Structured Critical Review on Market Basket Analysis using Deep Learning & Association Rules

Iqra Rehman, Dr. Hamid Ghous

Abstract— Market Basket Analysis (MBA) is a Data mining field to mine frequently buying patterns, helping retailers while proposing new business strategies. As demand of customer changes with growth of needs, sales data has increased massively in various businesses. To handle this massive data with changing frequently consumer buying habits there is demand of deep learning methods. In past many researchers performed MBA using deep learning and association rules on offline and online retail datasets. Association rules are used to mine frequently purchased items from retail dataset. There is a need of comprehensive review on MBA using deep learning and association rules to provide directions for future research. Therefore, in this paper a critical review is conducted with Association rules and Deep learning methods structured on online and offline retail datasets. In each section, summary table with dataset dimensions, pre-processing techniques, Results and future directions are delineated. The objective is to provide guide to future research while proposing hybrid framework considering MBA review table and future directions.

Index Terms— Market Basket Analysis, Association rules, Deep Learning, Frequent Pattern Mining, Retail Dataset

1 Introduction

Retailers are interested in customer buying behavior to compete with growing needs of consumer and remain stable in market. Dataset is increasing in volume day by day due to increasing demand of customers. As variety of items introduced to meet future needs so customer buy different set of items in a single visit of store. Traditional method takes more time to find purchasing behavior due to large volume of products and customers. So there is need of data mining methods for market basket analysis [1].
Association rules is one of the data mining method [2] to find frequent patterns for retailers. Market basket analysis performed to identify purchasing behavior using transactional dataset [2]. Transactional dataset consists on transaction ids and list of purchased items. In transactional dataset association of one item with another item is elevated. Data is explored and preprocessing techniques are applied to clean and normalized data. After cleaning, association rule mining algorithm applied to generate frequent item sets and association rules generated to find consumer buying habits in purchasing of goods [2].

As Retail is growing at rapid pace, researchers are concentrating on deep learning methods to propose new online and offline shopping experiences due to massive size of data [3]. Applications of deep leaning in MBA are next basket recommendations sales prediction, customer segmentation and churn prediction using offline and online retail datasets. Deep learning methods worked on structure of human brain called artificial neural networks. These neural networks can learn from unlabelled data without interaction of human beings. The neural network consists on layers, the learning algorithm and activation function. In Feed forward neural networks data flows only in one direction from input node to hidden node and to output node while Recurrent neural network works in bi-directional mode [4].

2 Significance of Market Basket Analysis

Now a day's Market basket analysis is a foremost problem in cross selling and upselling [5]. Cross selling is increasing sale of product and adding new features in current product. Upselling is introducing new category of product to attract more customers. Customer buying pattern is essential to know for upselling and cross selling. Here market basket analysis plays a vital role in identifying strong and weak items engendering of profit and loss.
Market basket analysis is used to find following problems for retailers [6]:

- Find product association to other product
- To optimize store and product layout
- Optimized arrangement of inventory on online store
- To predict future growing demands of product

3 Research Methodology

The literature search was conducted on the descriptor "market basket analysis" and "retail dataset". After reviewing

- *Dr. Hamid Ghoush is currently working as assistant professor at Institute of Southern Punjab Multan.He did his PHD from University of Technology Sydney.He got more than ten years of research experience from oversease and Pakistan.E-mail: hamidghous@isp.edu.pk*
- *Iqra Rehman is an M. Phil student at institute of Southern Punjab Multan9. She did bachelor at Punjab University College of Iinformation Technology Lahore. E-mail: iqrarehman504@gmail.com*

the complete text, the articles were eliminated not related to market basket analysis on retail dataset. The review is structured on the following criteria:

- Section I reviewed on offline retail dataset
- Section II reviewed on online retail dataset
- In Each section articles classified as shown in Fig:1 on data mining methods Association rules, Deep learning, Association rules and Deep learning
- Each section contains summary table delineating author name and year, method name, preprocessing techniques, dataset with dimensions, Results, Limitations and appealing work of author

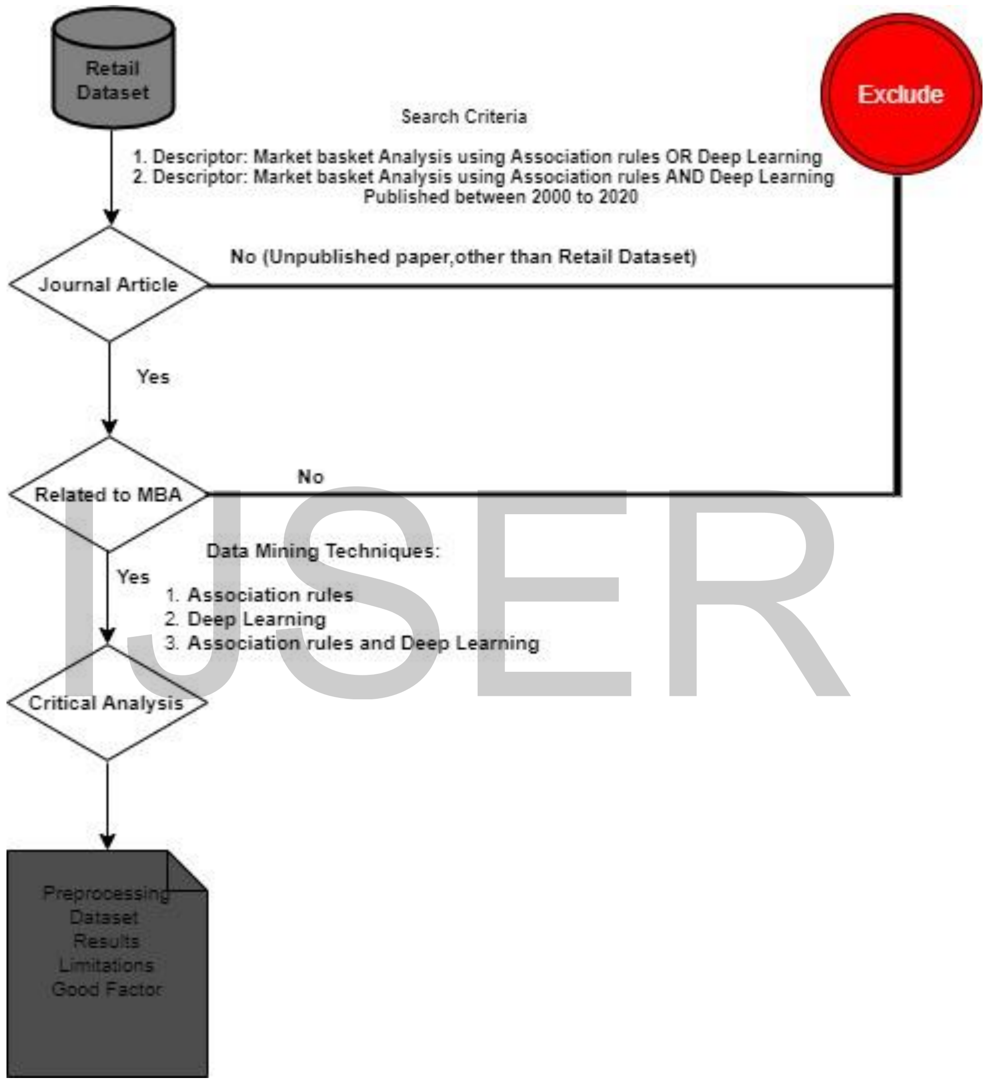

Fig. 1. Selection and Evalution Framework

4 LITERATURE REVIEW

4.1 Offline Retail Dataset

The offline retail dataset of purchased goods collected at storefront location by physically visiting the customer to mine frequent purchased intentions for offline retailers. In past the market basket analysis on offline retail dataset conducted using association rules and deep learning methods.

4.1.1 Market Basket Analysis using Assoication Rules

Association rules used to mine frequent purchase behaviour of customers that exist in offline transactional dataset [7]. Market basket analysis using association rule mining algorithm on offline retail dataset conducted by [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30].

Chen, Chiu and Chang [8] mined customer purchase patterns using Apriori algorithm by integrating customer demographic variables as Recency, frequency, and monetary (RFM) values with transaction dataset of foodmart over time. While G¨urdal Ertek [9] proposed association rules visulization framework on real time datset of supermarket. Relatively Nafari and Shahrabi. [10] worked on planograming regarding dynamic change of product price using Apriori-TdMl on real time dataset. While Avcilar and Yakut [11] worked on clothe and accessory dataset using apriori algorithm on both in-category and cross categories of product. The association rules visulized using web graphics to see relationship among items. Similarly Zulfikar *et al* [12] applied market basket analysis to find frequent items using Apriori algorithm at eight clusters of XMART retail company at Indonesia. Comparatively Abusida and Gültepe [13] designed business stretagey on order and purchase dateset of electricity company appling apriori algorithm. The association rules generated at high confidence value of 100% showing frequent demand spare parts. Comparatively Nur,Triayudi and Diana [14] analyzed sales dataset of resturent to mine customer intentions at mimum support 50%. A dynamic application develoed to abserve customer purchase patterns within specific time. Likewise Sutisnawati and Reski [15] mined association among items of dining place using small transaction dataset. The association rules found for frequenlty and rarely purchased items at mimum confidence 65%. While Ariestya, Supriyatin and Astuti [16] analyzed small grocery store dataset applying association rules algorithm Apriori and FP growth.The experiment shows that FP growth produced more rules than Apriori at minum support 45% and confidence 60% .

Anggraeni, Iha, Erawati, Khairunnas [17] analyzed sales dataset of electronic store to find most frequenly purchased item by customer. The assocaition rules extracted at mimum support 9% and confidence 40%. While Adalı and Balaban [18] followed CRISP-DM process to mine customer purchase behaviour independent of time and area using apriori algorithm on electronic dataset.The visulization of association rules incorporated in dynamic application. Comparatively Rizqi [19] proposed product bundeling strategy for electronic accessories using FP growth. The experiment perforemd on small retial dataset at minum support 30% and confidence 60%. While Halim, Halim and Felecia [20] designed business strategy for resturent with 4P (Price,Product,Place,Promotion) using power BI for data visulization. The data is analyzed considering effects of transaction day , month , weekend or non-weekend , morining or night. On the other hand Wijana and Finandhita [21] analyzed cafe sales data using CT-Pro algorithm to design business promotion.The experiment perfromed on small transaction dataset at mimum confidence 60%.Kurniawati and Widianti [22] propsed product recommendation system for clothing shop at support and confidence of 15%. While Bilqisth and Mustofa [23] analyzed supermarket transaction datset considering fasting month, Christmas and new year effects on customer buying behaviour. The association rules extracted at minimum support 70% and confidence 30% applying apriori algorithm.

M.Kavitha and Subbaiah [24] worked on groceries sample dataset avaiable in R to see associations among products at support 0.007, confidence 0.6 and lift 0.5. The association rules extracted using apriori algortihm visualized on group matrix. While Nurzani and Tania [25] analyzed real transaction dataset of mart using Eclat algorithm.The perfromance of Eclat compared with Apriori and FP-growth. The analysis of result is shown using confusion matrix with 91.2% accuracy , 92% precision, and 88% recall. On the other hand Liansyah and Destiana [26] analyzed large transaction dataset of cafe with detail simulation of appplied apriori algorithm. The association rule extracted at support 16% and confidence 100%. Comparatively Alfiqra and Khasanah [27] analyzed transaction dataset of retail over 4 different time periods.The association rules using apriori algorthm analyzed using overll Variability of Association Rule (OCVR).The OCVR value small than 30% show that these rules are less vurnerable to change so can these rules be used to make business stratgey at any time. While Yudhistyra, Risal, Raungratanaamporn and Ratanavaraha [28] analyzed big data of gold and silver metal company using apriori and CARMA.The CRISP-DM process is followed with proper data visulization on item frequency using web graph. comparatively Efrat, Gernowo and Farikhin [29] worked on minimarket transaction dataset using apriori algorithm at minium support 10% and confidence 65%. Similalry Rachmatika and Harefa. [30] analyzed small sales transaction dataset at minium support 105 and highest confidence 100% to lowest 84.

