Business Intelligence Integration and Architecture using Artificial Intelligence

Tanmayee Tushar Parbat[1], Honey Jain[1], Rohan Benhal[2], Dr. Lalit Kulkarni[3]

[1]B.E IT, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

[2]BBA IT, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

[3]Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

ABSTRACT

During the last decade of the twentieth century and among the first one and half decades of the third millennium-up to the best knowledge of the author of this thesis, the issue of Business Intelligence and its involvement in business environment and in Electronic Commerce environment in particular for better achievement, has 'often' been dealt with as several individual unidirectional problems, or quite seldom as bidirectional problem. This thesis presents however a research project that has differently handled the said issue as a 'single' multidirectional problem. The general trend to utilize the concept of Business Intelligence was through considering it as a tool or a means that is actively reliable in decision making concerning business organizations with respect to their partial activates besides their strategic and executive planning.

Keywords : - Recommendation System, Machine Learning, Deep Learning

I. INTRODUCTION

In the decades resulted the emergence of the World-Wide Web, using the Internet has passed away from its originally envisaged scope and altered from a distributed system to interchange knowledge towards a new application environment [1, 2]. This lead to a dramatic and fast growth of web, more recently activated by firms providing web based services and products and thus starts the real era of Electronic Commerce applications. Commerce denotes merchandise buying and selling. The processes of dealing with selling and buying, transferring, or exchanging products, services, and or information via computers including the Internet is referred to as Electronic Commerce (EC). EC presents an assortment of substantial benefits comprising: sales increase, costs reduction, better services and support from suppliers, achieving higher customer satisfaction, and ability to participate in international markets. The majority of commercial organizations have followed the trend of adopting EC at the begging of the nineties. This was due to the emergence of the Internet and the facilities / benefits offered by its associated cyber services [3, 4]. Those firms were in fact hopping to achieve these benefits, as they have mostly been influenced in their

decision to adopt EC by several factors. Amongst these are:

(i) The desire to strengthen their relationship with business partners,
(ii) The need to improve business efficiency,
(iii) Pressure from competitors who use EC,
(iv) Pressure from customers and suppliers,
(v) To assist firm managers in evaluating the rate of commercial growth of their companies.

Commercial products and pertinent services can be provided to customers by an organization through adopting EC. This can be done anywhere worldwide by costumers accessing international markets easily through merely simple mouse clicks. Nonetheless, EC has faced also some problems. The main drawbacks yielding out of the adoption of EC may be embodied in: extreme reliance on technology, maintenance costs, on-going costs, security costs, and lack of technical support [5].

Businesses understand in this extremely competitive environment, quickly paced and continuously changing business conditions and Life-Cycle, a key competitive quantity is how fast they react and adjust to change. BI provides best way in order to use information combined to quickly and constantly respond to changes. Fig. 1 is showing the basic understand of BI functions.
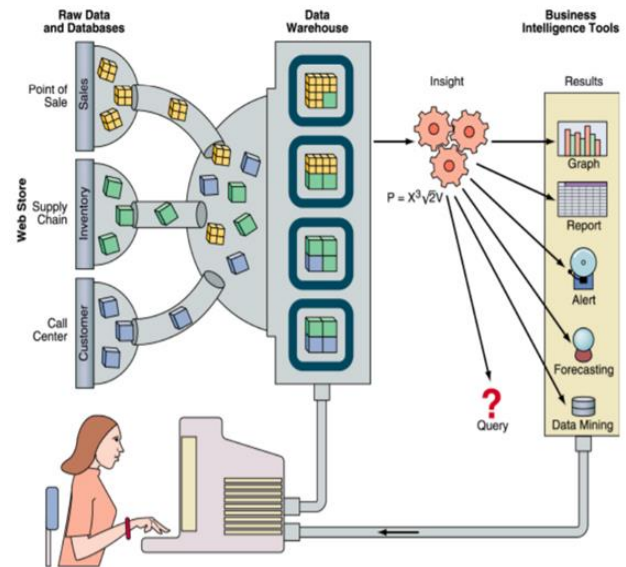

