iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** X **Month of publication:** October 2021

DOI: https://doi.org/10.22214/ijraset.2021.38550

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue X Oct 2021- Available at www.ijraset.com

A Systematic Review of the Literature on Customer Analytics to Predict Dynamic Purchase behavior of Consumers through Credit Cards

Suvigya Jain[1], Tushar Kaushik[2]

Abstract: The Credit card industry is flooded with customer data and also their diversified purchasing behavior. These multifariousness analysis helps different industry to maximize revenue through data and to identify right cluster of audience based on the products and services they offer. To analyse this data efficiently and erect efficient co-relation and streamline process execution, expertise in customer segmentation specifically through customer analytics, is required. This domain of analytics uses a wide variety of concepts related to data mining, Passive and active data collection, Media planning, Regression analysis, Customer based corporate-valuation, Market-structure, Probability Models, Optimization, visualization and Implementation of model spreadsheet for predictions on data set. Moreover, customer analytics is the subset of a larger range of business analytics techniques used for better statistical and predictive analysis to transform and make smarter, data-driven business decisions.
Keywords: Regression analysis, Market-structure, Optimization, Business analytics, Model spreadsheet.

I. INTRODUCTION

Previously, methodologies for evaluating enormous data sets and other systematic reviews were focused on outcomes that might be used to educate the organization for segmentation of customers. These segmentation outcomes are often applicable to variety of organizational communities, to build the ability to foresee and make predictions for the customer behaviour and their purchases. Dynamic purchase behaviour helps to make acquisition of purchases made by the customer's in a dynamic sequence and this study of behaviour purchase depicts how people make decisions of where to spend their limited resources. This spending's can be systematically and statistically analyse by transactions done by users, but to make a accurate analysis a new trail of data should be inspected which is left behind by internet users and computer systems where transactions is executed, processed or accomplished during their online activities which is referred as Data Exhaust. These will result in accurate perspective analysis for a customer that will result in better decisions about an individual and will give a way to attract and keep consumers engaged, and will result in discovery of high-value clientele satisfaction. The more a customer's purchase habits and lifestyle preferences are understood, the more accurate and precise prediction behaviours emerge, and the better the customer journey gets.

II. DISCUSSION

Credit card transactions details lucidly delineates shopper spending interests, and act as a fundamental for the research. This investigation particular illustrate aspects of customer and provide insights on how customer activities are performed and interplay with the consumer interests. These different experiences of consumer and dynamics associated with it helps to derive cluster of customers based on activities of customer and their narratives which in conjunction with customer journey mapping produce detailed data analysis and insights will help to design active patterns for enhancing customer experiences.

Actionable decisions from the processed information helps in forming up links between the right set of consumer and product generation as per the need by the organizations. These is achieved by assessing active and passive data sets dynamicity and making decisions from it and by further examining those decisions to interpret and compare for formulating better alternatives. These alternatives hypothesis can be gained by data collection survey techniques and segmenting the consumers based on the preference, value, demographics, reactions, responses, sentiments and interests and further forming cluster and targeting cluster by the satisfying necessity. The satisfying necessity of cluster can be determined by Net promoter score and is calculated by score of recommendation of the product/service by members of cluster to another members of different groups, higher the recommendation score, higher the net promoter score of the product/service. Historical data Transactions of credit cards can help in statistically understanding the spending behaviour for a consumer and will further relate in making prediction for future purchases, this includes cleansing and combining of data for analytical purpose. Cleansing focuses on finding the missing data within data sets and removing outliers from it so that completion and correlation within the primary data source is accurate for building a model for future preference.

Then, these models are utilized by analytical professionals to perform tasks which otherwise can consume significant time and expertise. The underlying data-driven model prepared can be combined with intelligence to learn neural behaviour statistics of a cluster and make intelligent predictions to make decisions.

Intelligent Predictive Decision-Model learning may be approached in three ways:

1) Supervised learning
2) Unsupervised learning
3) Reinforcement learning

Training data which is well labelled and has defined output for a input and is used in predictive models for making decisions has access to wide range of statistical methods. These model make use of training sets to predict and classify data effectively which train the algorithm to learn over time and make model result orientated for better predictions. These learning technique is divided into two types-" Classification and Regression". In Classification data algorithm precisely forms categories of data and recognize identical values within data-sets to streamline decision in form of output for inputs and in case of Regression a coordination is established to understand correspondence between inter-dependency and Independency between variable, which helps in making projections.