Table 1 Summary of MBA using Association Rules on Offline Retail dataset

Author Year	MBA using AR	Preprocessing	Dataset	Result	Improvement	Good
Mu-Chen Chena, Ai-Lun Chiu, Hsu-Hwa Chang 2005 [8]	Association rules with apriori algorithm	Data transformation, integration and segmentation	Transaction data of foodmart including customer, product and transaction database	At minimum support and confidence of 20% generated 111 and 78 rules for first and second period respectively	Visualizations of results Use of large real transactional dataset	developed online system to mine customer patterns at different time
G¨urdal Ertek1, Ayhan Demiriz 2006 [9]	Association rules using apriori algorithm	Not mentioned	Belgian supermarket dataset with 88,163 transactions and 16,470 unique items	At support 2% and confidence 20% generated 27 rules.	Integrating association mining and visualization in single model Exploring association rules in 3D	Application of visualization framework on real dataset with simulation
Maryam Nafari, Jamal Shahrabi Amirkabir 2010 [10]	Association rules using Apriori-TdMl	Not mentioned	Retail dataset of supermarket in Iran of four months with 1241 products 7305 sales transaction.	Using Apriori-TdMl 102, 329, and 1579 frequent itemsets found for categories, subcategories and products respectively at minimum support 0.1 and confidence 0.1. Cross selling profit effect for frequently buying products are shown for 6 products. At lift 2.75 shelf space allocated for high cross selling products	Deep learning models to predict shelf arrangement of products. Use of large transactions dataset.	Product shelf arrangement considering price, product cross category purchase and temporal effects.
Mutlu Yüksel Avcilar, Emre Yakut 2014 [11]	Apriori Algorithm	Calculated Descriptive statistics of data at product category level with Mean, Median, Mod, Maximum and Minimum values. Calculated Sales frequency and percentage of products	Sales transaction dataset of store in Turkey with 9.000 product of 35 categories and 42.390 transaction records from 01.01.2012 and 31.12.2012	At minimum support 0,05% and confidence 50% extracted 8 best rules	Time base mining of association rules. Predicting customer Demographic information using deep learning methods.	Web graphic view to show association among products Association rules defined clearly
Wildan Budiawan Zulfikar, Agung Wahana, Wisnu Uriawan, Nur Lukman 2016 [12]	Apriori Algorithm	Combining different products on Frequency	XMART Retail in Indonesia 25 products 145.548 record	Generated frequent 10 rules at minimum support of 12% and confidence of 60% for 8 clusters of store	Considering sale's day and hour effects on sales	Followed standard process CRISP-DM. Worked on Real Dataset
Ashraf Mohammed Abusida, Yasemin Gültepe 08 October 2019 [13]	Apriori Algorithm	Data selection, Data splitting , Data filtering , Data Transformation to ARFF format.	General electricity company of Libya(GECOL) Dataset properties ORDER_NO, PART_NO, PART_DESCRIPTION, SITE,UNIT	At confidence 100% 10 best rules are generated for each site	Data visualization of association rules Scatter plot for confidence and support values	Simulation of results

Yuli Nur Indah Sari, Agung Triayudi, Ira Diana Sholihati July 2019 [14]	Apriori Algorithm	Binarization	Sales transaction data of Fonzu Premium restaurant of month June 2018	With manual calculation at confidence 100% Chicken Teriyaki Don and Gyoza Saikoro Yakimeshi With built-in system calculation at confidence 100% Chicken Teriyaki Don and Gyoza Saikoro Yakimeshi	Considering Seasonal and holidays effects on sales of food	Results are mentioned clearly Simulation of association rules shown Defined research methodology clearly
Y Sutisnawati, M Reski 2019 [15]	Apriori Algorithm	Removed unnecessary Attributes Date, Price, Amount, Total , Location.	XYZ dataset with 157 transaction and 32 items with smallest frequency 2 and largest frequency 68	At minimum confidence 65% and support 23% 2 pairs are often purchased	Graph based visualization of association rules Scatter plot of confidence and support Considering Seasonal and holidays effects on sales of food Use of large transaction dataset	Comprehensive Simulation of Results Dynamic analysis of association rules using developed application
Winda Widya Ariestya, Wahyu Supriyatin, Ida Astuti December 2019 [16]	FP-Growth, Apriori Algorithms	Data cleaning from 620 to 577 data, Data integration, Data selection, Data Transformation	Dataset of 620 items	FP-Growth: At minimum support 45% and confidence 60% 5 rules Apriori: At minimum support 45% and confidence 60% 3 rules	Dataset with proper dimensions Performance plot of two algorithms Use of large dataset	Followed standard Research methodology Results mentioned with support and confidence
Sita Anggraeni, Marlina Ana Iha, Wati Erawati, Sayyid Khairunnas April 2019 [17]	Apriori, FP-Growth Algorithms	Not mentioned	Electronic Sales data from January 2016 to December 2016 116 transactions containing 14 items	Using Apriori at minimum support of 9% and confidence of 40% 10 best rules	Association rules with FP-growth not mentioned properly Performance plot of two algorithms Visualization of association rules Use of large dataset	Use of real dataset
Gökçe KARAHAN ADALI, M. Erdal BALABAN April 2019 [18]	Apriori Algorithms	39 missing records in Group-Code and Typecode areas filled with most repeated area "Illumination Region". Outlier Analysis decreases number of observations from 183403 to 183402,Duplicated Observations, Binarization	Dataset1 from January 2014 to December 2014 with 19 attributes 177393 transaction Dataset2 of year 2015 with 23 attributes 183401 transaction	At support 0.1% and leverage 185.0291 product HESNYA-03 and HESNYA-02 in different color packages can be used for promotions	More than one dynamic model can be added applying different algorithms	Followed CRISP-DM methodology Dynamic model for rules independent of time and region Association rules can download for analysis confidence Results mentioned clearly. Proposed business strategy on obtained results
Zakka Ugih Rizqi 2019 [19]	FP-Growth	Data cleaning, Recapitulating data, removing duplicate data, checking inconsistent data, Data Transformation, cleaning of noise and missing data	Retail Z dataset of 5 Departments with 58 transactions	At minimum support of 30% and 60% generated 6 rules.	Scatter plotting of confidence and support. Graphical visualization of association rules Use of Large dataset Considering price attribute for product bundling	Problem statement mentioned clearly Results with support and confidence Research methodology defined clearly Worked on real dataset Proposed business strategy on mined frequent patterns

Karina Kusuma Halim, Siana Halim, Felecia 2019 [20]	Association rules with matrix incidence	Data Cleaning, Data Aggregating, Deleted duplicate data for each transaction	Surabaya restaurant sales with 15087 records from March 23 to June 30, 2018 and September 23 to October 16, 2018 May 17 until June 16, 2018	At minimum support and confidence of 4% and 15% 10 rules	Use of Dimension reduction or feature selection method on dataset	Data visualization using power BI Business strategy discussed briefly Research method and flow defined clearly Results are mentioned properly Dashboard for making business strategy in different time period Data analysis using matrix incidence Considering transaction day week, month and food type on customer purchase
Heri Wijana, Alif Finandhita 2019 [21]	Association rules using CT-Pro algorithm	Data Cleaning	57 sales transaction data for October 2018	At minimum confidence 60% generated 4 packages with menu and price	Scatter plotting of confidence and support. Graphical visualization of association rules. Use of large dataset	Problem statement is defined clearly Source of Dataset mentioned Followed CRISP-DM simulation of algorithm considering price in designing menu package
Desti Kurniawati, Utami Dewi Widianti 2019 [22]	Association rules using apriori	Not mentioned	Dataset from period 2 February 2019	At minimum support 15% extracted 12 best association rules	Dataset not mentioned with dimensions Considering Seasonal effects on purchase of clothes	Results not mentioned briefly with Support and confidence
Shona Chayy Bilqisth, Khabib Mustofa 2020 [23]	Temporal Association rules using Apriori Algorithm	Data selection. Data Cleaning , Data Transformation	Maharani Supermarket 3 year sales transaction data in 2016,2017 and 2018 with 429,832 records. In first year 126,590 records ,In second year 126,590 records, In Third year 153,324 records.	At minimum support 0.07 and confidence 0.3 rule generated for 12 months,6 months and 3 months is Milk with snack. 7 rules appeared at Christmas 2017 and 2018 recommended as product layout for next new year upcoming events	Visualization of association rules Scatter plotting of confidence and support	Temporal analysis of association rules with fasting month, Christmas and new year. Results mentioned clearly for every evert. Dataset mentioned with proper dimensions
Mrs. M. Kavitha, Dr. S. Subbaiah 2020 [24]	Apriori Algorithm	Not mentioned	Groceries Dataset 9835 transactions and 169 items.	At support of 0.007,confidence 0.6 butter and yogurt most frequent item	Limited Research scope	Visualization of association rules Use of scatter plot to show support, confidence and lift values

Zikri NURZANI, Ken Ditha TANIA 2020 [25]	Eclat algorithm	Data aggregation, Data Binarization	212 Mart Sales transaction data from February 21,2018 to February 2019 with 58068 transactions 140733 records and 47 attributes.	In Quarter 1 from 21 February 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 the quantity as 355 pieces. In Quarter 3 from August 22, 2018 November 2018 the bestselling item is 60ml milo stick with quantity as 268 pieces.	Scatter plot of confidence and support Dynamic Framework to analyze association rules with time	Performance plot of Eclat with Apriori and FP-growth algorithms Results are mentioned clearly Graphical visualization of association rules
Ovi Liansyah, Henny Destiana April 2020 [26]	Apriori Algorithms	Not mentioned	sales data in the April 2019 period in Lotteria.	At support of 16% and confidence 100% Hot/Ice coffee and float most buying product	Large transaction dataset to check reliability of Apriori algorithm Data visualization of association rules	Results are shown with support and confidence value Simulation of apriori algorithm shown in detail
Alfiqra, A UKhasanah 2020 [27]	Apriori algorithm	Data cleaning, Data reduction, Data integration	Supermarket X dataset of 57784 transactions in a month including 41248 items	At 1%<OCVR<30% generated 17 rules	Scatter plotting of confidence and support. Graphical visualization of association rules	Followed standard research methodology Results mentioned clearly Proposed Business strategy
Wecka Imam Yudhistyra, Evri Marta Risal, I-soon Raungratanaamporn, Vatanavongs Ratanavaraha June 10,2020 [28]	Apriori, CARMA	Measuring central tendency, removing null tuples, Sorting hashing, Exploring Aggregating and visualizing the relationship. Dataset transformed to 80 variables	3,986,872 observations from 248,856 customers	Apriori: At minimum support of 1% and confidence of 50% 21 rules CARMA: At minimum support of 1% and confidence of 50% 5 rules	A flexible model for development of new model to propose business strategy in different time period	Web graph visualization of item sets with weak, medium and strong link Worked on real dataset Application of two algorithms with performance plot Followed standard CRISP-DM process
A R Efrat, R Gernowo and Farikhin 2020 [29]	Apriori algorithm	Not mentioned	Minimarket 20 transaction form 1 June-2019 until 20 June-2019	At support of 10% an confidence 100% 4 rules are found most frequently	Graphical representation of Association rules. Use of large dataset. Time and seasonal effects on product purchase.	Results are mentioned clearly
Rinna Rachmatika, Kecitaan Harefa 2020 [30]	Apriori algorithm	Not mentioned	Dataset consisting of 38 sales transaction	At minimum support 10% and confidence 100% to 84% 4 rules	Scatter plotting of confidence and support. Graphical visualization of association rules. Use of large transaction dataset	Results mentioned clearly

4.1.2 Market Basket Analysis using Deep Learning

The varying needs of customers with time and season introduced use of deep learning mthods in market basket analysis on large retail dataset. Market basket anaysis using deep learning methods conducted by [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42].