Fig. 1: Basic Understanding of BI

II. MACHINE LEARNING

Neural networks are the data handling frameworks, which are built and executed to show the human cerebrum. The main object of the neural network analysis is to develop a process of computing device for modeling the brain to execute various process of computing tasks at a faster rate than the traditional systems. Artificial Neural Networks execute various tasks such pattern matching and classification, optimization function and data clustering. These errands are exceptionally troublesome for conventional PCs, which are quicker in algorithmic procedure of registering undertakings and exact number juggling tasks. ANNs gangs substantial number of exceedingly interconnected preparing components called hubs or unit or neuron, which more often than not work in parallel and are arranged in standard models [6, 7]. Every neuron is associated with the other by an association interface. Every association interface is related with weights, which contain data about the info flag. This data is required by neuron net to take care of a specific issue. ANN, s aggregate conduct is portrayed by their capacity to learn, review and sum up preparing examples or information like that of human mind [8].

Neural network, with their astounding capacity to get importance from muddled or uncertain information, could be utilized to remove designs and distinguish patterns that are too perplexing to be in any way saw by either people or other PC method. Favorable circumstances of neural system incorporate their high resistance to clamors information and in addition their capacity to order designs on which they have not been prepared. Other advantages of working with ANN include:

- **Adaptive Learning:** An ANN is enriched by the capacity to recognize how to do undertakings in light of the information given for preparing or starting background.
- **Self-Organization:** An ANN can make its own affiliation or depiction of the information it gets in the midst of learning time.
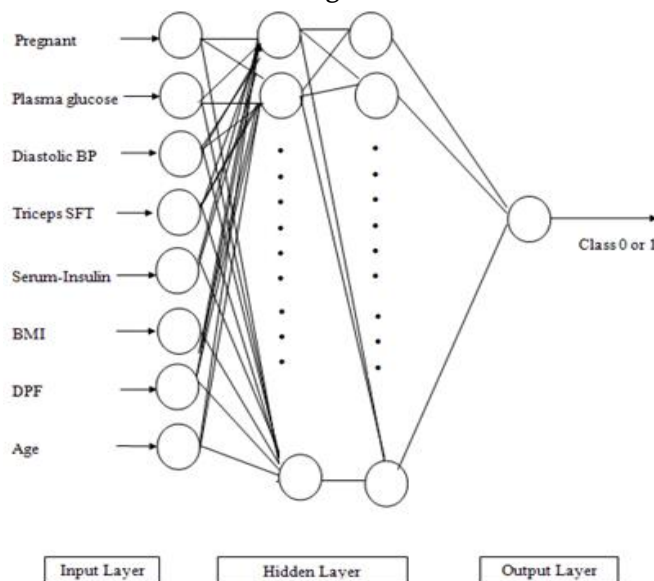

Fig. 2: Feed forward NN model for diabetes diagnosis

- **Real time operation:** Extraordinary equipment gadgets are being outlined and produced to exploit this ability of ANNs.
- **Fault tolerance:** Fractional obliteration of a neural system prompts the comparing debasement of execution. Be that as it may, some system capacities might be retrained even after real system harm.

III. PROPOSED METHODOLOGY

The key estimate for time series analysis is to decompose the original time series into several independent parts. Typically, business time series are divided into the following four parts:

- Trend – overall direction of the series, i.e. upwards, downwards etc.
- Seasonality – monthly or quarterly patterns
- Cycle – long-term business cycles
- Irregular remainder – random noise left after extraction of all the elements

Interference of these ingredients produces the last series and the most popular model in time series analysis is an autoregressive integrated moving average (ARIMA) model. It is a linear model used for analyzing the time series data given as input and predicts the future points in the series. While predicting the future data, three terms are used to refine the ARIMA model namely:

- Number of autoregressive terms (p)
- Number of non-seasonal differences (d)
- Number of lagged forecast errors (q).
- Granting to the time series modeling method of the ARIMA (p, d and q) model, the number of autoregressive term data is drawn first.
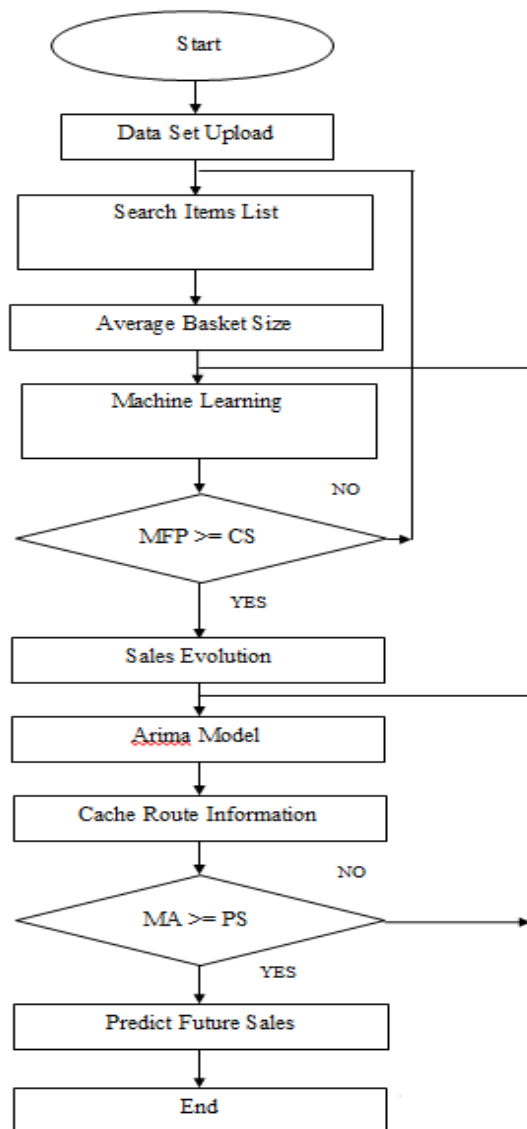
Fig. 3: Flow Chart of Proposed Methodology

ARIMA model has the suspicion that the time arrangement is stationary. On these lines, information change is done to create a stationary time arrangement; hence the mean and autocorrelation structure needs to be consistent after several iterations. This model designation used to gain the resources of inputs called as a class of model identification done by getting the inputs. Year of publication and total number of publications for the concept are considered as input respectively. Initially, the appropriate ARIMA model has to be placed for the particular data sets and the parameters should have smallest possible values such that it can break down the information properly and

forecast accordingly. When comparing two or more models, the one with the lowest AIC is generally considered closer to real data. Two widely used measures can be used to assess the model namely

- Alkaline Information Criteria (AIC)
- Bayesian Information Criterion (BIC) The AIC does not penalize model complexity as heavily as the BIC does.

The authors showed that AIC and BIC can be derived in the same framework and using AIC for model selection is theoretically proved more effective than using BIC for selecting a model. According to Box-Jenkins method, in ARIMA (p, d, q) the value of p and q hould be 2 or less or total number of parameters should be less than 3 (Ly Pham, 2012) Hence, for checking AIC of the good example it needs to be checked for p and q values 2 or less. Depending on AIC, model ARIMA (1, 0, and 2) is selected, since it has the least AIC value and experimental results are done for a selection of various learning objects.

Algorithm for Proposed Methodology:-
Step 1: Start
Step 2: Data set upload
Step 3: Search Item List
Step 4: Average Basket Size
Step 5: Machine Learning
    (5.1) Hierarchical routing protocol
    (5.2) Calculate Product
    (5.3) arranged themselves
    (5.4) Control unbalanced product
Step 6: Calculate moving frequency product,
Customer size (CS)
    (6.1) if
    (6.2) (MFP >= CS)
    (6.3) end if;
Step 7: Sales Evolution
Step 8: Cache route information
Step 9: Moving average (MA), Product size (PS)

(9.1) if

(9.2) (MA >= PS)

(9.3) end if;

Step 10: Predict Future Sales

Step 10: End

IV. SIMULATION RESULT

Python is a high level general programming language and is very widely used in all types of disciplines such as general programming, web development, software development, data analysis, machine learning etc. Python is used for this project because it is very flexible and easy to use and also documentation and community support is very large.

Question 1 - How Many Items Are There In Customer Orders?

```
no_of_Items = item['product_id'].nunique()
print("Number of items in customer orders: "+str(no_of_Items))


    Number of items in customer orders: 32951
```

Question 2 - What Is The Average Basket?

```
No_of_products = len(item.product_id)
Number_of_invoice = len(item.order_id.unique())
average_basket = No_of_products/Number_of_invoice
print("Total no unit sold = "+str(No_of_products))
print("Number of invoices = "+str(Number_of_invoice))
print("Aerage Basket Size is "+str(average_basket))


    Total no unit sold = 112650
    Number of invoices = 98666
    Aerage Basket Size is 1.1417306873695092
```

Question 3 - What Are The Most Frequent Product Categories?

```
There are total 71 categories of products.
Top 10 Frequent categories are :-
product_category_name
bed_bath_table              11115
health_beauty                9670
sports_leisure               8641
furniture_decor              8334
computers_accessories        7827
housewares                   6964
watches_gifts                5991
telephony                    4545
garden_tools                 4347
auto                         4235
```

V. CONCLUSION

Among the different phases of the work presented in this thesis, several aspects have been articulated. Amongst these are:

1- Most firms nowadays are keen to operate effective business processes as well as to permit for changing business conditions at an increasing rate.

2- Operational Business Intelligence is complicatedly conjugated to transactional business processes. The limit of this integration depends on the implementation level.

3- Business Process Management is commonly concentrated on enhancing the automation and business processes modeling, aiming to enable for fast and cost-effective process execution.

4- The interest of customers in managing process execution becomes higher, as more processes become more automated.

5- This research has taken both managerial and technical factors under consideration. It has also encompassed different modes and possibilities for the growth of Business Intelligence in Indian organizations.

6- Business Intelligence is applied by selected Dot Com companies for factors analysis involving data affecting

their business inside and outside the organization. This will help them in decision making.

VI. REFERENCES

[1] Betru, B. T., Onana, C. A., & Batchakui, B. (2017). Deep learning methods on recommender system: A survey of state-of-the-art. International Journal of Computer Applications, 162(10), 17-22.

[2] Doha, A., Elnahla, N., & McShane, L. (2019). Social commerce as social networking. Journal of Retailing and Consumer Services, 47, 307-321.

[3] Osadchiy, T., Poliakov, I., Olivier, P., Rowland, M., & Foster, E. (2019). Recommender system based on pairwise association rules. Expert Systems with Applications, 115, 535-542.

[4] Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert systems with applications, 36(2), 2592-2602.

[5] Rajarajeshwari, K., & Radhamani, G. (2019). A Naïve Bayes Model using Semi-Supervised Parameters for Enhancing the Performance of Text Analytics. International Journal of Advanced Networking and Applications, 10(6), 4083-4089.

[6] Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications, 97, 205-227.

[7] Li, W., Bai, Q., Zhang, M., & Nguyen, T. D. (2018). Automated influence maintenance in social networks: an agent-based approach. IEEE Transactions on Knowledge and Data Engineering, 31(10), 1884-1897.

[8] Khan, Z. A., Zubair, S., Imran, K., Ahmad, R., Butt, S. A., & Chaudhary, N. I. (2019). A New Users Rating-Trend Based Collaborative Denoising Auto-Encoder for Top-N Recommender Systems. IEEE Access, 7, 141287-141310.

[9] Zhang, S., & Zhong, H. (2019). Mining users trust from e- commerce reviews based on sentiment similarity analysis. IEEE Access, 7, 13523-13535.

[10] Jia, S. S., & Wu, B. (2019). User Generated Information on Mobile Channels With More Concise Reviews and More Extreme Ratings. IEEE Access, 7, 83495-83503.

[11] Wang, S., Tang, X., Zhang, Y., & Chen, J. (2019). Auditable Protocols for Fair Payment and Physical Asset Delivery Based on Smart Contracts. IEEE Access, 7, 109439-109453.

[12] Scholta, H., Mertens, W., Kowalkiewicz, M., & Becker, J. (2019). From one-stop shop to no-stop shop: An e- government stage model. Government Information Quarterly, 36(1), 11-26.

[13] Helo, P., & Shamsuzzoha, A. H. M. (2020). Real-time supply chain—A blockchain architecture for project deliveries. Robotics and Computer-Integrated Manufacturing, 63, 101909.

[14] Da Costa, A. F., Manzato, M. G., & Campello, R. J. (2019). Boosting collaborative filtering with an ensemble of co- trained recommenders. Expert Systems with Applications, 115, 427-441.

[15] AbdulHussien, A. A. (2017). Comparison of machine learning algorithms to classify web pages. International Journal of Advanced Computer Science and Applications (ijacsa), 8(11).