-pooling and mean-pooling. Max pooling selects the maximum value in a corresponding convoluted vector.
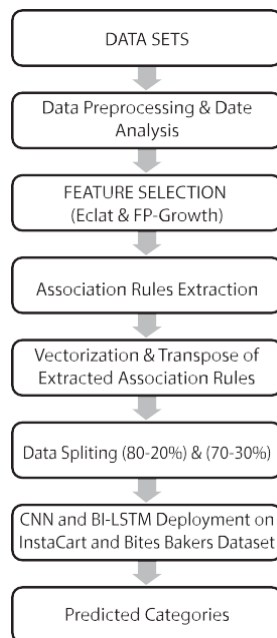

Figure 2: Proposed Model – Flow Diagram

3.6   Long Short-Term Memory (LSTM):

The Long Short-Term (LSTM) neural network is a variation of RNN that defeats the Vanishing Gradient problem faced in traditional RNN [41]. [42]. It can also be used to grasp the time-related dependency from the dataset. LSTMs uses internal memory unit to learn from previous observations to make a prediction. The basic structure of LSTMs consists of three gates: input, output, and forget gates [43]. The flow of training information is controlled across these gates by inserting information in the input gate, removing unnecessary information in forget gate, or passing it to the output gate. The input, output, and forget I represent gates o and f, respectively. The cell state is defined with C, and the cell output is represented by h, while the cell input is given by x. The equations to calculate gates and states are shown below.

$$f_t = \sigma\ (w_f\ .[h_{t-1}, x_t] + b_f) \qquad \dots \quad (1)$$
$$i_t = \sigma\ (w_f\ .[h_{t-1}, x_t] + b_i) \qquad \dots \quad (2)$$
$$c_t = \tanh\ (w_\phi\ .[h_{t-1}, x_t] + b_C) \qquad \dots \quad (3)$$
$$o_t = \sigma\ (w_o\ .[h_{t-1}, x_t] + b_o) \qquad \dots \quad (4)$$
$$h_t = o_t\ \times \tanh(C_t) \qquad \dots \quad (5)$$

The weights of gates represented by w and C is the revised cell state. These weights are upgraded using backpropagation for a long time. The forget gate decreases over-fitting by not permitting whole information from the prior time instant. In traditional LSTM, the input of the hidden layer relies on the computing of the cell controlling the data at the earlier time interval. On the other hand, in Bidirectional LSTM, information flows in two ways: one forward and the other in the opposite direction.

Based on the above methods, the proposed framework is implemented with two datasets, one from previous research and the real-life data of a grocery store, which will be preprocessed into transaction form to generate one hot encoded basket of purchased items. The AR will be extracted using Eclat and FP-Growth algorithms on two datasets. These rules will be converted into vector form to feed into deep learning methods. The extracted rules will be split into (80,20) and (70,30) ratios of training and testing. The experiments on each split will be conducted using CNN, Bi-LSTM, and CNN-BiLSTM methods on two datasets by parameter tuning. The simulation of experiments and results will be mentioned in tabular form with performance metrics

4.   Results and Discussion

The AR is only used to find customer purchase patterns [44]. Deep learning methods must be implemented to make these rules' predictions and classifications. This section describes the simulation of the proposed Model using AR and DL methods—the experiments conducted on two datasets, InstaCart and real-life data from Bites Bakers grocery store. The AR was extracted using Eclat and FP-Growth algorithms on two datasets. These rules are split into (80,20) and (70,30)

ratios to simulate experiments. The experiments are conducted using PyCharm and Rstudio while writing scripts in Python and R programming languages.

4.1   Experiment using InstaCart:

4.1.1   Preprocessing:

The AR algorithm takes data in a sparse matrix, a read. Transactions function applied to convert data into transactions. The transaction data summary shows the sparse matrix size and frequency of frequent items with corresponding product IDs. Unique product labels are assigned to get the corresponding name for product IDs.

4.1.2   Deployment of Eclat

Now the preprocessed data is ready to select features using AR. The frequent items of a maximum length of 15 with support of 0.001 were inspected by implementing the Eclat algorithm to generate these rules. The frequent item set was used to get 347 AR at confidence 0.3, from which 65 is a length 2 (LHS is one item and RHS is one item), 267 is a three dimensional (LHS is two products) and 15 of size 4 (LHS is three products) as shown in Figure 3. The top 15 rules with confidence, support, and lift values are shown in Figure. 4.
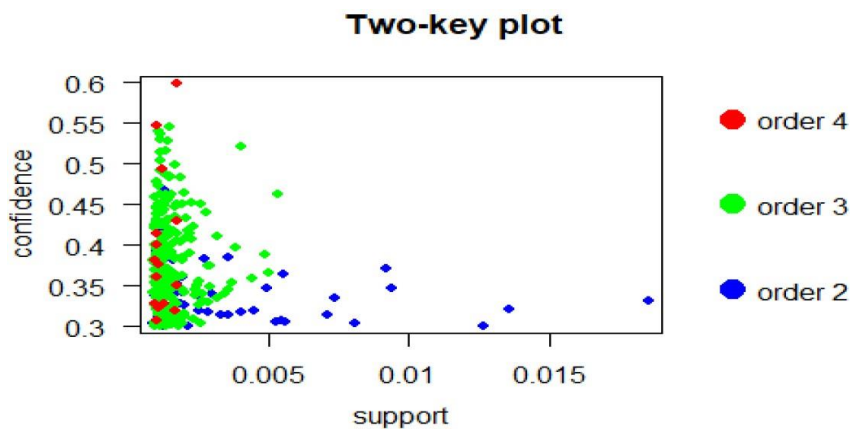