Doganis,Alexandridis and Sarimveis [31] proposed sales forecasting model for short life shelf milk product genetic selection algorithm for candidate variable selection and radial basis function(RBF) neural network predicts sales using previous and current sales information with less error rate of 4.61%. While Aburto and Weber [32] proposed demand forecasting model for supply chain supermarket using Multi-Layer Perceptron Networks. The demand forecasting model proposed taking into the effects of holidays, independence day, new year and summer vocation on sales of product. Comparatively Das and Chaudhury [33] worked on weekly sales dataset of footwear outlet to forecast using Feedforward and Elman network Considering store location and Seasonal variation on sales prediction. The weekly sales network 5-10-8-1 performs best with least MSE 0.00661666. While Ã, Choi and Yu [34] proposed sales forecasting model for fashion retail using Evolutionary Neural Network (ENN). The proposed approach outperforms than SARIMA model when fluctuations of season effects fashion retail. Similarly, Chen et al [35] proposed sales forecasting model using Ordinary day and holiday moving average method and back propagation neural network. The experiment performed on sales dataset of fresh food sale's store show that BPNN outperforms with less MSE and high precision.

Auon et al [36] proposed prediction model extracting temporal and aggregate features integrating LSTM and QR models into Mixture of Experts (ME) to classify repeated and non-repeated customers. The experiments on transaction dataset of 9 markets show that Mixture of Experts (ME) performs best with less MSE as compared to individual Quantile Regression and LSTM models. While Kaneko and Yada [37] proposed sales prediction model to predict increase and decrease in sales of retail store using deep learning methods considering change in number of product attributes on model accuracy. The experiment on real dataset show that to predict next day sales deep learning outperforms with 10% high accuracy than logistic regression. Similarly, Wang et al [38] proposed prediction model for multi task and multi class demographic attributes of customer for supermarket Retail dataset in china. The proposed model Structured Neural Embedding (SNE) performs best using average aggregating pooling function for Inactive users with high Precision 0.350, Recall 0.281, F1 0.312 Hamming Loss 0.431 and for medium users with high Precision 0.371, Recall 0.289, F1 0.324 Hamming Loss 0.411 and for active users with high Precision 0.361, Recall 0.299, F1 0.327 Hamming Loss 0.410.

Wang et al [39] proposed Multi Task Representation Learning Model (MTRLM) to predict user's demographic attributes extracting features automatically. The MTRL predicts user demographics characteristics from purchase history with high weighted F1 measure for medium group of user Gender 0.645, Marital status 0.802, Education Background 0.647. Comparatively Salehinejad and Rahnamayan [40] predicted Recency, Frequency and Monetary (RFM) values using RNN to mine customer shopping patterns. The auto-encoding method is used to extract features of input variables Customer ID, R, F and M. The proposed model RNN-ReLU outperforms than LSTM-RNN and SRNN to predict RFM values with accuracy 80%. Massaro, Likewise Maritati and Galiano [41] proposed prediction model for 45 store's of walmart using Artificial Neural Network (ANN). The Deep learning method ANN outperforms as best sales predicting algorithm with Correlation 73% / 97,4%, Average Absolute Error 2000 +/- 1250 and Relative Average Error 12,9% +/- 9,9% than other methods Gradient Boosted Trees, SVM, k-NN, Decision Trees and Random Forest. Comparatively Lismont et al [42] proposed model to predict less buying products using Customer-Product network as preprocessing technique. The hybrid Random Forest predicts less buying product with high AUC than neural network.

4.1.3 Market Basket Analysis using Aoociation rules and Deep Learning

The hybrid model used to mine unique customer behvaiuor using both association rules and deep learning methods. The Market basket anaysis using association rules and deep learning conducted by [43], [44], [45], [46], [47].

Matobobo and Osunmakinde [43] worked on real and publically available transaction dataset using Association rules (AR) and Artificial Neural Network (ANN). At minimum support 75% and confidence 70% generated 10 rules using AR with accuracy 83%. At threshold 50% ten rules generated using ANN with accuracy 67%. Comparatively Zekić-Sušac and Has [44] proposed integrated model using Association rules and Artificial Neural Natworks. The 9 best rules are extracted at confidence greater than 50%. The customer's profiles of most buying items are classified using Artificial Neural Networks (ANN) at accuracy of 98.73%. Similarly, Matobobo Courage [45] proposed Association Rule Artificial Neural Network(ARANN) Model for distributed and centralized Retail Enterprise. The transactional data passed to AR model to generated output values of support and confidence. The ANN model used these values as input multiply with weights and summed together. ARANN model identifies pattern at accuracy 83% and 100% on real and public dataset respectively. While Beheshtian-Ardakani, Fathian and Gholamian [46] proposed Product bundling model using clustering, Association rules and classification methods. The Apriori algorithm is used to determine association rules in product clusters. Then classification models support vector machine (SVM), Artificial neural network (ANN), K-nearest neighbor (KNN) and Logistic Regression (LR) are used to determine which product bundle suggested to the byers. The SVM outperforms with high accuracy above 90% than other methods. Comparatively Kilimci et al [47] proposed demand forecasting system on real dataset using deep learning, Time series analysis, Machine Learning and usage of boosting ensemble to forecast best model for demand prediction. The Apriori algorithm used to find correlated items. The third model with Deep learning performs best with 24.7% MAPE.

Table 2 Summary of MBA using Deep Learning on Offline Retail dataset

Author Year	MBA using Deep Learning	Preprocessing	Dataset	Result	Improvement	Good
Philip Doganis, Alex Alexandridis, Panagiotis Patrinos, Haralambos Sarimveis 2005 [31]	Radial Basis Function neural network architecture and Genetic algorithm	Selecting 14 candidate variable by variable selection genetic algorithm	Sales data of fresh milk company in Greece from 2001 to 2002	Adaptive RBF model outperforms with less error rate of 4.61%.	Considering promotion and price in sales forecasting	Use of Real dataset Predicting model for short shelf life milk product
Luis Aburto, Richard Weber 2005 [32]	Multi-Layer Perceptron (MLP) Neural Network	Not mentioned	On year Sales dataset of supply chain Economax supermarket	Neural networks and SARIMAX outperforms on training set with Percentage error 26.12 , Normalized error 0.2760 and on Test set with Percentage error 28.80 , Normalized error 0.3544	Using dataset collected from radio frequency identification	Demand forecasting taking into count the holidays , summer vacation , new year and independence day
Prasun Das, Subhasis Chaudhury 2006 [33]	Feed Forward and Recurrent Neural Network	Transformation of data from database to MS. Excel Min-max Normalization on sale values. Removing items belonging to accessories of footwear, belt, wallet, socks and "T-shirts" Dividing sales data into two parts from April 2002 to March 2004 for training total of 106 weeks and from April 2004 to March 2005.	weekly sales data of footwear company from 2002 to 2005	5-10-8-1 weekly sales network performs best with least MSE 0.00661666	Application of proposed model on other footwear outlet's dataset	Considering store location and Seasonal variation in sales prediction
Kin-Fan Au, Tsan-Ming Choi, Yong Yu 2008 [34]	Evolutionary Neural Network	Not mentioned	Sales data of two fashion products T-shirt and Jeans from 2002 to 2003	ENN outperforms than SARIMA when seasonal trend fluctuate in fashion retail	Taking into count the attribute of products color and size	Experiment performed on real dataset
Chen-Yuan Chen, Wan-I Lee, Hui-Ming Kuo, Cheng-Wu Chen, Kung-Hsing Chen 2010 [35]	Back Propagation Neural Networks (BPNN) And Logistic Regression(LR)	Not mentioned	35 days of fresh food sales at Hi-Life convenience stores. Fresh foods herein are comprised of four kinds of sandwich, three kinds of hand-made rolls, two kinds of rice balls and sushi.	The BPNN outperforms with smaller MSE and high precision to forecast sale	Demand forecasting taking into count the holidays , summer vacation , new year and independence day	Plotting Sale predictions applying BPNN
Gaurangi Anand Auon Haidar Kazmi Pankaj Malhotra Lovekesh Vig Puneet Agarwal Gautam Shroff 2015 [36]	Classification of temporal features using Long Short Term Memory (LSTM) Classification of aggregate features using Quantile Regression (QR)	Customer based feature, product based feature .customer product based feature extraction	Transaction dataset obtained from Kaggle Acquire Valued Shoppers Challenge with 1 year transaction history of shoppers including 9 markets and 38k customers.	Mixture of Experts (ME) performs best with less MSE as compared to individual Quantile Regression and LSTM methods to classify repeated and non-repeated customers for 9 markets.	Use of association rules to mine frequently and non-frequently buying customers Use of large dataset to measure efficiency of proposed model	Integration of deep learning and machine learning methods.

Yuta Kaneko, Katsutoshi Yada 2016 [37]	Deep learning and Logistic Regression	Aggregating sales data according to attributes of product category 1 ,category 2 and category 3 with 62 , 569 and 3312 attributes respectively	POS datasets of supermarket situated in Japan from 2002 to 2004	To predict next day sales deep learning performs best with 10% high accuracy than logistic regression. Using L1 regularization deep learning predictive accuracy increased from 1% to 3%	Use of Large sales transaction dataset. Predicting sales considering Seasonal and temporal effects	Effect of Number of changing product attributes on model accuracy
Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, Xueqi Cheng 2016 [38]	Structured Neural Embedding (SNE)	Filtered dataset with user who has five demographic attributes. Extracted transaction history of users and removed the items bough by less than five times	BeiRen retail dataset of supermarket in china from 2012 to 2013 with 49, 290, 149 transactions, 220, 828 items and 1, 206, 379 users.	The proposed model SNE performs best using average aggregating pooling function with high Precision 0.350, Recall 0.281, F1 0.312 and Hamming Loss 0.431 for Inactive users. For medium users with high Precision 0.371, Recall 0.289, F1 0.324 and Hamming Loss 0.411. For active users with high Precision 0.361, Recall 0.299, F1 0.327 and Hamming Loss 0.410.	Using deep neural architecture than shallow to predict demographic attributes from purchase history. Prediction of more demographic attributes to measure efficiency of proposed model Integrating Association rules to find frequent itemset with deep learning method.	Correlation between demographic attributes. Incorporating Multi task and multi class prediction problems. Turning Multiple prediction tasks into single structured prediction task.
Pengfei Wang, Jiafeng Guo(B), Yanyan Lan, Jun Xu, and Xueqi Cheng 2016 [39]	Multi Task Representation Learning Model (MTRLM) using Feed forward Neural networks	Filtered dataset with items bough by less than 10 times.	BeiRen retail transaction dataset of Chinese supermarket from 2012 to 2013 with 64097 items,80540 users.	MTRL predicts user demographics characteristics from purchase history with high weighted F1 measure for medium group of user Gender 0.645, Marital status 0.802, Education Background 0.647.	Prediction of more demographic attributes to measure efficiency of proposed model. Use of association rules to find frequent itemsets than Bag of items representations	Multitask Representation Learning model based on automatically extracting features than manually defining features
Hojjat Salehinejad Shahryar Rahnamayan 2016 [40]	Recurrent Neural Network (RNN)	Splitting dataset into 50% training, 25% validation and 25% testing.	Ta-Feng grocery shop transaction sales dataset with 817,741 transactions 32,266 users and 23,812 items.	The ReLU-RNN predicts R, F and M values with 80% than LSTM-RNN and SRNN.	Conducting Experiment on Large Training dataset. Using location and age variable for feature extraction. Proposing recommendation system using predicted RFM Values.	Customer behaviour prediction using RFM values.
Alessandro Massaro, Vincenzo Maritati, Angelo Galiano June 2018 [41]	Artificial Neural Network (ANN)	Using "Join" operator to make single dataset, Nominal to Binominal conversion, Split data into training and testing	Walmart 45 store sales forecasting dataset 140000 sales record for each store features.csv scores.csv train.csv test.csv	Deep learning with ANN outperforms as best sales predicting algorithm with Correlation 73% / 97,4% , Average Absolute Error 2000 +/- 1250 and Relative Average Error 12,9% +/- 9,9% than other methods Gradient Boosted Trees, SVM ,k-NN , Decision Trees and Random Forest	Hybrid model to predict sales using AR and Deep Learning. Use of Dimension reduction algorithm on dataset. Applying LSTM OR CNN to predict sales with more high correlation and less absolute and relative Errors.	Worked on Real sales dataset. Comparison of Deep Learning with other methods. Graphical performance comparison plot of applied method ANN with other algorithms Gradient Boosted Trees, SVM ,k-NN , Decision Trees and Random Forest.
asmien Lismont, Sudha Ram, Jan Vanthienen Wilfried Lemahieu, Bart Baesens 12 March 2018 [42]	Customer product network for extracting features, predicting less sold product using classification method Random Forest, Decision tree, Logistic regression, Neural network	Customer-Product network used to extract feature of product and customer	European Grocery shop transactional dataset with 6355 products, 406,678 customers ,100 million transactions during the period of 12 months	Random forest outperforms with average values of AUC using local features 0.6616, Network features 0.6902 and Hybrid features 0.6989. The hybrid Random Forest predicts less buying product with high AUC value.	Considering long time period effects on customer buying behaviour and seasonal effects on product purchase. Conducting current research on online dataset of retail. In local feature considering price and promotion other than RFM values.	First customer product network based study on offline retail dataset.