The technique of unsupervised learning in decision model is to form opinions from data-sets that contains input data without labels. These opinions of same characteristics are collected within the same group and named cluster. In unsupervised learning, cluster analysis is by far the most often used method of learning. Use it to discover hidden patterns or groups in data using exploratory data analysis (EDA). (EDA) make use of initial screening and perform critical process execution to invent patterns and spot anomalies and verify hypothesis to check assumptions using various data visualization techniques for graphical representations and forming synopsis for detailed statistical report. Accuracy of unsupervised learning depends Hierarchical agglomerative clustering(HAC) and by measuring performance metrics.

Reinforcement learning based intelligent predictive Decision model are based on situation of trial and error to come up with a decision for a problem, predictive intelligence decision gets either winning or losing points for the actions it performs. Its aim is to maximize the total wining point. Although trials are Completely random, followed by advanced methods and superhuman abilities. Reinforcement learning is the most efficient technique to hint machine creativity by utilizing the power of search and many trials and massive computer infrastructure is required to run the algorithm. This technique can handle complex relationships and are very flexible when examining a large set of data samples.

Traditional statistical regression procedures are well-known among statisticians and the scientific community that employs them. Choosing the "appropriate model" when utilizing traditional regression modelling can be problematic as well. Traditional regression models in the age of big data have constraints that non-traditional decision model learning algorithms & techniques can overcome, but they don't provide a complete solution because the algorithms and methods must be examined in connection to the data utilized in the research.

Although decision model learning techniques are used to analyse both population models and informed customer behaviour decision-making processes, it's important to note that the information, model, or outcomes used to analyse a behaviour need to meet the highest level research quality standards, because the decision made will almost certainly have an impact on both long-term and short-term spending outcomes. Although population-based estimates are prone to error, cluster-level models should be as accurate as feasible in order to give high-quality credit card spending's behaviour.

III. APPROACHES TO ADVANCE MODEL AND ITS VALIDATION IN GENERAL

More complex models, such as neural networks, can be utilized to generate predictive models in addition to regression models. Neural networks are statistical modelling techniques that are not linear. In general, neural networks can handle a myriad of variables than the regression techniques, and they also solve some of the other constraints of regression techniques, such as statistical concerns about dimensionality. A predictive model should be checked by out-of-sample data sets to test and guarantee that the result are as accurate and feasible. Out-of-sample testing splits data into two categories: in-sample data (data used to construct the model) and out-of-sample data (data not used to develop the model) which can be further used for segregation and targeting of consumer into different domains of offering. These will help to increase efficiency percentage for a model and will result in more precision. The common types of validation metrics used are: Confusion Matrix, F1-score, Gain-Lifts charts, Kolmogorov smirnov chart, AUC-ROC, Log-loss, Gini-Coefficient, Cordant-Discordant ratio, Root mean squared value, Cross validation.

IV. ANALYTICAL SOFTWARE FOR OPTIMIZATION AND VISUALIZATION

The methods utilized in these research for Dynamic Behaviour Analysis differed greatly, and no consistency was found. One study that used decision tree analysis used Quinlan's C5.0 decision tree method, while another used an older version of the same programme, as previously described (C4.5). In other decision tree investigations, different statistical methods were applied. For visualizing the data and creating interactive dashboards, Tableau, Microsoft Power BI, Looker, Zoho Analytics, and IBM Cognos Analytics were used. As a result, such insight may be simply derived from the data. For optimization, ADMB, ASCEND, CUTER, GNU Octave, and Scilab were employed.