Figure 3: Dimensions of Extracted Association Rules

```
        lhs                                              rhs                        support     confidence lift     itemset
[1]     {Organic Hass Avocado}                        => {Bag of Organic Bananas}  0.018443857 0.3318250  2.812560 2560
[2]     {Organic Raspberries}                         => {Bag of Organic Bananas}  0.013566143 0.3209520  2.720400 2500
[3]     {Organic Raspberries}                         => {Organic Strawberries}    0.012727785 0.3011179  3.626710 2501
[4]     {Honeycrisp Apple}                            => {Banana}                  0.009381978 0.3466629  2.428991 1996
[5]     {Organic Fuji Apple}                          => {Banana}                  0.009221928 0.3715075  2.603072 1967
[6]     {Organic Lemon}                               => {Bag of Organic Bananas}  0.008132064 0.3044223  2.580293 2067
[7]     {Organic Large Extra Fancy Fuji Apple}        => {Bag of Organic Bananas}  0.007415650 0.3365617  2.852709 1831
[8]     {Broccoli Crown}                              => {Banana}                  0.007049821 0.3154843  2.210530 1858
[9]     {Cucumber Kirby}                              => {Banana}                  0.005662721 0.3079155  2.157496 1299
[10]    {Organic Navel Orange}                        => {Bag of Organic Bananas}  0.005525536 0.3661616  3.103598 1068
[11]    {Blueberries}                                 => {Banana}                  0.005456943 0.3082221  2.159645 1261
[12]    {Organic Strawberries,Organic Hass Avocado}   => {Bag of Organic Bananas}  0.005411214 0.4613385  3.910321 2558
[13]    {Apple Honeycrisp Organic}                    => {Bag of Organic Bananas}  0.005235921 0.3050622  2.585717 1277
[14]    {Organic Kiwi}                                => {Bag of Organic Bananas}  0.004984414 0.3478723  2.948578 1140
[15]    {Organic Strawberries,Organic Raspberries}    => {Bag of Organic Bananas}  0.004946307 0.3886228  3.293980 2498
```

Figure 4: Association Rules with High Lift Values

The scatter plot of the 347 rule, as shown in Figure.5, depicts that the rules with high lift have low support.
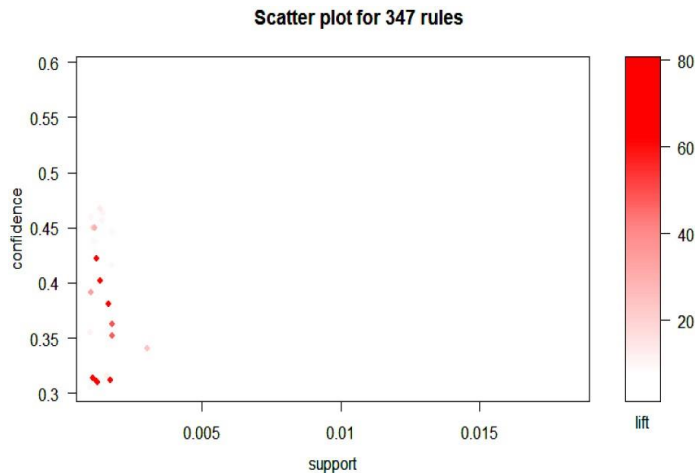

Figure 5: Scatter Plot – Association Rules with High Lift Value

The AR is only used to find customer buying patterns [45]. There is no use of these rules to predict and classify future business needs. In this work, the experiments are conducted to classify and predict 347 rules using CNN and Bi-LSTM. The AR is converted into vector form by making a data frame of unique products on rows and association rules on columns. The class labels are assigned for multi-categorical classification for the cross-ponding products in association rules marked as 1s and the remaining products with 0s. These rules passed to the CNN method consisting of Conv1D layer with activation function' relax, GlobalMaxPool1D layer and two Dense layers with activation function' relax, and the last Dense classification layer used activation function' softmax' to perform multi-classification. The experiments were conducted at epochs five and batch-size 10 using the 'Adam' optimizer by splitting data into (80,20) and (70,30) ratios listed in Table 1. The second method, Bi-LSTM, is implemented with the SpatilaDropout1D layer, Bi-LSTM with dropout

0.2, and the Dense classification layer using activation function' softmax' for multi-categorical classification. The experiment was performed at epochs 50 and batch_size 10 using the 'Adam' optimizer by dividing data into (80, 20) and (70, 30) ratios listed in Table 2.

Table 2: Results using InstaCart

Split	Method	Training Loss	Testing Loss	Training Accuracy	Testing Accuracy
80%-20%	CNN	0.61	0.62	0.80	0.83
80%-20%	Bi-LSTM	0.66	0.35	0.78	0.92
70%-30%	CNN	0.67	0.46	0.77	0.91
70%-30%	Bi-LSTM	0.57	0.46	0.83	0.89

The best results considering simulation at both splits show that Bi-LSTM performs with a high accuracy rate on the InstaCart dataset. The prediction label shows that the customer will buy yogurt, milk, fruits, vegetables, egg, and bread in their next purchase

4.2 Experiment using Bites Bakers:

4.2.1 Preprocessing:

The Sales Detail Report (SDR) is processed by omitting attributes that are not required, such as date, price, time, and total amount. The required data from SDR was extracted by applying Excel formulas—the unique products with product names extracted and stored in the products file. The unique product IDs are assigned to each product, as shown in Figure. 6.

The preprocessing steps performed on the Bites Bakers dataset are grouping transactions based on product id and Sale No, merging the top 100 special Bites product's name on product id as shown in Figure. 7

product_i	product_name
1	BITES CAKE CHOCOLATE/ 1POND
2	DATES 1-KG
3	ITEM
4	STING 300ML
5	BITES SPECIAL BARFI
6	BITES BREAD LARGE-M
7	EGG
8	BITES PASTRY 50
9	LU TUC SNACK PACK
10	PF RIO HALF ROLL
11	LU GALA MANI HALF ROLL
12	SPRITE CAN 250-ML
13	MAZA PANI 1.5L
14	CAPSTAN
15	COCA COLA 1-LITER
16	GLASS
17	LAYS YOGURT & HERB RS 20
18	BITES BREAD LARGE-P
19	SUNDIP CHILLI GARLIC SAUCE 250-G
20	NATIONAL TOMATO KETCHUP 250G

Figure 6: Names and IDs of Bites Products

Sr No	Sale No	Product Name	product_id	QTY
1	SAL-244277	BITES CAKE CHOCOLATE/ 1POND	1	1
2	SAL-244277	DATES 1-KG	2	1
3	SAL-244276	ITEM	3	1
4	SAL-244275	STING 300ML	4	1
5	SAL-244274	BITES SPECIAL BARFI	5	0.25
6	SAL-244274	BITES BREAD LARGE-M	6	1
7	SAL-244274	EGG	7	12
8	SAL-244273	BITES PASTRY 50	8	2
9	SAL-244273	LU TUC SNACK PACK	9	1
10	SAL-244273	PF RIO HALF ROLL	10	1
11	SAL-244273	LU GALA MANI HALF ROLL	11	1
12	SAL-244272	SPRITE CAN 250-ML	12	1
13	SAL-244271	MAZA PANI 1.5L	13	1
14	SAL-244270	CAPSTAN	14	1
15	SAL-244269	COCA COLA 1-LITER	15	1
16	SAL-244269	GLASS	16	4
17	SAL-244269	LAYS YOGURT & HERB RS 20	17	1
18	SAL-244268	BITES BREAD LARGE-P	18	1
19	SAL-244267	EGG	7	6
20	SAL-244267	SUNDIP CHILLI GARLIC SAUCE 250-G	19	1
21	SAL-244267	NATIONAL TOMATO KETCHUP 250G	20	1
22	SAL-244266	BITES HARD NOKKA CHANA	21	1

Figure 7: Sales Transaction Report – Transformation to Transaction

4.2.2   Implementation of FP-Growth

The baskets of the top 100 products were generated by grouping transactions on Sale No and Product Name where Quantity to count items purchased by the customer. The top 20 products based on frequency are shown in Figure. 8. One hot encoding is applied on baskets to generate the frequent item. The one hot encoded baskets feed into the FP-growth algorithm at support 0.001, and the first 20 frequent item sets of size 203 are shown in Figure. 9. These frequent itemsets are used to extract AR.

4.2.3   Implementation of CNN-Bi-LSTM

The rules extracted after implementing FP-Growth must be converted into vector form, as discussed in section 4.1.4, to feed deep learning methods. These rules are converted into vector form and transposed. The data is split into (70,30) and (80,20) ratios to feed into the DL method. The rules are converted into a vector to predict and classify using CNN-BiLSTM. The structure of CNN-BiLSTM consists of a Conv1D layer with the activation function' relax, a MaxPool1D layer, a Dropout layer of 0.2 dropouts, a Bi-LSTM layer, and the last Dense layer with the activation function' softmax'. The experiment used optimizer 'adam,' at epochs 50 and batch size 10 to classify the 170 AR and predict the next customer purchase category, as in Table 3.

```
Bites Total Products 2233
      product_id  frequency                  product_name
0             24       3389        BITES SPECIAL SWEET
1              7       2053                        EGG
2            116       1856        BITES SIMPLE COOKIES
3              1       1753   BITES CAKE CHOCOLATE/ 1POND
4              5       1498        BITES SPECIAL BARFI
5             91       1382            BITES CAKE RASS
6             18       1360        BITES BREAD LARGE-P
7             32       1299            BITES MIX SWEET
8              8       1205            BITES PASTRY 50
9             42       1124       BITES FRUITE CAKE BOX
10            63       1122   BITES CAKE PINEAPPLE /1POND
11            39       1026        BITES BREAD SMALL-P
12           114        895       BITES FF FRY SANDWITCH
13           203        718            FACE MASK 3-PLY
14            14        682                    CAPSTAN
15            67        664        BITES MIX NIMKO 1.KG
16            38        661           BITES RASS SMALL
17           119        653           BITES PIZZA PUNCH
18             6        646        BITES BREAD LARGE-M
19           161        620     BITES SPECIAL ORDER CAKE
20            75        619        BITES BREAD SMALL-M
```

Figure 8: Transaction Group based on the Product ID

```
      support                              itemsets
0     0.042308                   (BITES CAKE RASS )
1     0.062850                             (EGG )
2     0.018950             (BITES BREAD SMALL-M )
3     0.031410             (BITES BREAD SMALL-P )
4     0.006674         (FRESHLEY WATER 600-ML )
5     0.020879                       (CAPSTAN  )
6     0.020236               (BITES RASS SMALL )
7     0.021981                 (FACE MASK 3-PLY )
8     0.006521   (KOLSON SLANTY VEGETABLE RS 20 )
9     0.041635             (BITES BREAD LARGE-P )
10    0.036890                 (BITES PASTRY 50 )
11    0.019991               (BITES PIZZA PUNCH )
12    0.056819             (BITES SIMPLE COOKIES )
13    0.018644               (BITES RASS SLICE )
14    0.045859             (BITES SPECIAL BARFI )
15    0.009888             (FUDGE CREAMY TOFFEE )
16    0.103750             (BITES SPECIAL SWEET )
17    0.007684     (KOLSON SLANTY JALAPNO RS.10 )
18    0.020328             (BITES MIX NIMKO 1.KG )
19    0.007990                   (BITES PEANUT )
Frequent itemsets: (203, 2)
```

Figure 9: Frequent Item Set with Support

The 170 AR generated with support, confidence, and lift values. The top 20 rules based on high lift values are shown in Figure. 10.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
165	(KOLSON SLANY SALTED RS 20)	(KOLSON SLANTY VEGETABLE RS 20)	0.005633	0.006521	0.001500	0.266304	40.839585	0.001463	1.354075
164	(KOLSON SLANTY VEGETABLE RS 20)	(KOLSON SLANY SALTED RS 20)	0.006521	0.005633	0.001500	0.230047	40.839585	0.001463	1.291465
83	(SPRITE CAN 250-ML)	(COCA COLA CAN 250-ML)	0.004990	0.009888	0.001347	0.269939	27.298904	0.001298	1.356203
82	(COCA COLA CAN 250-ML)	(SPRITE CAN 250-ML)	0.009888	0.004990	0.001347	0.136223	27.298904	0.001298	1.151929
97	(BD CANDLE CHORI RS 30)	(BD BALLOON 50 PACK)	0.007500	0.006521	0.001163	0.155102	23.785954	0.001114	1.175857
96	(BD BALLOON 50 PACK)	(BD CANDLE CHORI RS 30)	0.006521	0.007500	0.001163	0.178404	23.785954	0.001114	1.208014
152	(LAYS PAPRIKA RS 20)	(LAYS FRENCH CHEESE RS 20)	0.005786	0.008021	0.001102	0.190476	23.747728	0.001056	1.225386
153	(LAYS FRENCH CHEESE RS 20)	(LAYS PAPRIKA RS 20)	0.008021	0.005786	0.001102	0.137405	23.747728	0.001056	1.152584
157	(LAYS YOGURT & HERB RS 20)	(LAYS PAPRIKA RS 20)	0.007378	0.005786	0.001010	0.136929	23.665613	0.000968	1.151950
156	(LAYS PAPRIKA RS 20)	(LAYS YOGURT & HERB RS 20)	0.005786	0.007378	0.001010	0.174603	23.665613	0.000968	1.202600
161	(COCA COLA 1.5-LITER)	(SPRITE 1.5-LITER)	0.012980	0.005572	0.001653	0.127358	22.858050	0.001581	1.139561
160	(SPRITE 1.5-LITER)	(COCA COLA 1.5-LITER)	0.005572	0.012980	0.001653	0.296703	22.858050	0.001581	1.403419
162	(KOLSON SLANTY VEGETABLE RS 10)	(KOLSON SLANTY SALTED RS 10)	0.009062	0.011388	0.002327	0.256757	22.545590	0.002223	1.330132
163	(KOLSON SLANTY SALTED RS 10)	(KOLSON SLANTY VEGETABLE RS 10)	0.011388	0.009062	0.002327	0.204301	22.545590	0.002223	1.245368
39	(KOLSON SLANTY JALAPNO RS.10)	(KOLSON SLANTY VEGETABLE RS 10)	0.007684	0.009062	0.001531	0.199203	21.983014	0.001461	1.237440
38	(KOLSON SLANTY VEGETABLE RS 10)	(KOLSON SLANTY JALAPNO RS.10)	0.009062	0.007684	0.001531	0.168919	21.983014	0.001461	1.194006
34	(KOLSON SLANTY SALTED RS 10)	(KOLSON SLANTY JALAPNO RS.10)	0.011388	0.007684	0.001714	0.150538	19.590884	0.001627	1.168169
35	(KOLSON SLANTY JALAPNO RS.10)	(KOLSON SLANTY SALTED RS 10)	0.007684	0.011388	0.001714	0.223108	19.590884	0.001627	1.272521
154	(LAYS MASALA RS 20)	(LAYS PAPRIKA RS 20)	0.010011	0.005786	0.001133	0.113150	19.555766	0.001075	1.121062
155	(LAYS PAPRIKA RS 20)	(LAYS MASALA RS 20)	0.005786	0.010011	0.001133	0.195767	19.555766	0.001075	1.230974

Figure 10: Top Twenty AR with High Lift

Table 3: Results using Bites Bakers

Split	Method	Training Loss	Testing Loss	Training Accuracy	Testing Accuracy
80%-20%	CNN-Bi-LSTM	0.03	1.34	1.0	0.75
70%-30%	CNN-Bi-LSTM	0.06	0.55	1.0	0.77

The experiment at (80,20) shows that the best result considering minimum loss is with a training accuracy of 1.0 and testing accuracy of 0.75, as shown in Figure. 11.
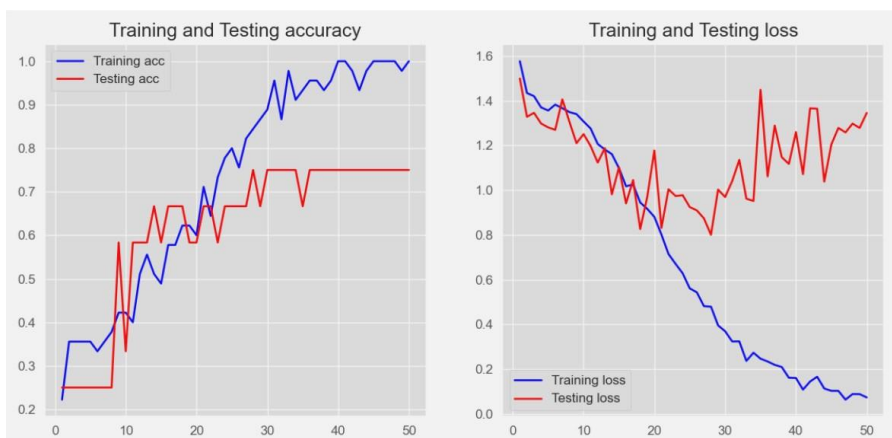

Figure 11: Graphical Representation of Bikes Bakers Dataset using 80-20 Split

The best result at (70,30) shows that CNN-BiLSTM classifies AR with a minimum training loss of 0.06, testing loss of 0.55 and training accuracy of 1.0, and testing accuracy of 0.77, as shown in Figure. 12.
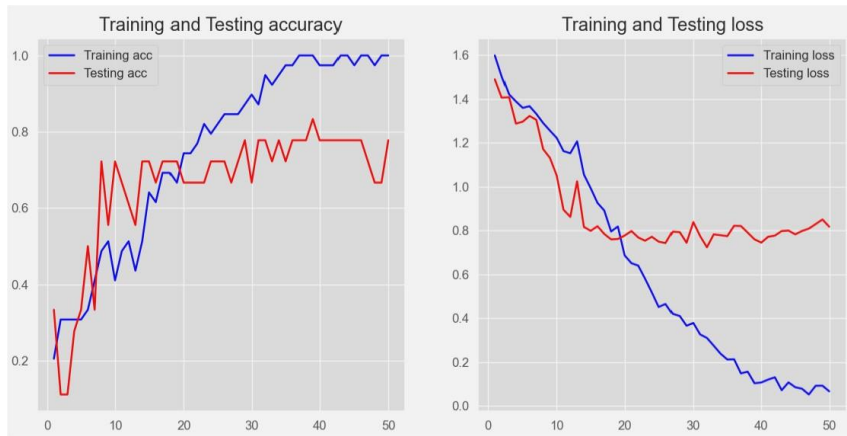

Figure 12: Graphical Representation of Bikes Bakers Dataset using 70-30 Split

The high accuracy rate is 0.77 while considering both splits by implementing CNN-BiLSTM. The prediction label shows that Bites fast food, Bites Ice Bar ice cream, and the Bites-shake will be the next items to purchase. So more variety in these items can be introduced while proposing a new business strategy to increase the profit.