Table 3 Summary of MBA using AR and DL on Offline Retail dataset

Author Year	MBA using AR and Deep Learning	Preprocessing	Dataset	Result	Improvement	Good
C. Matobobo, I.O. Osunmakinde 2014 [43]	Association rules and Artificial Neural Network	Data Consolidation, Data Propagation, Data Federation, Deleting unnecessary rows. Removing out-of-range values, Dealing with missing values , changing dataset format into format acceptable by the model	Real dataset of Retail Enterprise 85 records and 24 products. Public data downloaded from Informatics (2010) 1000 records and has 11 products five branches with about 200 records in each branch	Using AR at minimum support 75% and confidence 70% 10 rules generated Using ANN at minimum threshold 50% 10 rules are generated. AR identifies pattern at accuracy of 83% and ANN with accuracy of 67% on real data. AR identifies pattern at accuracy of 100% and ANN with accuracy of 75% on Public data.	Large dataset not used	Simulation of proposed method on datasets
Marijana Zekić-Sušac, Adela Has Aug 05, 2015 [44]	Association rules and Neural Networks	After Data cleaning and filtering transaction reduced to 7006 with number of items 278 , customer receipts 3158.	Transactional dataset of Retail store with 14012 transactions	At minimum support 10% and confidence 10% extracted 36 rules, found 9 best rules with confidence larger than 50%. Customer profile classified using Neural networks at accuracy of 98.73%.	Testing Model efficiency using more dataset with large transactions. Seasonal and time stamp effects on purchase of items. Integrating machine learning methods with Association rules.	Integrating Association rules and deep learning with Knowledge management. Beneficial for both retail and marketing managers in designing business strategy using Knowledge management system.
Courage Matobobo 2016 [45]	Association Rules and Artificial Neural Network	Data collection. Data cleaning. Data integration, Data Transformation	**Centralized Dataset**: Real life South Africa Retail Enterprise Three Branches Each branch contains 66 records and 24 products. **Public Data** : 1000 records including seven products five branches 200 records in each branch **Distributed Dataset**: Real life South Africa Retail Enterprise eight branches 66 records from each branch. Public Dataset 1000 transactions each branch contained 200 transactions	**Real Dataset**: AR identifies 6 number of patterns at accuracy 83%ANN identifies 6 number of patterns at accuracy 67%, ARANN identifies 5 number of patterns at accuracy 83%. **Public Dataset**: AR identifies 4 Number pattern at accuracy 75%, ANN identifies 4 number of Patterns at accuracy of 75%, ARANN Identifies rules at Accuracy 100%	Performance plot Of proposed Methods in terms of time used to find patterns.	Simulation of proposed method on datasets. Results are mentioned clearly.

Beheshtian-Ardakani, A., Fathian, M. and Gholamian, April 2017 [46]	Association rule with Apriori algorithm , Artificial neural network (ANN)	Data integration of three databases customers' profile, transactions, and products, Data cleaning by removing missing and noisy values records decreased from 541910 to 406742, Z-score normalization for number of product category. Z-score Normalization of Recency, frequency, and monetary values	Electronic Retail company sales transaction including 541910 records 340 customers during December 2014 to December 2015 in the cities of Golestan province in Iran.	Using K-means clustering highest Silhouette value 0.501 calculated for scenario 8. Five association rules generated for scenario 8 and 1 using Apriori algorithm. SVM performs best to recommend product bundle with accuracy greater than 90%.	Improving accuracy of proposed model Using self-attention with Recurrent Neural network. Mining item to user intention. Proposing price bundling model. Proposing next item recommendation system using customer segmentation and customer loyalty analysis.	Use of three methods Association rules, clustering and classification. Use of customer segmentation and loyalty analysis for product bundling.
Zeynep Hilal Kilimci, A. Okay Akyuz, Mitat Uysal, Selim Akyokus M. Ozan Uysal, Berna AtakBulbul , Mehmet Ali Ekmis 2019 [47]	Multilayer Feedforward Artificial Neural Network (MLFANN), Apriori, Support Vector Regression	Data cleaning, Principle Component Analysis as dimension reduction	SOK's market real sales data with 106 weeks of sales data including 7888 products for 5500 branches. On every branch 1500 products sold out. 2 year data of each store with data size of 875 million records with 155 features	The integration model S1 performs best with 42.4% MAPE on average. The integration model S2 performs best with 25.8% MAPE. The third model with Deep learning performs best with 24.7% MAPE	Extending feature set using location of store, social media and new year event. Applying Deep learning with CNN and RNN. Optimization of weights using heuristic method.	Blending of model based on three approaches. Integration using boosting. Use of real dataset. Effects of Season and Events on sales.

4.2 Online Retail Dataset

The online retail dataset collected from e-commerce site by click-through rate and sales transaction to find customer purchasing habits, likes and dislikes while proposing instant business strategy. In past many researchers performed market basket analysis on online retail dataset using association rules and deep learning methods

4.2.1 Market Basket Analysis using Association Rules

Market basket analysis using association rule mining algorithm on online retail dataset conducted by [48], [49], [50], [51]. M. and Ahmed H [48] proposed association rule mining algorithm simple association rules with multiple minimum (SARMSMC) to mine customer buying habits in less time than previous methods. While Sivri and Cem [49] worked on online store dataset using clustering algorithm Expectation Maximization (EM) to classify data into groups. The Apriori algorithm used to find 6 best rules at support 50%, confidence 50% and Lift 100%. Similarly, Fang et al [50] worked on product bundeling problem using Association rules and clustering algorithms. The Apriori algorithm used to find 3 best rules at support 0.1, confidence 0.01. The clustering method K-means applied to find 4 clusters of Customers on values of R, F and M to recommend product bundle. Comparatively Lok, Wang and Xu [51] worked on visualization of association rules. Graph base visualization for large dataset is not enough when customer purchasing pattern changes frequently. The R packages Igraph and visNetwork are used for dynamic visualization of association rules.

4.2.2 Market Basket Analysis using Deep Learning

Market basket anaysis using deep learning methods on online retail dataset while proposing next item recommendadation system , predicting customer purchase behaviour , sales forecasting by [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [67], [68], [69], [70], [71], [72].

Kim, Kim and Lee [52] proposed sales prediction model using genetic based classification of neural networks on sales dataset. The results of
three neural networks NN1, NN2 and NN3 integrated into single GA-based
method with classification rate 76.5% and error rate 13.5 to predict
purchased product. Comparatively Bhargav, Mathur and Bhargav [53] worked on transaction dataset of apparel using ANN Feedforward network. Market basket analysis using ANN reduces number of scanning database to find candidate itemset. While Wang et al [54] proposed next item recommendation using aggregating complicated interactions of item and user using nonlinear max pooling function at both first and second layers. The experiment performed on three dataset shows that HRM performs best on three datasets with high F1score, Hit Ratio and NDCG. Similarly, Wan et al [55] proposed novel neural network approach NN-Rec to predict next basket recommendation of two Real Retail datasets Tafeng and Beiren. The feature vector transformed from user id and items id's feed into embedding layer. This model captures long dependencies and input layer more flexible to add other features than previous methods. The Experiment performed on Beiren dataset with accuracy below 0.1 when baskets k= 2 and for Tafeng dataset with accuracy below 0.06 when basket k=2. Likewise, Yu et al [56] proposed novel model named Dynamic Recurrent Basket Model (DREAM), based on Recurrent Neural Network (RNN) considering dynamic user's interest and global interactions of all baskets of user over the time. Result shows non liner operation max pooling outperforms than Avg pooling method because it measures the interactive relationship of basket items.

Gangurde, Kumar and Gore [57] proposed novel model to find frequent item set from seasonal shopping dataset by changing the weights through backpropagation minimizing the time and cost at accuracy of 0.90. While Li, Hwang and Chang [58] dynamically predicted preference for next item purchasing using purchase information of customer extracted from Personal Purchasing Preference Pattern (PPPP). The Recurrent Neural Network (RNN) applied to extraction info of PPPP to find next purchasing item of customer. Comparatively Wang et al [59] proposed attention based recommendation model using neural network approach. ATEM consists of three layers input layer, item embedding layer, Attention Layer, context embedding layer, and output layer. The performance of ATEM evaluated using two real sales transactional dataset of Statistics IJCAI-15 and Ta-Fang. While Yu et al [60] proposed Sequential Hierarchical Attention Network (SHAN) combining long and short term intentions of customer. The experiment performed on two dataset show that SHAN On Gowalla dataset with high AUC 0.801 and SHAN-S with high Recall 0.156 on T-Mall dataset. Comparatively Xia et al [61] proposed multitask learning model using Long short-term memory networks (LSTM) to mine purchasing behaviour of customers. The novel work is extracting items, users and product categories embedding to predict conversion rate for product recommendation. While Bai et al [62] proposed a novel Attribute-aware Neural Attentive Model (ANAM) to predict next basket recommendation. The novel work is attention base mechanism applied on item and attribute information. The experiment performed on two datasets first encoding item and basket information then assigning weights to integrate item and its attribute. Comparatively Sakar et al [63] proposed hybrid model using Multilayer Perceptron Network (MLP) and RNN-LSTM to predict purchasing behaviour of customer. The novel work is use of filter based feature selection methods. The MLP outperforms than base classifiers at Accuracy 87.24, F1-score 0.86, True-positive rate 0.84 and True-negative rate 0.92.

Chen and Li [64] proposed Attention Base Recurrent Neural Network to recommend item for repurchase. The novelty of work is introducing self-attention mechanism rather than feature engineering base recommendation model. The other distinguish factor is use of LSTM to mine periodic purchase behaviour of users. The Result shows that AttRNN outperforms than LSTM at confidence level 0.2 with F1-score 0.4723.

While Le, Lauw and Fang [65] proposed Basket Sequence Correlation Network model for next basket prediction. The novel work is use of correlation matrix to measure frequency of co-occurrences between items. The experiment performed on three real dataset shows propose Beacon model outperforms than other baseline models with high F1 score. Comparatively Zhang et al [66] proposed Feature-level Deeper Self- Attention model (FDSA) to mine patterns between items and features of items. The novel work is capturing Sequential pattern between item and both implicit and explicit features using self-attention block. The output of these two attention blocks are passed into next layer to predict next item recommendation. On the other hand, Sreenivasa and Nirmala [67] proposed hybrid model to mine customer buying behaviour on retail dataset of T-Mall online shopping store using RNN and FNN. The novel work is considering location base information of user and using dynamic short and long term intention of user. To attain short term intention transition matrices are used from past transaction of customers. While Cirqueira, Helfert and Bezbradica [68] predicted single, multiple and pre-trained embedding of next day, hour and purchase category using LSTM model on InstaCart shopping store's dataset. The novel work is guide to preprocess multi-intent of customer using neural network embedding strategies.

Bai et al [69] proposed novel Long-Short Demands-aware Model (LSDM) to predict next item. The novel work is considering long and short demands of user over different weekly time scale implementing hierarchical neural network structure. While Liu, Li, et al [70] worked on multi intent pattern mining using new framework named as Multi-Intent Translation Graph Neural Network (MITGNN) for next basket recommendation. Experiments performed on two real datasets of grocery store InstaCart and Walmart. Similarly, Liu, Wan, et al [71] proposed basket recommendation framework BasConv based on graph convolution Neural Network. The novel work is heterogeneity of user, basket and item in interactive layer as compared to previous work. The proposed framework performed at Recall 0.2092, NDCG 0.2281 and HR 0.7712 on InstaCart dataset and at Recall 0.0530, NDGC 0.0841 and HR 0.3394 on Walmart dataset. Comparatively Lee et al [72] proposed multi period recommendation model using Recurrent Neural Network (RNN). The novel work is Proposed model evaluated by multiple time periods because customer purchasing behaviour changes over time but traditional recommendation model evaluated only once.

4.2.3 Market Basket Analysis using Aoociation rules and Deep Learning

The Market basket anaysis using association rules and deep learning on online retial dataset conducted by [73], [74], [75], [76], [77].

Changchien and Lu [73] proposed online recommendation system for electronic store using clustering and rule extraction module. The SOM neural network based architecture used to find 9 clusters from O-ID, Buyer, Receiver, Product table. The rule extraction model used to find 99 association rules at minimum confidence 0.25. While Shim, Choi and Suh [74] proposed business strategy for small online shop using classification and Association rules. The classification method used to classify VIP and Non VIP customers using RFM values. The classification with decision tree outperforms with accuracy 99.83%. The buying behaviour of VIP customers determined by finding 11 association rules at minimum support 4% and confidence 40% for categories, minimum support 3% and confidence 20% for subcategory. Comparatively Mansur and Kuncoro [75] worked on sale of antique furniture accessories. Using Karomah Brass sale's transaction 21 rules are generated at minimum support 3 and confidence 80%. ANN backpropagation is used to predict product quantity of most buying products for next year backpropagation at epoch 7 and MSE 0.000788252. On the other hand, Itkar and Uday [76] proposed efficient algorithm for mining frequent itemsets using Artificial neural network based approach auto associative memory. The Proposed correlation matrix memory (CMM) algorithm takes less time while generating candidate itemset than Apriori, FP growth, LCM and CT-PRO at minimum support 50% to

90%. Comparatively Ghadekar and Dombe [77] proposed hybrid recommendation model using Deep Learning and Association rules. The Convolution Neural Network used to classify input image and then this image is passed as input to recommendation system. The most buying product applying association rule recommended using product id.

Table 4 Summary of MBA using Association Rules on Online Retail dataset

Author Year	MBA using AR	Preprocessing	Dataset	Result	Improvement	Good
Walaa M., Ahmed H., Hoda K 2004 [48]	simple association rules with multiple minimum supports (SARMSMC)	Not mentioned	Clickstream dataset of Two ecommerce sites BMS-Web View-1 with Number of transactions: 27,736 Number of individual Items: 348 Xml File size: 6.07 MB AdventureWorks DW with Number of transactions: 21,255 Number of individual Items: 37 Xml File size: 5.58 MB	The proposed algorithm consumes less time to generate association rules and faster by 95% than other methods	Seasonal and temporal effects on purchasing behaviour of customers using large online transaction dataset	Proposed new algorithm to generate association rules in less time
Elif Şafak Sivri, Mustafa Cem Kasapbaşı, Fettullah Karabiber 2014 [49]	Clustering algorithm Expectation Maximization (EM) to classify demographic attributes , Association rules extraction using Apriori algorithm	Data filtering	Online sales transaction dataset of retail company in Turkey with 14000 transactions	At support and confidence 50% and Lift 100% extracted 4 association rules.	Seasonal and temporal effects on purchasing behaviour of customers using large online transaction dataset	Use of clustering and Association rules to mine frequent patterns. Extracting demographic attributes of customers using Clustering
Yan Fang, Xinyue Xiao, Xiaoyu Wang, Huiqing Lan 2018 [50]	Association rule using Apriori Algorithm, K-means clustering find customer segment using Recency frequency and monetary (RFM) model	Integration data from four database customer, Transaction, products and products classed. Z-score normalization	Electronic Sales Transaction dataset with10281 customers, 251396 records, 1560 products and 110 product categories.	At minimum support 0.1 and confidence 0.01 three rules generated using Apriori to find necessary products for bundling. 4 cluster of Customers using K-means identified on values of R, F and M.	Proposing price bundling method by using customer attribute and demands.	Dynamic product bundling to save both time and energy of customer.
Jun Haur LOK, Xingwen WANG, Shifeng XU 2020 [51]	Market basket Analysis With network graph	Not mentioned	InstaCart website Dataset of 33.8 million from over 200,000 unique customers	Dynamic visualization of association rules with Network graph helps to design flexible business strategy	Limited Research scope	Problem statement defined clearly Source of Dataset with proper dimensions

Table 5 Summary of MBA using Deep Learning on Online Retail dataset

Author Year	MBA using Deep Learning	Preprocessing	Dataset	Result	Improvement	Good
Eunju Kim, Wooju Kim, Yillbyung Lee 2002 [52]	Genetic based classification of neural networks	Not mentioned	Sales dataset of EC company in Korea. Data set consists of 10 demographic features , 5 transactional feature during one year.	Integration of results from three neural networks NN1,NN2 and NN3 into Single GA-based method show prediction of target product purchase with classification rate 76.5% and error rate 13.5	Applying rank and abstract level classifiers	Experiment performed on two datasets
Anshul Bhargav, Robin Prakash Mathur, Munish Bhargav 2014 [53]	Market basket Analysis with ANN	Not mentioned	Handmade 4 inputs Jacket, Sweater, Jeans and T-Shirt	Valid combination for summer and winter seasons are (3,4) (1,3,4) and (1,2	Simulation on real retail dataset Performance plot of ANN Simulation of frequently buying product	Simulation of candidate itemset. Reduced time and number of scanning database
Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, Shengxian Wan, Xueqi Cheng CAS 2015 [54]	hierarchical representation model (HRM) based on neural network	Removed items bought by less than 10 User for Ta-Feng and BeiRen dataset Removed items bought by less than 3 users for T-Mall dataset Splitting dataset into training and testing data. In Testing set only last transaction of each user and in Training all remaining transaction of users	Ta-Feng dataset with users |U| 9238 items |I| 7982 transactions T 67964 avg. transaction size 7.4 avg. transaction per user 5.9. BeiRen dataset with users |U| 9321 items |I| 5845 transactions T 91294 avg. transaction size 9.7 avg. transaction per user 5.8 T-Mall dataset with users |U| 292, items |I| 191 transactions T 1805 avg. transaction size 5.6 avg.transaction per user 1.2	$HRM_{Max,Max}$ with nonlinear max pooling function at both first and second layer performs best on three dataset with high F1score,Hit Ratio and NDCG	Dimension reduction of dataset. Aggregating other complicated information like transaction timestamp.	Resolving limitations of sequential and general recommendation systems by taking nonlinear combination among purchase
Shengxian Wan, Yanyan Lan, Pengfei Wang, Jiafeng Guo, Jun Xu, Xueqi Chen 2015 [55]	Novel neural network NN-Rec	Removed Items bought less than 10 times and users bought less than 10 items or 4 baskets. Transformation of inputs to user matrix and item matrix form feature vector	Tafeng grocery store in china # of users 9,238, # of items 7,973, # of baskets 77,202. Beiren contains user shopping history from 2011 to 2013 # of users 13,736, # of items 5,920, # of baskets 522,963	For Beiren dataset accuracy is below 0.1 when baskets k= 2. For Tafeng dataset accuracy is below 0.06 when basket k=2.	Time series analysis of dataset	Worked on two real datasets. Performance plot of Proposed method on two real datasets.

Feng Yu, Qiang Liu, Shu Wu, Liang Wang, Tieniu Tan 16 July 2016 [56]	Dynamic Recurrent basket Model (DREAM), based on Recurrent Neural Network (RNN)	Dataset preprocessed for each item purchased by at least k users. For Ta-Feng k=10 and For T-mall k=3	**Ta-Feng:** 817,741 transactions belonging to 32,266 users and 23,812 items. **T-mall:** 4,298 transactions of 884 users and 9,531 brands	**Using Avg-Pooling:** For Ta-Feng dataset with dimensions {50, 100, 150} f1-Score 0.061 NDCG 0.082, f1-score 0.064 NDCG 0.081, f1-score 0.067 NDCG 0.083 respectively. For T-mall dataset with dimensions {10, 15, 20} f1-Score 0.058 NDCG 0.141, f1-score 0.063 NDCG 0.154, f1-score 0.066 NDCG 0.160 respectively. **Using Max Pooling:** For Ta-Feng dataset with dimensions {50, 100, 150} f1-Score 0.065 NDCG 0.084, f1-score 0.068 NDCG 0.085, f1-score 0.070 NDCG 0.086 respectively. For T-mall dataset with dimensions {10, 15, 20} f1-Score 0.070 NDCG 0.162 , f1-score 0.071 NDCG 0.168, f1-score 0.073 NDCG 0.173respectively	Heterogeneity of Basket Recommendation Problem. Effects of season and time on datasets.	Use of linear and nonlinear pooling methods for basket representation Performance plot of applied methods with metrics on two datasets
Roshan Gangurde, Dr. Binod Kumar, Dr. S. D. Gore 2017 [57]	Market basket Analysis with Feed forward Neural network	Data cleaning using proposed EHCleaner algorithm	Handmade 4 inputs Biscuit, Cold drinks, Tea and Fast food	4 Valid combinations for rainy phase are (1,3,4), (1,3), (1,4) and (3,4)	Source of dataset not mentioned Visualization of data is missing using graph base and scatter plot	Performance plot of FFNN with MBA algorithm Simulation on Transactional dataset
Yun-Rui Li, Ting-Kai Hwang and Shi-Chung Chang 2018 [58]	Market Basket Analysis using Recurrent Neural Network with Long Short Term Memory	Splitting transactional data into purchase history and shopping cart data	InstaCart Online Grocery Shopping Dataset1 in 2017 206,209 Customer's transaction history including 3,421,083 orders categorizes 49,685 products into 21 departments and 134 aisles	RNN based Personal Purchasing Preference Pattern achieves 18.29% higher Individual Prediction Accuracy than Baseline	Simulation of proposed method using transactional dataset Visualization of data Prediction results of next item not shown	Performance metrics of proposed method. Real time marketing strategies using Proposed model

Shoujin Wang, Liang Hu, Longbing Cao, Xiaoshui Huang, Defu Lian, Wei Liu 2018 [59]	Attention based neural network	Removed Transactions that contains only one item	Two Real world transaction dataset IJCAI-15 and Ta-Fang **IJCAI-15**: #Transactions 144,936, #Items 27,863, Avg. Transaction Length 2.91, #Training Transactions 141,840, #Training Instances 412,679, #Testing Transactions 3,096, #Testing Instances 9,030 **Ta-Fang**: #Transactions 19,538, #Items 5,263, Avg. Transaction Length 7.41, #Training Transactions 18,840, #Training Instances 141,768, #Testing Transactions 698,#Testing Instances 3,150	For IJCAI-15Attention Based Transaction Embedding Model performs at REC@10 0.3542, REC@50 0.5134, MRR 0.2041. For Ta-Fang Attention Based Transaction Embedding Model performs at REC@10 0.1089 , REC@50 0.2016 , MRR 0.0347.	Proposing Hybrid model using Attention based mechanism With RNN and CNN	Dataset mentioned with proper dimensions Worked on Recommending novel item other than rigid order assumption using attention mechanism
Haochao Ying, Fuzhen Zhuang, Fuzheng Zhang, Yanchi Liu, Guandong Xu, Xing Xie, Hui Xiong and Jian Wu July 2018 [60]	Sequential Hierarchical Attention Network (SHAN)	Items observed by less than 10 users are removed from dataset	T-Mall #user 20,648 #item 25,129 avg. session length 2.73 #train session 71,892 #test session 3,534 Gowalla #user 15,171 #item 13,193 avg. session length 2.97 #train session 129,225 #test session 3,635	SHAN performs best on Gowalla dataset with AUC 0.987, Recall 0.439. On T-Mall dataset SHAN performs best with high AUC 0.801 and SHAN-S performs best with high Recall 0.156 on T-Mall.	Capturing Feature in long and short term intentions Worked on Large datasets	Combining user dynamic long and short term intentions. Hierarchical attention network to capture long and short term intentions. Capture high level complex interactions between item-item and user-item factors.
Qiaolin Xia, Peng Jiang, Fei Sun, Yi Zhang, Xiaobo Wang, and Zhifang Sui 2018 [61]	Long Short-Term Memory Networks (LSTM)	Extracting numerical and embedding features of user, items and user-item interactions. Excluded user with less than 10 buying behaviour	Data collected from Taobao.com from 7 days' consumer buying behaviour with 24,634 number of customer, user click rate 442.7, buy rate 4.3.	To predict consumer buying behaviour best result achieved with three classes category embedding, multi embedding and behaviour embedding in terms of AUC 0.8589, P@5 0.1952, MAP@5 0.5173, MAP@10 0.5185, MAP@20 0.4890	Predicting buying behaviour considering seasonal and temporal features on large sales transaction dataset	First step to predict consumer buying behaviour in recommendation systems considering item, category and multi embedding
Ting Bai, Jian-Yun Nie, Wayne Xin Zhao, Yutao Zhu, Pan Du, Ji-Rong Wen 27 June 2018 [62]	Attentive Recurrent Neural Network (ARNN) Long Short Term Memory (LSTM)	Users with Purchase less than 10 and 25 in Ta-Feng and JingDong datasets are removed.	Ta-Feng 4 months' transaction from November 2000 to February 2001. # Users 9,238 #Items 7,973 # Transactions 464,118 # Category 1,074 JingDong product reviews dataset of 4 months from January 2012 to April 2012 # Users 4,832 #Items 3,283 # Transactions 41,932 # Category 165	ANAM models outperforms than baseline methods as experiment performed on two datasets with F1-score@5 0.146, NDCG@5 0.190 for Ta-Feng and with F1-score@5 0.1313, NDCG@5 0.1842 for JingDong.	Adding more attribute Information. Dimension reduction of datasets.	Attention mechanism applied on item and attribute. Item attribute considering while recommending next basket item

C. Okan Sakar , S. Olcay Polat, Mete Katircioglu, Yomi Kastro May 2018 [63]	Multilayer perceptron (MLP), Long Short Term Recurrent Neural Network (LSTM)	Oversampling and feature selection	Online Retail data consisting of 185,000 Web pages visited in 9800 sessions of 3500 visitors.	In first module MLP performs best than SVM and Random forest to predict purchasing intention of visitor at Accuracy 87.24, F1-score 0.86, True-positive rate 0.84 and True-negative rate 0.92. In second module LSTM-RNN estimated visitor intention to leave the site without transaction.	Integrating recommendation system on user or item based.	Real time customer behaviour analysis using hybrid model.
Pengda Chen, Jian Li 2019 [64]	Attention Recurrent Neural Network (AttRNN) Long Short Term Memory (LSTM)	Not mentioned	InstaCart Dataset with 10,000 shopping record of users, average 4 to 100 shopping baskets per user.	AttRNN performs better than LSTM with F1-score 0.4723 at confidence level 0.2	Hybrid model to mine periodic purchase rules Attention based mechanism with CNN	Periodic purchase rule Use of self-Attention mechanism
Duc-Trong Le, Hady W. Lauw and Yuan Fang August 2019 [65]	RNN using correlation matrix	Number of items and users for Tafeng, Delicious and foursquare are 10,5,5 respectively. Splitting Tafeng dataset into training, validation and test as (3, 0.5, 0.5) months. Splitting Delicious dataset as (80, 2, 2) months. Splitting Foursquare dataset as (10, 0.5, 0.5) months. Removing baskets longer than 30 using prefix cut off.	Tafeng grocery transaction dataset from Nov 2000 to Feb 2001. #Sequence 77 209, #Item 9 964, Average length 7.0 and Average basket size 5.9. Delicious dataset consists on user's sequence of bookmark. #Sequence 61 908, #Item 6 520, Average length 21.4 and Average basket size 3.8. Foursquare user's check ins from August 2010 to Jul 2011. #Sequence 100 980, #Item 5527, Average length 22.2 and Average basket size 1.8.	Basket Sequence Correlation Networks (BEACON) performed better with high F1 score on three datasets Tafeng, Delicious and Foursquare.	Hybrid model to recommend next item using AR and Deep learning. Reducing dataset Dimensions.	Use of correlation matrix to predict correlation between items. Experiments om three real datasets.
Tingting Zhang, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Deqing Wang, Guanfeng Liu and Xiaofang Zhou, 2019 [66]	Feature Level Deeper Self-attention (FDSA)	Filtered Toys and Games dataset with users who rated less than 5 items and items rated by less than 10 users. Filtered T-Mall dataset with users who rated less than 15 items and items rated by less than 30 users.	T-Mall dataset with # users 16,257 # items 18,678 # avg. actions/user 15.98 # Ratings 276,117. Toys and Game dataset with # users 35,124 # items 28,351 # avg. actions/user 5.51 # Ratings 228,650.	FDSA model outperforms than baseline methods on T-Mall dataset with item based self-attention block 4 and feature based self-attention block 2 on Toys and Games dataset, on T-Mall dataset with item based self-attention based block 2 and feature based self-attention block 4.	General recommendation with self-attention Dimension reduction of dataset. Transaction timestamp self-attention block	Sequential pattern mining between item and both implicit and explicit features using self-attention mechanism. Capturing implicit features from heterogeneous attributes of item. Vanilla attention mechanism to attain implicit features.

B. R. Sreenivasa, C. R. Nirmala 12 June 2019 [67]	Recurrent Neural Network (RNN), Feed Forward Neural Network (FNN)	Splitting 70% dataset into training and 30% into testing.	T-Mall online shopping dataset 200,000 shopping records with 1000 users and 10,000 items. Dataset contains four behaviour of customer purchasing, clicking, add to cart and add to favorite.	The Hybrid Location Centric Prediction model HLCP outperforms than other hybrid models with Hit Ratio(HR) 0.08547365 and Mean reciprocal rate (MRR) 0.12935566.	Proposed model evaluation using Amazon dataset. Considering time centric base information to predict next item recommendation.	Using Customer Dynamic purchase behaviour considering location base information of user while product recommendation.
Douglas Cirqueira, Markus Helfert, Marija Bezbradica September 2019 [68]	Multi task Long Short Term Memory based model	Splitting Transaction attribute into four bins "morning", "afternoon", "night" and "dawn time". Splitting dataset into 90% of transaction for training dataset and 10% as testing data.	InstaCart online grocery shopping cart sales transaction data from 2017 with 3,421,083 million of orders.	LSTM predicted single transaction embedding of next day purchase accuracy 44%, next hour 56% and next purchase category 82%. Predicted multiple transaction embedding of next day 30%, Next purchase 57% and purchase category 81%. Predicted Pre-trained embedding of next day 32%, next hour 62% and purchase category 82%	Experiment conducting on Multi context embedding using additional dimensions as location and customer age and gender. Use of large dataset to evaluate embedding strategies.	Guide to preprocess customer multi-intention behaviour using embedding. Strategies for learning customer's multi-intent embedding. Multitask LSTM model to learn embedding of multitask transaction.
Ting Bai, Du Pan, Wayne Xin Zhao, Ji-Rong Wen, Jian-Yun Nie 12 Feb 2019 [69]	Long short Term (LSTM) to predict next item recommendation	In Tafeng dataset removed the products bought less than 15 times. Same filtering method applied on BeiRen dataset. Removed product purchased less than 5 times.	Ta-Feng transaction dataset from December 2000 to February 20017 with 7,044 items 1,951 users total 90,986 purchase records. Average purchase record of users 50 average purchase time 14. BeiRen online shopping dataset from April 2013 to July 2013 with 211,519 purchase records 3,264 users and 5,818 items. Average purchase record of user 65. Amazon dataset collected from product reviews from January 1st, 2014 to June 30th, 2014. 6,092 items 1,443 users with 15,811 purchase record. average product reviewed by user is 11.	LSDM performance metrics are On Ta-Feng dataset Hit@5 0.1194, Hit@10 0.1281, NDCG@5 0.0824, NDCG@10 0.0890. On BeiRen dataset Hit@5 0.2187, Hit@10 0.2290, NDCG@5 0.1617, NDCG@10 0.1646. On Amazon dataset Hit@5 0.0182, Hit@10 0.0265, NDCG@5 0.0119, NDCG@10 0.0147.	Incorporating attribute of item as price and category. Detecting Best time scale from dataset automatically.	incorporated user's demand towards products over specific time. Experiments performed on three real e-commerce datasets. Considering multiple time scale for long demands using hierarchical neural structure. Use of Attention mechanism to capture user's intention.

Zhiwei Liu, Xiaohan Li, Ziwei Fan, Stephen Guo, Kannan Achan, and Philip S. Yu 22 October 2020 [70]	Multi-Intent Translation Graph Neural Network (MITGNN)	Filtered basket for InstaCart dataset with less than 30 items and User with less than 5 baskets. Filtered basket for Walmart data with less than 40 items and user with less than 5 baskets.	**InstaCart** # Edges 1, 271, 195 32, #Baskets 32, 201 2, #Users 2, 904, #Items 30, 754, Density 0.0293% **Walmart** # Edges 1, 225, 155, #Baskets 27, 797, #Users 7, 110, #Items 50, 408, Density 0.0168%	**Transductive Recommendation:** InstaCart dataset improved Re-call, HR, and NDCG by 12.03%, 9.43%, and 9.71%, respectively. On Walmart dataset improves the Recall, HR, and NDCG by 17.82%, 17.17%, and 10.75%, respectively. **Inductive Recommendation:** On InstaCart data improved Recall, HR, and NDCG on average by 6.85%, 0.63%, and 6.80%, respectively. On Walmart dataset the Recall, HR, and NDCG on average by 16.92%, 5.74%, and 6.41%, respectively	Effect of season and time on dataset. Simulation for Next basket predicted items	Multi intent pattern mining. Performance plot of proposed method on real time datasets. Solved Transductive and Inductive basket Recommendation Problem Performance of proposed method with metrics
Zhiwei Liu Mengting Wan Stephen Guo Kannan Achan Philip S. Yu 8 May 2020 [71]	BasConv based on the graph convolutional neural network	Filtered Basket with less than 30 items and 40 items for InstaCart and Walmart datasets	**InstaCart:** #User 22, 168, #Item 40, 044, #Basket 65, 672, Avg. Basket 2.96, Avg. Size 37.0, #Interaction 2, 495, 695. **Walmart:** #User 44, 218 77, #Item 77, 599, #Basket 130, 707, Avg. Basket 2.96, Avg. Size 52.5 2, #Interaction 6, 997, 572.	On InstaCart within basket recommendation using BasConv at Recall 0.2092, NDCG 0.2281 and HR 0.7712. On Walmart within basket recommendation using BasConv at Recall 0.0530, NDGC 0.0841 and HR 0.3394.	Effect of season and time on dataset. Out of basket recommendation	Simulation for next predicted item Heterogeneity of Basket Recommendation Problem
Hea in Lee, Il Young Choi, Hyun Sil Moon and Jae Kyeong Kim 29 january 2020 [72]	Recurrent Neural Network (RNN)	Not mentioned	Real dataset of Fresh Food delivery company with 7716 customers and 10 shopping carts	The LSTM based recommendation model outperform collaborative filtering based model with accuracy at 21% high at T and 10% high at T+4.	Use of diverse transactional dataset Proposing hybrid model using CF and RNN based recommendation model	Multi period product recommender Results mentioned clearly Use of real dataset Dataset mentioned with proper dimensions

Table 6 Summary of MBA uisng Association Rules and Deep Learning on Online Retail dataset

Author Year	MBA using AR and Deep Learning	Preprocessing	Dataset	Result	Improvement	Good
S. Wesley Changchien, Tzu-Chuen Lu 2001 [73]	Clustering module using neural networks, Self-organization map and rule extraction module using rough set theory	Creating fact table, from database for mining, selecting dimensions and attributes, filtering data with noise and handling missing values Data transformation and normalization	Electronic store dataset for online marketing	using SOM find 9 clusters from O-ID, Buyer, Receiver and product. 99 association rules extracted at minimum confidence 0.25	Applying proposed approach on sales and product dataset considering seasonal effects	Use of hybrid approach for online recommendation Dimension reduction of dataset
Beomsoo Shim, Keunho Choi, Yongmoo Suh Business 2012 [74]	Classification using Decision Tree, Artificial Neural Network, Logistic regression ,bagging, Association rules	Removed duplicated and missing values, Computed normalized values of R,F and M.	Sales transaction dataset of Online shopping mall in Korea from 2008 to August, 2010 with customer 3445, purchased item 14,782 and 11,033 transaction data	To classify VIP and Non VIP customer Decision Tree and Decision tree with bagging outperforms with accuracy of 99.83%. From VIP transaction data 16 Association rules are extracted at minimum support 4% and confidence 40% for categories, minimum support 3% and confidence 20% for subcategory.	Use of Large sales transaction dataset. Improving accuracy to classify VIP customers. Use of RNN and LSTM as classifier.	Customer purchasing pattern identification using RMF, classification and Association rules
Agus Mansura, Triyoso Kuncoroa 2012 [75]	Market basket Analysis with Association rules and ANN Backpropagation	Not mentioned	Antique Furniture accessories named Karomah Brass from September 2010 to August 2011, 564 sales transaction	At minimum support 3 and confidence 80% 21 rules are generated containing 17 items. Antique cabinet accessories most buying product. Product quantity predicted at epoch 7 and MSE 0.000788252.	Product cost not considered while prediction for inventory. Simulation of predicting product quantity	Results are mentioned clearly. Network performance graph of training data. Predicted Quantity list of most buying products at lowest MSE.

Suhasini Itkar, Uday Kulkarni 2014 [76]	Market basket Analysis with Association Rules and ANN based Auto associative memory using CMM	Not mentioned	Dataset download from Frequent Itemset Mining Implementations Repository (FIMI'04) website Mushroom dataset 8124 transaction ,119 items average transaction length 23 Chess dataset 3196 transaction 75 items average transaction length 37 Accidents datasets 340183 transactions 468 items and average transaction length 34 BMS- Web View -2 dataset	At minimum support threshold from 50 to 90% the proposed CMM algorithm shown better efficiency than Apriori, FP-growth, CT-PRO and LCM algorithms.	Improving efficiency of proposed algorithm using Large dataset	Simulation of proposed algorithm Performance plot of performed experiments Reducing number of database scans in generation of candidate itemset
Premanand Ghadekar, Anay Dombe September 2019 [77]	Convolutional Neural Networks (CNN) And Association Rules using Apriori Algorithm	Splitting 80% data for training and 20% for testing, Normalization on images Singular value decomposition to identify features of purchased products	Image dataset consisting of 44000 images. Manually generated transaction dataset from Amazon	The image classification using Convolution Neural Network is 70%. The most frequent buying item generated at 80% confidence to recommend items.	Resolving overfitting issue in proposed model. Use of large dataset. Improving accuracy of proposed model	Proposed Hybrid model using Association rules and CNN. Product recommendation Based on item to item, user to user similarity and user interest based on social media

CONCLUSION & FUTURE WORK

Application of data mining techniques in market basket analysis is an emerging trend in retail. This paper identified seventy articles related to market basket analysis in retail published between 2000 to 2020.It aims to give a structured review on market basket analysis in retail using association rules and deep learning methods. Market basket analysis using association rule consists on twenty-seven articles. Of these, twenty-three and four articles related to offline and online retail dataset respectively. Market basket analysis using deep learning methods included thirtythree artcles. Of these, twenty-one and twelve articles related to online and offline retail dataset respectively. Market basket analysis using association rules and deep learning methods consist of ten articles.

The limitations identified as need of dimension reduction methods on offline retail dataset using association rules, use of attention and embedding layers in deep learning methods on offline retail datasets to achieve more accurate results. In future more research work is required to develop hybrid framework due to growing retail dataset at rapid speed and frequently changing customer buying habits.

REFERENCES

[1] Blattberg R.C., Kim BD., Neslin S.A. (2008) "Market Basket Analysis. In: Database Marketing". International Series in Quantitative Marketing, vol 18. Springer, New York, NY.

[2] Berry, M.J.A., Linoff, Gordon.S. (2004) "Data Mining Techniques: for Marketing, Sales and Customer Relationship Management" (second edition), Hungry Minds Inc.

[3] D. Grewal, A. L. Roggeveen, and J. Nordf ̈alt, "The future of retailing, "Journal of Retailing, vol. 93, no. 1, pp. 1–6, 2017.

[4] Zhang G.P. (2009) "Neural Networks for Data Mining". In: Maimon O., Rokach L. (eds) Data Mining and Knowledge Discovery Handbook.

[5] Kubiak, B. F., & Weichbroth, P. (2010). "Cross- And Up-selling Techniques in ECommerce Activities". Journal of Internet Banking and Commerce, 15(3).

[6] Svetina, M., and J. Zupancic. (2005) "How to Increase Sales in Retail with Market Basket Analysis." System Integration. pp. 418-428.

[7] Ahmed, S. R. (2004) "Applications of Data Mining in E-Business and Finance", *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04)*, p. 157.

[8] Chen, M. C., Chiu, A. L. and Chang, H. H. (2005) "Mining changes in customer behavior in retail marketing", *Expert Systems with Applications*, 28(4), pp. 773–781. doi: 10.1016/j.eswa.2004.12.033.

[9] G ̈urdal Ertek, A. D. (2006) "A Framework for Visualizing Information", pp. 593–602. doi: 10.1007/978-94-017-0573-8.

[10] Nafari, M. and Shahrabi, J. (2010) "A temporal data mining approach for shelf-space allocation with consideration of product price", *Expert Systems with Applications*. Elsevier Ltd, 37(6), pp. 4066–4072. doi: 10.1016/j.eswa.2009.11.045.

[11] Avcilar, M. Y. and Yakut, E. (2014) "Association Rules in Data Mining: An Application on a Clothing and Accessory Specialty Store", *Canadian Social Science*, 10(3), pp. 75–83. doi: 10.3968/4541.

[12] Zulfikar, W.B.; Wahana, A.; Uriawan, W.; Lukman, N. "Implementation of association rules with apriori algorithm for increasing the quality of promotion". In Proceedings of the 4th International Conference on Cyber and IT Service Management, Bandung, Indonesia, 26–27 April 2016, pp. 4–8. doi: 10.1109/CITSM.2016.7577586.

[13] Abusida, A. M. and Gültepe, Y. (2019) "An Association Prediction Model: GECOL as a Case Study", *International Journal of Information*

Technology and Computer Science, 11(10), pp. 34–39. doi: 10.5815/ijitcs.2019.10.05.

[14] Nur, Y., Triayudi, A. and Diana, I. (2019) "Implementation of Data Mining to Predict Food Sales Rate Method using Apriori", *International Journal of Computer Applications*, 178(35), pp. 22–28. doi: 10.5120/ijca2019919228.

[15] Sutisnawati, Y. and Reski, M. (2019) "Looking for Transaction Data Pattern Using Apriori Algorithm with Association Rule Method", *IOP Conference Series: Materials Science and Engineering*, 662(2). doi: 10.1088/1757-899X/662/2/022078.

[16] Ariestya, W. W., Supriyatin, W. and Astuti, I. (2019) "Marketing Strategy for the Determination of Staple Consumer Products Using Fp-Growth and Apriori Algorithm", *Jurnal Ilmiah Ekonomi Bisnis*, 24(3), pp. 225–235. doi: 10.35760/eb.2019.v24i3.2229.

[17] Anggraeni, S., Iha, M. A., Erawati, W., & Khairunnas, S. (2019). "Analysis of Sales by Using Apriori and FP Growth at PT. Panca Putra Solusindo". Riset Dan E-Jurnal Manajemen Informatika Komputer, 3(2), 41–46.

[18] Karahan Adalı, G. and Balaban, M. E. (2019) "A Dynamic Application of Market Basket Analysis with R and Shiny in The Electric Materials Sector", *Bilişim Teknolojileri Dergisi*, (April), pp. 93–102. doi: 10.17671/gazibtd.448245.

[19] Rizqi, Z. U. (2019) "'Implementation of association rule-market basket analysis in determining product bundling strategy:" Case study of retail businesses in Indonesia', *Proceedings of the International Conference on Industrial Engineering and Operations Management*, pp. 153–160.

[20] Halim, K. K., Halim, S. and Felecia (2019) "Business intelligence for designing restaurant marketing strategy: A case study", *Procedia Computer Science*. Elsevier B.V., 161, pp. 615–622. doi: 10.1016/j.procs.2019.11.164.

[21] Wijana, H. and Finandhita, A. (2019) "APPLICATION OF DATA MINING USING ASSOCIATION RULE" Teknik Informatika - Universitas Komputer Indonesia'.

[22] Kurniawati, D. and Utami Dewi Widianti (2019) "APPLICATION OF CUSTOMER RELATIONSHIP MANAGEMENT IN SALES STRATEGY WITH THE ASSOCIATION RULES METHOD".

[23] Bilqisth, S. C. and Mustofa, K. (2020) "Determination of Temporal Association Rules Pattern Using Apriori Algorithm", *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(2), p. 159. doi: 10.22146/ijccs.51747.

[24] M.Kavitha and Subbaiah, D. S. (2020) "Association Rule Mining using Apriori Algorithm for Extracting Product Sales Patterns in Groceries", *International Journal of Engineering Research and Technology (IJERT)*, 8(3), pp. 5–8. Available at: www.ijert.org.

[25] NURZANI, Z. and TANIA, K. D. (2020) "Analysis of Transactions 212 Mart Kuto Palembang to Find Frequent Patterns Among Itemset Using Association Rule Mining", 172(Siconian 2019), pp. 325–332. doi: 10.2991/aisr.k.200424.049.

[26] Liansyah, O. and Destiana, H. (2020) "The Use of Apriori Algorithm in the Formation of Association Rule at Lotteria Cibubur", *SinkrOn*, 4(2), p. 76. doi: 10.33395/sinkron.v4i2.10526.

[27] Alfiqra and Khasanah, A. U. (2020) "Implementation of Market Basket Analysis based on Overall Variability of Association Rule (OCVR) on Product Marketing Strategy", *IOP Conference Series: Materials Science and Engineering*, 722(1). doi: 10.1088/1757-899X/722/1/012068.

[28] W. I. Yudhistyra, E. M. Risal, I. Raungratanaamporn, and V. Ratanavaraha, "Using Big Data Analytics for Decision Making: Analyzing Customer Behavior using Association Rule Mining in a Gold, Silver, and Precious Metal Trading Company in Indonesia," *Int. J. Data Sci.*, vol. 1, no. 2, pp. 57–71, 2020.

[29] Efrat, A. R., Gernowo, R. and Farikhin (2020) "Consumer purchase patterns based on market basket analysis using apriori algorithms", *Journal of Physics: Conference Series*, 1524(1). doi: 10.1088/1742-6596/1524/1/012109.

[30] Rachmatika, R. and Harefa, K. (2020) "Analysis of Determination of Strategy Promotion using Apriori Algorithm", *Journal of Physics: Conference Series*, 1477(2). doi: 10.1088/1742-6596/1477/2/022032.

[31] Philip Doganis, Alex Alexandridis, P. P. and Sarimveis, H. (2005) 'Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing', *Journal of Food Engineering*, 75(2), pp. 196–204. doi: 10.1016/j.jfoodeng.2005.03.056.

[32] Aburto, L. and Weber, R. (2005) "Improved supply chain management based on hybrid demand forecasts". doi: 10.1016/j.asoc.2005.06.001.

[33] Das, P. and Chaudhury, S. (2006) "Prediction of retail sales of footwear using feedforward and recurrent neural networks", *Neural Computing and Applications*, 16(4–5), pp. 491–502. doi: 10.1007/s00521-006-0077-3.

[34] Ã, K. A., Choi, T. and Yu, Y. (2008) "Fashion retail forecasting by evolutionary neural networks", 114, pp. 615–630. doi: 10.1016/j.ijpe.2007.06.013.

[35 Chen CY, Lee WI, Kuo HM, Chen CW, Chen KH (2010d). "The study of a forecasting sales model for fresh food'. Expert Syst. Appl., 37: 7696-7702.

[36] G. Anand, A. H. Kazmi, P. Malhotra, L. Vig, P. Agarwal, and G. Shroff. "Deep temporal features to predict repeat buyers". In NIPS 2015 Workshop: Machine Learning for eCommerce, 2015.

[37] Kaneko, Y. and Yada, K. (2016) "A Deep Learning Approach for the Prediction of Retail Store Sales", *IEEE International Conference on Data Mining Workshops, ICDMW*, 0, pp. 531–537. doi: 10.1109/ICDMW.2016.0082.

[38] Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. 2016. "Your Cart tells You: Inferring Demographic Attributes from Purchase Data". In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining. WSDM, pages 173–182.

[39] Pengfei Wang., Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. (2016) "Multi-task Representation Learning for Demographic Prediction", pp. 88–99. doi: 10.1007/978-3-319-30671-1.

[40] Salehinejad, H. and Rahnamayan, S. (2016) "Customer shopping pattern prediction: A recurrent neural network approach", 2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016. doi: 10.1109/SSCI.2016.7849921.

[41] Massaro, A., Maritati, V. and Galiano, A. (2018) "Data Mining Model Performance of Sales Predictive Algorithms Based on Rapidminer Workflows", *International Journal of Computer Science and Information Technology*, 10(3), pp. 39–56. doi: 10.5121/ijcsit.2018.10303.

[42] Jasmien Lismont, Sudha Ram, Jan Vanthienen, Wilfried Lemahieu, and Bart Baesens. "Predicting interpurchase time in a retail environment using customer-product networks: An empirical study and evaluation". Expert systems with applications, 104:22–32, 2018.

[43] Matobobo, C. and Osunmakinde, I. O. (2014). "A Comparative Market Basket Analysis For Centralized Retail Enterprises Using The ANN and AR Models".C. *CONTEMPORARY MANAGEMENT IN Proceedings of the 26 th Annual SAIMS Conference*.

[44] Zekić-Sušac, M. and Has, A. (2015) "Data Mining as Support to Knowledge Management in Marketing", *Business Systems Research Journal*, 6(2), pp. 18–30. doi: 10.1515/bsrj-2015-0008.

[45] Matobobo Courage (2016). ."Development Of An Intelligent Analytics-Baced Model For Product Sales Optimisation In Retail Enterprises" 'at the UNIVERSITY OF SOUTH AFRICA SUPERVISOR: PROFESSOR ISAAC O. OSUNMAKINDE 03 July 2016', (July).

[46] Beheshtian-Ardakani, A., Fathian, M. and Gholamian, M. (2017) "A novel model for product bundling and direct marketing in e-commerce based on market segmentation", *Decision Science Letters*, 7(1), pp. 39–54. i:

10.5267/j.dsl.2017.4.005.
[47] Zeynep Hilal Kilimci, A. Okay Akyuz, Mitat Uysal, Selim Akyokus, M. Ozan Uysal, Berna Atak Bulbul, Mehmet Ali Ekmis. 2019. "An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain". Complexity 2019, 1-15.
[48] M., W. and , Ahmed H, H. K. (2004) "Combined Algorithm for Data Mining using Association rules", *Currents*, 10(7), pp. 1–14.
[49] Sivri, E. Ş. and Cem, M. (2014) "Association Rule Mining to Extract Knowledge from Online Store Transactions of a Turkish Retail Company: A Case Study", *International Journal of Electronics, Mechanical and Mechatronics Engineering (IJEMME)*, 4(4), pp. 861–865.
[50] Fang, Y., Xiao, X., Wang, X. and Lan, H. (2018). "Customized bundle recommendation by association rules of product categories for online supermarkets", 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), pp. 472–475.
[51] Lok, J. H., Wang, X. and Xu, S. (2020) "Grocery Maze Navigating through grocery maze by using Interactive Network Visualization".
[52] Kim, E., Kim, W. and Lee, Y. (2002) "Combination of multiple classifiers for the customer ' s purchase behavior prediction", 34, pp. 167–175.
[53] Bhargav, A., Mathur, R. P. and Bhargav, M. (2014) "Market basket analysis using artificial neural network", *2014 International Conference for Convergence of Technology, I2CT 2014*, pp. 1–6. doi: 10.1109/I2CT.2014.7092091.
[54] Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, Shengxian Wan, X. C. (2015) "Learning Hierarchical Representation Model for Next Basket Recommendation", *Pediatria Polska*, 56(5), pp. 487–495.
[55] Shengxian Wan, Yanyan Lan, Pengfei Wang, Jiafeng Guo, Jun Xu, and Xueqi Cheng. 2015. "Next basket recommendation with neural networks". In Proceedings of the 9th ACM Conference on Recommender Systems. ACM, 1–2.
[56] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan." A dynamic recurrent model for next basket recommendation". In SIGIR '16, pages 729–732, 2016.
[57] Gangurde, R., Kumar, B. and Gore, S. D. (2017) "Optimized Predictive Model using Artificial Neural Network for Market Basket Analysis". doi: 10.090592/IJCSC.2017.014.
[58] Li, Y.-R., Hwang, T.-K. and Chang, S.-C. (2018) "Dynamic Inference of Personal Preference for Next-to-Purchase Items by Using Online Shopping Data", pp. 23–34. doi: 10.5121/csit.2018.81403.
[59] S. Wang, L. Hu, L. Cao, X. Huang, D. Lian, and W. Liu, "Attention-based transactional context embedding for next-item recommendation," in AAAI, 2018.
[60] Haochao Ying, Fuzhen Zhuang, Fuzheng Zhang, Yanchi Liu, Guandong Xu, Xing Xie, Hui Xiong, and Jian Wu. 2018. "Sequential Recommender System based on Hierarchical Attention Networks". In IJCAI.
[61] Qiaolin Xia, Peng Jiang, Fei Sun, Yi Zhang, Xiaobo Wang, and Zhifang Sui. 2018. "Modeling Consumer Buying Decision for Recommendation Based on Multi-Task Deep Learning". In CIKM. ACM, 1703–1706.
[62] Ting Bai, Jian-Yun Nie, Wayne Xin Zhao, Yutao Zhu, Pan Du, and Ji-Rong Wen. 2018. "An attribute-aware neural attentive model for next basket recommendation". In SIGIR'18. ACM, 1201–1204.
[63] C. O. Sakar, S. O. Polat, M. Katircioglu, and Y. Kastro, "Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and lstm recurrent neural networks," Neural Computing and Applications, vol. 31, no. 10, pp. 6893–6908, 2019.
[64] Chen, P. and Li, J. (2019) "A recurrent model with self-attention for product repurchase recommendation", *ACM International Conference Proceeding Series*, (10), pp. 199–203. doi: 10.1145/3325730.3325763.
[65] Le, D. T., Lauw, H. W. and Fang, Y. (2019) "Correlation-sensitive next-basket recommendation", *IJCAI International Joint Conference on Artificial Intelligence*, 2019-Augus, pp. 2808–2814. doi: 10.24963/ijcai.2019/389.
[66] Tingting Zhang, Pengpeng Zhao, Yanchi Liu, Victor Sheng, Jiajie Xu, Deqing Wang, Guanfeng Liu, and Xiaofang Zhou. 2019. "Feature-level Deeper Self Attention Network for Sequential Recommendation". In IJCAI. 4320–4326.
[67] Sreenivasa, B. R. and Nirmala, C. R. (2019) "Hybrid location-centric e-Commerce recommendation model using dynamic behavioral traits of customer", *Iran Journal of Computer Science*. Springer International Publishing, 2(3), pp. 179–188. doi: 10.1007/s42044-019-00040-3.
[68] Cirqueira, D., Helfert, M. and Bezbradica, M. (2019) "Towards Preprocessing Guidelines for Neural Network Embedding of Customer Behavior in Digital Retail", *ACM International Conference Proceeding Series*. doi: 10.1145/3386164.3389092.
[69] Ting Bai, Pan Du, Wayne Xin Zhao, Ji-Rong Wen, and Jian-Yun Nie. 2019. "A Long-Short Demands-Aware Model for Next-Item Recommendation". arXiv preprint arXiv:1903.00066 (2019).
[70] Z. Liu, X. Li, Z. Fan, S. Guo, K. Achan, and P. S. Yu, "Basket recommendation with multi-intent translation graph neural network," arXiv preprint arXiv:2010.11419, 2020.
[71] Z. Liu, M. Wan, S. Guo, K. Achan, and P. S. Yu, "Basconv: Aggregating heterogeneous interactions for basket recommendation with graph convolutional neural network," in Proceedings of the 2020 SIAM International Conference on Data Mining. SIAM, 2020, pp. 64–72.
[72] Lee, H.I.; Choi, I.Y.; Moon, H.S.; Kim, J.K. "A Multi-Period Product Recommender System in Online Food Market based on Recurrent Neural Networks". Sustainability 2020, 12, 969.
[73] Changchien, S. and Lu, T. C. (2001) "Mining association rules procedure to support on-line recommendation by customers and products fragmentation", *Expert Systems with Applications*, 20(4), pp. 325–335. doi: 10.1016/S0957-4174(01)00017-3.
[74] Shim, B., Choi, K. and Suh, Y. (2012) "CRM strategies for a small-sized online shopping mall based on association rules and sequential patterns", *Expert Systems with Applications*. Elsevier Ltd, 39(9), pp. 7736–7742. doi: 10.1016/j.eswa.2012.01.080.
[75] Mansur, A. and Kuncoro, T. (2012) "Product Inventory Predictions at Small Medium Enterprise Using Market Basket Analysis Approach-Neural Networks", *Procedia Economics and Finance*. Elsevier B.V., 4(Icsmed), pp. 312–320. doi: 10.1016/s2212-5671(12)00346-2.
[76] Itkar, S. and Uday, K. (2014) "Efficient Frequent Pattern Mining Using Auto-Associative Memory Neural Network", *British Journal of Applied Science & Technology*, 4(22), pp. 3160–3178. doi: 10.9734/bjast/2014/10707.
[77] Ghadekar and Dombe (2019) "Image base Product Recommendations Using Market Basket Analysi"'.