V. WEAKNESSES AND STRENGTHS TAKEN INTO CONSIDERATION FOR ANALYSIS

Numerous articles assessed the relative merits and demerits of a Predictive Behavioural decisions model learning methods employed. Decision model learning techniques have been lauded for their simplicity and low complexity. The use of cluster learning methods to large data sets was both successful and efficient. It was noted that variables that were significant only for credit card transaction to predict behaviour spending were included in this study, even if they would not be significant at the population level using traditional regression analysis model building. According to one publication, decision model learning's effectiveness is highly dependent on the model selection technique & parameter optimization, and therefore that behavioural predictive model learning alone would not result in better predictions unless these steps are done properly.

Behaviour model learning methodologies, even when properly developed, may have limits that should be considered in future study using these techniques. Model over-fitting and an excessive amount of data were recognized as flaws in the qualifying papers. Furthermore, limitations imposed by the data sources used to make predictions, such as the lack of all necessary variables and partial data, may impede the performance and growth of these models.

VI. COMMENT

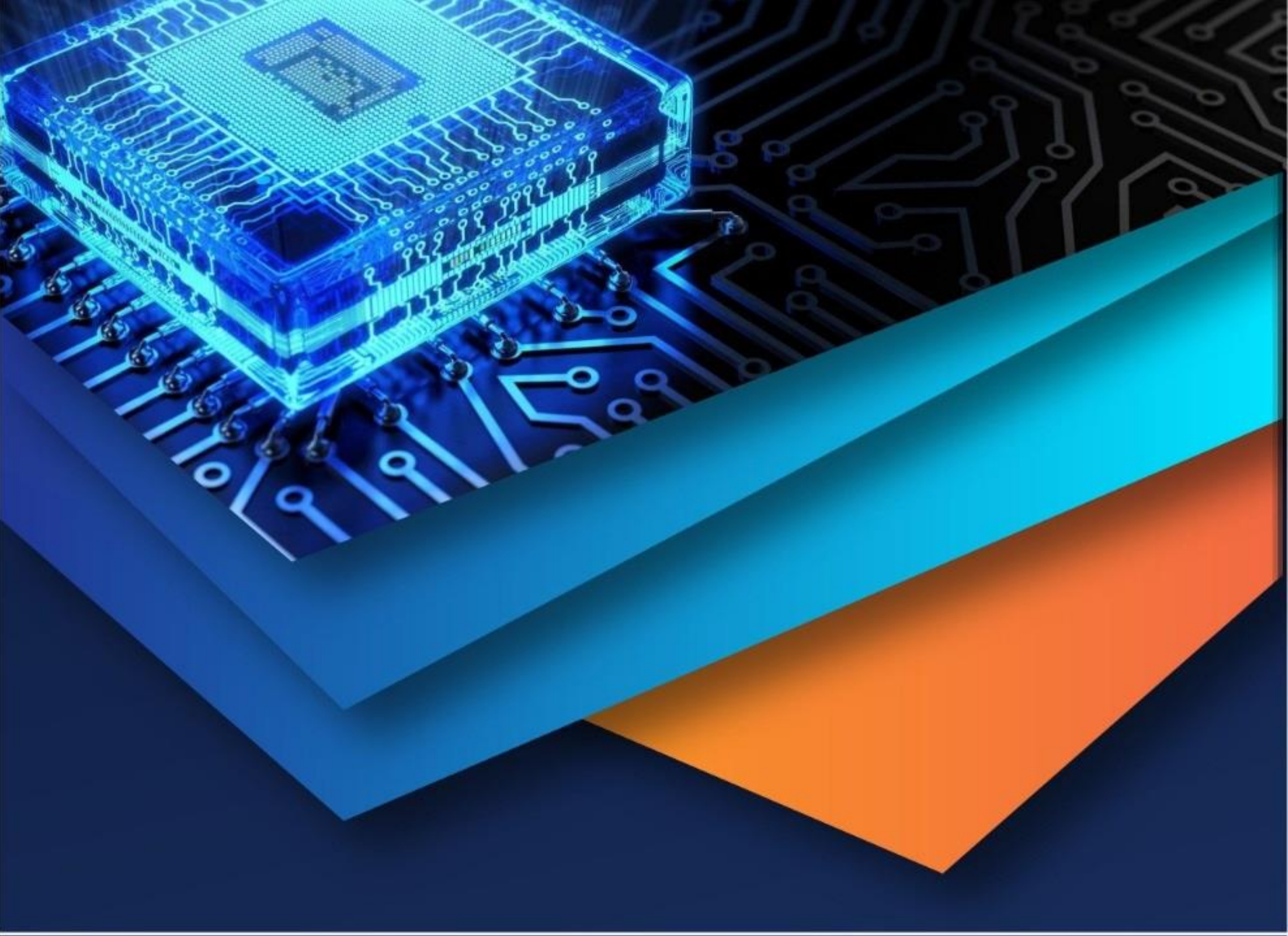
As a result of the application of Predictive and Customer Analytics techniques for decision making for individuals is much more frequent than compared with observational studies, the aim of this findings was to conduct a thorough assessment and review of the situation regarding dynamic analytical approaches sources of secondary methods, data, and methodologies that might be used to aid decision-making based on preference of cluster. As a result, the explanation of this research does not really apply equally to any and all situations. Numerous dynamic models depicts the drawbacks of utilizing population-based samples to forecast decisions. A population summary statistic, to be more precise, has nothing to do with any particular in that cluster. Population predictions are a single point on a potentially broad axis, and each individual consumer can fall anywhere along that axis and have a value that differs dramatically from the average determined value. A homogeneous analytical method was used in the majority of the studies reviewed. Using a single approximation technique has historically been recognized to introduce variability. To circumvent this drawback, multiple procedures and versions of the produced models are required. This, combined with the recent advancement of more complicated analytics, has resulted in the establishment of proposed rules for selecting and developing machine learning algorithms. Under certain cases, a linear model may be able to fit the data and provide an adequate answer; in these cases, additional techniques such as cognition approach may be used to increase the model's accuracy. Effective framework is required to ensure model accuracy; yet, it must have been rarely done in the articles included in study. This could have been due to a variety of issues, such as a lack of appropriate data-sets or a failure to understand critical need of validation and attention. Because the use of machine learning in model building grows, independent validation assessment of models before they can be used in Analyzing behavioural spending. The generalization of modelling theories cannot be evaluated without such information.