5. Discussion

The experiment simulation shows that Bi-LSTM performs best with a high accuracy rate considering performance metrics at both splits (80,20) and (70,30) on InstaCart. The simulation performed on the proposed framework shows that AR can be used as a feature selection to find customer buying habits. These rules can be used to predict and classify using deep learning methods. So market basket analysis can be conducted using both AR and DL methods. The research objective is extracting AR by implementing Eclat and FP-Growth algorithms on two datasets, InstaCart, and real-life data Bites Bakers. These extracted rules feed into CNN, Bi-LSTM, and CNN-BiLSTM methods for classification and prediction. The results of simulation at both splits (80,20) and (70,30) show the accuracy and loss of both datasets. Compared to related work, a comparative analysis is conducted with a proposed framework. The author used an ensemble approach to predict the purchase probability of retail items. Table 4 shows that the performance of the proposed framework is better than related work.

Table 4: Comparison of Proposed Model with Existing Techniques

Author	Technique	Accuracy
V. Umayaparvath, K. Iyakutti [46]	Baseline, FNN, CNN	71.68%
Ghadekar (2019) [47]	CNN	70.00%
Sharma & Omair Shafiq [48]	Random Forests, CNN, XGBoost	88.84%
Seymen, Omer Faruk [49]	NN, Logistic Regression	89.00%
Proposed Framework	Association rules as feature selection using CNN for next-item prediction	91.00%

6.   Conclusion and Future Work

The relationship among buyers' preferences can be deduced using a data mining technique called market basket analysis. The analysis of these associations assists the retailer in proposing a business strategy by assessing the frequently buying items. The comparison between the customer buying pattern and the production pattern of the company helps in buying the product. Analyzing the consumer buying pattern is crucial; many data mining methods can attain this. AR is a data mining approach to understanding the frequent activities of purchasers. In previous research, many authors used this method to detect only customer frequent purchase patterns. There is a lack of predicting and classifying the AR to propose new business strategies for the growing needs of consumers. This work introduced a framework using AR and DL methods to overcome the problem. The AR is used as a feature selection to find customer purchase behavior from transactional datasets. These rules are classified and predicted using deep learning methods. The experiments are simulated in Python and R programming languages on InstaCart and Bites Bakers datasets. To conclude, the best accuracy is achieved using Bi-LSTM with a 0.92 accuracy rate on InstaCart. The performance metrics on Bites Bakers expose CNN-BiLSTM classifies AR at a high accuracy rate of 0.77.

Market Basket Analysis helps the manager understand which items customers purchase during their visits to the store. The results can be used to propose marketing strategies and design different store layouts to attract customers. MBA on Bites Bakers dataset predicts Bites fast food, Bites Ice cream, and Bites shake will be the next product category to purchase. So these products can be placed at stores where new products can be placed to capture customer intentions. The new items in predicted categories can be launched to increase revenue or put on Sale.

The proposed model is slightly overfitting on Bites Bakers dataset. However, the model performed well on Bites Bakers. This happened due to the inconsistency of the real time dataset. In the future, the model will be improved to overcome the inconsistency of the real-time dataset for MBA by introducing new preprocessing methods. Furthermore, the proposed model will be implemented on a large grocery store's transactional dataset for future work. The dimension reduction methods should be used to reduce the dimensions of the baskets. Many ARs can be used to evaluate the performance of the proposed framework. The accuracy of the real dataset will be improved using multiple preprocessing techniques.

References

[1] Grewal, Dhruv, Michael Levy, and Vijay Kumar, *"Customer experience management in retailing: An organizing framework,"* Journal of Retailing, vol. 85, no. 1, pp. 1--14, 2009.

[2] Wang, Zhanpeng, et al, *"Optimal retail sales strategies for old and new products in monopoly and horizontal competition scenarios,"* Journal of Retailing and Consumer Services, vol. 71, p. 103218, 2023.

[3] Nouinou, Hajar, et al., *"Decision-making in the context of Industry 4.0: Evidence from the textile and clothing industry,"* Journal of Cleaner Production, p. 136184, 2023.

[4] Blattberg, Robert C., et al., *"Market basket analysis," Database Marketing: Analyzing and Managing Customers,* pp. 339--351, 2008.

[5] Raorane, Abhijit, and R. V. Kulkarni., *"Data mining techniques: A source for consumer behavior analysis,"* arXiv preprint arXiv:1109.1202, 2011.

[6] Bihl, Trevor, et al., *"Artificial Neural Networks and Data Science,"* Encyclopedia of Data Science and Machine Learning, pp. 899--921, 2023.

[7] Sharma, Moolchand, et al., *"Intelligent Data Analysis for Medical Applications," Intelligent Data Analysis: From Data Gathering to Data Comprehension,* pp. 333--346, 2020.

[8] Sarhan, Amany M., *"Data Mining in the Internet of Things Systems: A Literature Review,"* Journal of Engineering Research, vol. 6, no. 5, pp. 252--263, 2023.

[9] Solanke, Abiodun Abdullahi. *"Digital Forensics AI: on Practicality, Optimality, and Interpretability of Digital Evidence Mining Techniques."* (2022).

[10] Kaur, Manpreet, and Shivani Kang. *"Market Basket Analysis: Identify the changing trends of market data using association rule mining."* Procedia computer science 85 (2016): 78-85.

[11] Kaur, Manpreet, and Shivani Kang. *"Market Basket Analysis: Identify the changing trends of market data using association rule mining."* Procedia computer science 85 (2016): 78-85.

[12] Kaur, Manpreet, and Shivani Kang. *"Market Basket Analysis: Identify the changing trends of market data using association rule mining."* Procedia computer science 85 (2016): 78-85.

[13] Patel, Suresh B., Sanjay M. Shah, and Mahendra N. Patel. *"An Efficient Search Space Exploration Technique for High Utility Itemset Mining."* Procedia Computer Science 218 (2023): 937-948.

[14] H. H. Nasereddin, *"New Data Mining Techniques to deal with both Static and dynamic databases,"* 2007.

[15] Joshi, ASHISH P., and BIRAJ V. Patel "*Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process,"* Orient. J. Comput. Sci. Technol, vol. 13, no. 0203, pp. 78--81, 2021.

[16] Guha, Abhijit, et al. *"How artificial intelligence will affect the future of retailing."* Journal of Retailing 97.1 (2021): 28-41.

[17] Rehman, Iqra, and Hamid Ghous. *"Structured Critical Review on Market Basket Analysis using Deep Learning & Association Rules."* International Journal of Scientific & Engineering Research 12.1 (2021): 1-24.

[18] SEYMEN, Omer Faruk, et al. *"Customer Churn Prediction Using Ordinary Artificial Neural Network and Convolutional Neural Network Algorithms: A Comparative Performance Assessment."* Gazi University Journal of Science (2023).

[19] Alzubaidi, Laith, et al. *"Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions."* Journal of big Data 8 (2021): 1-74.

[20] Rita, Paulo, Tiago Oliveira, and Almira Farisa. *"The impact of e-service quality and customer satisfaction on customer behavior in online shopping."* Heliyon 5.10 (2019).

[21] D. a. O. K. a. O. B. a. M. A. Y. a. O. T. Alcan, *"A Comparative Analysis of Apriori and FP-Growth Algorithms for Market Basket Analysis Using Multi-level Association Rule Mining,"* in Industrial Engineering in the Covid-19 Era: Selected Papers from the Hybrid Global Joint Conference on Industrial Engineering and Its Application Areas, GJCIE 2022, October 29-30, 2022, Springer, 2023, pp. 128--137.

[22] L.-C. Ma, *"Novel next-group recommendation approach based on sequential market basket information,"* Electronic Commerce Research, pp. 1--20, 2022.

[23] S. a. R. A. a. K. R. a. N. N. Karnila, *"MARKET BASKET ANALYSIS ON TRANSACTION DATA USING THE APRIORI ALGORITHM,"* Jurnal TAM (Technology Acceptance Model), vol. 13, no. 1, pp. 34--43, 2022.

[24] G. M. I. a. R. X. D. M. a. L. Y. J. G. a. G. J. A. T. Caldas, *"A Product Network Analysis Using A Priori Algorithm for Extending the Market Basket in Retail,"* in Proceedings of the First Australian International Conference on Industrial Engineering and Operations, Sydney, Australia, 2022.

[25] M. R. a. S. M. a. I. H. a. C. J. C. a. S. A. Pradana, *"Market Basket Analysis Using FP-Growth Algorithm On Retail Sales Data,"* in 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), IEEE, 2022, pp. 86--89.

[26] A. a. P. S. Tripathi, *"Market Basket Analysis of Cosmetic Products using Apriori Algorithm,"* Journal) Volume, vol. 2, no. 1, 2022.

[27] Y. A. Unvan, *"Market basket analysis with association rules,"* Communications in Statistics-Theory and Methods, vol. 50, no. 7, pp. 1615--1628, 2021.

[28] S. C. a. M. K. Bilqisth, *"Determination of Temporal Association Rules Pattern Using Apriori Algorithm,"* IJCCS (Indonesian Journal of Computing and Cybernetics Systems), vol. 14, no. 2, pp. 159--168, 2020.

[29] Z. a. T. K. D. NURZANI, *"Analysis of Transactions 212 Mart Kuto Palembang to Find Frequent Patterns Among Itemset Using Association Rule Mining,"* in Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019), Atlantis Press, 2020, pp. 325--332.

[30] P. a. H. K. A. a. G.-H. E. a. Z. H. a. F.-F. M. Ebrahimi, *"Consumer knowledge sharing behavior and consumer purchase behavior: Evidence from E-commerce and online retail in Hungary,"* Sustainability, vol. 13, no. 10, p. 10375, 2021.

[31] D. a. S. I. Gal, *"redicting consumers' choices in the age of the internet, AI, and almost perfect tracking: Some things change, the key challenges do not,"* Consumer Psychology Review, vol. 4, no. 1, pp. 135--152, 2021.

[32] N. K. a. S. K. Chauhan, *"A review on conventional machine learning vs deep learning,"* in 2018 International conference on computing, power and communication technologies (GUCON), IEEE, 2018, pp. 347--352.

[33] S. C. a. L. J. a. Y. N. Chintala, *"Browsing the Aisles or Browsing the App? How Online Grocery Shopping Is Changing What We Buy,"* How Online Grocery Shopping is Changing What We Buy (August 8, 2022), 2022.

[34] M. a. B. M. a. E.-g. S. a. A. R. a. E.-d. A. Shawkat, *"An optimized FP-growth algorithm for discovery of association rules,"* The Journal of Supercomputing, pp. 1--28, 2022.

[35] L. a. F.-V. P. a. Y. R. a. D. I. {Shabtay, *"A guided FP-Growth algorithm for mining multitude-targeted item-sets and class association rules in imbalanced data,"* Information Sciences, vol. 553, pp. 353--375, 2021.

[36] I. a. R. M. a. H. M. a. N. N. a. L. F. a. N. F. Fauziah, *"Market Basket Analysis with Equivalence Class Transformation Algorithm (ECLAT) For Inventory Management Using Economic Order Quanty (EOQ),"* 2022.

[37] M. J. a. G. K. Zaki, *"ast vertical mining using diffsets,"* in Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 2003, pp. 326--335.

[38] W. a. J. M. a. M. M. a. A. Z. a. M. F. Bakar, "*Postdiffset: an Eclat-like algorithm for frequent itemset mining,"* International Journal of Engineering \& Technology, vol. 7, pp. 197--199, 2018.

[39] L. Schmidt-Thieme, *"Algorithmic Features of Eclat.,"* FIMI, 2004.

[40] L. Z. J. H. A. J. A.-D. A. D. Y. A.-S. O. S. J. F. M. A. A.-A. M. F. L. Alzubaidi, *"Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions,"* Journal of big Data, vol. 8, pp. 1--74, 2021.

[41] V.-D. a. L. C.-M. a. T. D. A. Ta, *"Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading,"* Applied Sciences, vol. 10, no. 2, p. 437, 2020.

[42] B. T. a. H. T. T. Khoa, *"Long Short-Term Memory Recurrent Neural Network for Predicting the Return of Rate Underframe the Fama-French 5 Factor,"* Discrete Dynamics in Nature and Society, p. 2022, 2022.

[43] M. a. D. Q. a. H. N. u. a. J. K. a. S. S. M. a. J. H. T. Ishfaque, *"Use of recurrent neural network with long short-term memory for seepage prediction at Tarbela dam, KP, Pakistan,"* Energies, vol. 15, no. 9, p. 3123, 2022.

[44] A. a. D. G. a. B. F. M. a. S. N. a. o. Bahtiar, *"Data mining techniques with machine learning algorithm to predict patients of heart disease,"* in IOP Conference Series: Materials Science and Engineering, vol. 1088, IOP Publishing, 2021, p. 012035.

[45] R. C. A.-M. M. C. I. Lixuandroiu, *"An analysis of the impact of personality traits towards augmented reality in online shopping,"* Symmetry, vol. 13, no. 1, p. 416, 2021.

[46] V. a. I. K. Umayaparvathi, *"Automated feature selection and churn prediction using deep learning models,"* International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 3, pp. 1846--1854, 2017.

[47] P. a. D. A. Ghadekar, *"Image-Based Product Recommendations Using Market Basket Analysis,"* in 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA), IEEE, 2019, pp. 1--5.

[48] A. Sharma, *"Retail Customer and Market Proclivity Assessment using Historical data and Social Media Analytics,"* 2020.

[49] O. F. a. D. O. a. H. A. Seymen, *"Customer churn prediction using deep learning,"* Proceedings of the 12th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2020), pp. 520--529, 2021.