VII. CONCLUSIONS

This research uncovered a wide range of methodologies, procedures, statistical software, and validation measures for dynamic behaviour predictions using decision learning approaches to train models using secondary resources. When constructing learning models for analysis, it is necessary, according to some resources, to incorporate a wide range of modelling methodologies. In order to properly enable share proof decision making by comprehending dynamic spending through credit cards, these models must also adhere to high research rules. Models must be evaluated against well-defined selection criteria and then validated both internally and externally before being used to drive future purchases. Just a few researches have reported the level of evidence required to analyze transactions and provide in making future decisions. Also, there are various analytical software which are further used for visualization and optimization of data and brings accuracy and precision in making insights which helps in further analysis to train knowledge data centres to form an integrating market structure based on research spending's of users through their historical data of spending's by credit cards.

REFERENCES

[1] T. H. Davenport and J. G. Harris, “The dark side of customer analytics,” Harv. Bus. Rev., vol. 85, no. 5, pp. 37–48, 2007.
[2] N. Singh, K. Lai, M. Vejvar, and T. C. E. Cheng, “Data-driven auditing: A predictive modeling approach to fraud detection and classification,” J. Corp. Account. Financ., vol. 30, no. 3, pp. 64–82, 2019, doi: 10.1002/jcaf.22389.
[3] M. Hafiz Bakar, S. Norbaya Yahaya, and C. C. Men, “The Critical Factors Influencing Consumer Spending by Using Credit Card,” 2012.
[4] B. Rao, “Machine Learning Algorithms: A Review,” Int. J. Comput. Sci. Inf. Technol., vol. 7, no. 3, pp. 1174–1179, 2016, doi: 10.21275/ART20203995.
[5] B. Gill, B. Borden, and K. Hallgren, “A conceptual framework for data- driven decision making,” https://pdfs.semanticscholar.org/9c24/f3863c4b3c3bd02d0a1fb3d1be9a17d7204f.pdf?_ga=2.111147269.1056954673.1579876302-434277816.1579876302.
[6] K. Hjort, B. Lantz, D. Ericsson, and J. Gattorna, “Customer segmentation based on buying and returning behaviour,” Int. J. Phys. Distrib. Logist. Manag., vol. 43, no. 10, pp. 852–865, 2013, doi: 10.1108/IJPDLM-02-2013-0020.
[7] K. M. Archana Patel and P. Thakral, “The best clustering algorithms in data mining,” Int. Conf. Commun. Signal Process. ICCSP 2016, no. 2, pp. 2042–2046, 2016, doi: 10.1109/ICCSP.2016.7754534.
[8] Y. O. Sayad, H. Mousannif, and H. Al Moatassime, “Predictive modeling of wildfires: A new dataset and machine learning approach,” Fire Saf. J., vol. 104, no. January, pp. 130–146, 2019, doi: 10.1016/j.firesaf.2019.01.006.
[9] Iyer Aurobind Venkatkumar and Sanatkumar Jayantibhai Kondhol Shardaben, “Estudio comparativo de algoritmosagrupamientodemineríadedatos,”2016,[Online].Available:http://bdbib.javerianacali.edu.co:2208/document/7823946/.
[10] Y. Xie, X. Li, E. W. T. Ngai, and W. Ying, “Customer churn prediction using improved balanced random forests,” Expert Syst. Appl., vol. 36, no. 3 PART 1, pp. 5445–5449, 2009, doi: 10.1016/j.eswa.2008.06.121.
[11] J. H. Schwab, “Predictive analytics,” Spine J., vol. 20, no. 7, pp. 1152–1153, 2020, doi: 10.1016/j.spinee.2020.03.003.
[12] J. Chen, G. Kou, and Y. Peng, “The dynamic effects of online product reviews on purchase decisions,” Technol. Econ. Dev. Econ., vol. 24, no. 5, pp. 2045–2064, 2018, doi: 10.3846/tede.2018.4545.
[13] E. T. Bradlow, M. Gangwar, P. Kopalle, and S. Voleti, “The Role of Big Data and Predictive Analytics in Retailing,” J. Retail., vol. 93, no. 1, pp. 79–95, 2017, doi: 10.1016/j.jretai.2016.12.004.
[14] E. Diecidue, N. Rudi, and W. Tang, “Dynamic purchase decisions under regret: Price and availability,” Decis. Anal., vol. 9, no. 1, pp. 22–30, 2012, doi: 10.1287/deca.1110.0227.
[15] N. I. Fisher and R. E. Kordupleski, “Good and bad market research: A critical review of Net Promoter Score,” Appl. Stoch. Model. Bus. Ind., vol. 35, no. 1, pp. 138–151, 2019, doi: 10.1002/asmb.2417.
[16] R. J. E. James and R. J. Tunney, “The need for a behavioural analysis of behavioural addictions,” Clin. Psychol. Rev., vol. 52, pp. 69–76, 2017, doi: 10.1016/j.cpr.2016.11.010.
[17] N. D. Shah, E. W. Steyerberg, and D. M. Kent, “Big data and predictive analytics: Recalibrating expectations,” JAMA - J. Am. Med. Assoc., vol. 320, no. 1, pp. 27–28, 2018, doi: 10.1001/jama.2018.5602.
[18] N. Sun, J. G. Morris, J. Xu, X. Zhu, and M. Xie, “ICARE: A framework for big data-based banking customer analytics,” IBM J. Res. Dev., vol. 58, no. 5–6, pp. 1–9, 2014, doi: 10.1147/JRD.2014.2337118.
[19] V. Victor, J. J. Thoppan, R. J. Nathan, and F. F. Maria, “Factors influencing consumer behavior and prospective purchase decisions in a dynamic pricing environment-an exploratory factor analysis approach,” Soc. Sci., vol. 7, no. 9, 2018, doi: 10.3390/socsci7090153.
[20] J. Wexler, M. Pushkarna, T. Bolukbasi, M. Wattenberg, F. Viegas, and J. Wilson, “The what-if tool: Interactive probing of machine learning models,” IEEE Trans. Vis. Comput. Graph., vol. 26, no. 1, pp. 56–65, 2020, doi: 10.1109/TVCG.2019.2934619.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY