International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 1 Volume No.7 January - February – 2023
DOI:10.46647/ijetms.2023.v07i01.052 ISSN: 2581-4621

Churn Detection Using Machine Learning in the Retail Industry

Ms.Prithi Madhavan[1], Dr.K. Tamizharasi[2]
[1]Assistant Professor, PG Department of Computer Applications,Marudhar Kesari Jain College for Women, Vaniyambadi, Tirupattur District, Tamil Nādu
2 Research Supervisor, Computer Science, Periyar University Constituent College of Arts & Science, Idappadi, Salem, Tamil Nādu, India
Orchid ID : 0000 - 0001 - 8933 - 2092

Abstract
The top priority of any business is a constant need Increase sales and profitability. one of the causes of a reduction in profit occurs when an existing customer stops trading. When a customer leaves or terminates the company, potential sales or cross-selling opportunities are lost. When the customer leaves the store without any advice. It can be difficult for companies to respond and take corrective action. Ideally, companies should act proactively and identify themselves chances are you will churn before they leave. customer retention strategies have proven to be less expensive than attracting new ones client. Through data available at the POS(POS) system, extract customer transactions, you can analyze their buying behavior. In this paper Features are created through transactional data and how they are created Identified as important for predicting retail churn industry. Data provided in this document refer to local resident's supermarket. Thus, dropouts are identified and results are obtained. The results obtained are based on real scenarios. The novelty of this paper is the concept of implementing deep learning algorithms. Convolutional Neural Networks and Restricted Boltzmann Machine is the deep learning technique of choice or restricted Boltzmann machine gave the best results 83% in predicting customer attrition.
Keywords: customer attrition, deep learning, retail grocery industry.

1. INTRODUCTION

Customer retention is an important issue. It is found in various industries. Common reason why customers leave: competition is offer. Similar products at lower prices. word of mouth or negative marketing via social media. there may have been competition could have been better customer service or customers moved research [1] shows that the cost of retaining customers is high less than put a new one. Marketing costs are the reason required for attracting new customers. together for this as competition intensified, this became a key factor existing customer base will be preserved. regular customer transition gradually, rather than abruptly. this is, analyze past customer buying patterns for adoption a proactive approach to predicting churn. because all transactions inserted via POS and saved in the database. Customer needs and patterns can be grasped as data Accessible. This paper focuses on two aspects of predicting customer Churn for grocery stores industry. The first is based on the following features: passed to the model. instead of buying customers the tendency to group individuals creates these values passed to the model as a feature.

So to everyone various functions are created for customers to enable the model learn and recognize per capita patterns for this reason 2 Datasets are created to test and evaluate how data should be plotted to predict churn. The second aspect is Algorithm implementation. New in this study predict Grocery Churn Using Deep Learning industry. To our knowledge, this is the first study. implements deep learning in this industry. the strength of the use of deep learning is that it can reveal hidden patterns inside available datasets.

II. CONTEXT

A. Definition of Churn

Churn is a marketing term for customers who have switched to a competitor or who have stopped buyings from you. Churn can be defined as customers who are likely to stop doing business with the

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 1 Volume No.7 January - February – 2023
DOI:10.46647/ijetms.2023.v07i01.052 ISSN: 2581-4621

company within a certain period of time can also be defined as:customer's average shopping cart i.e. if the average amount spent over a period of time decreases It has been below the threshold for a given period of time or The average basket is formally expressed in Equation 1.Purchase Amount C (i)Customer C for week i. n = number of |P| Weeks of period P. Pinpointing the exact moment is difficult. Customers migrate to supermarket retail outlets. Customers suddenly stop shopping at your store they are partially broken. In fact, they tend to switch to rivals gradually [6]. To be more factual, this context churner is defined by the following rules: Let P1 be the next period. Observing where a customer's typical purchases are Examined. This will be the input for future churn detection model. Immediately after P1 is P2, which is the period evaluation, changes in customer buying habits you can see P2 is unknown to the predictive model. During this investigation, P2 is permanently set to 12 to reach consensus. Requests for marketing services. What do you mean the evaluation period is always 3 months. The marking method is as follows. Average 20% of the average purchase amount during the purchase period when P2 is for P1, this customer is flagged as churn. If not, He/she is called a non-modifier. who is that 20% allowable waste, called reduction factor. that's what it means Churn can be partial or total. formal definition of Here is the churn: Let α be the reduction factor and C be the customer.

From these definitions and observations we can deduce three things termination case.

1) Simple churn cases These cases have a negative slope.The noise is lower Let S be the gradient Time series values as yˆi = Sxi+c. every time i die Error ei = |yi −yˆI | | must respect ei ≤ yes n is Number of periods or length of time series. In other words, noise does not distort trends. This The type of charner Simple linear regression with threshold.

2) Termination cases with hidden variables These cases are similar to the previous one, with the following differences:Includes hidden variables, i.e. promotions, soccer championships, weather or other external events influence consumer habits. These hidden variables distort the trend, Classification by linear regression model. was suggested. The model for these cases should be readable. Additional information about inputs, hidden possibilities variable. 3) Difference between churn case and sporadic case sales records often include 20% to 30% of customers. Those who come to shop sometimes. In contrast for regular customers who come regularly. Once or several times a week, or even every two weeks. Regular customers take our business foremost They do their biggest shopping there. If even then you can even make small purchases at once if you want closer to or more accessible than competitor stores. Basically, churners are regular customers become sporadic. This study does not focus on sporadic cases.

III. Methodology

A. Data collection

The study includes sales data for his 5,115,472 lines distributed among his 105,488 customers in his chain of French supermarkets. Overall, pseudonymization has taken place datasets that offer the possibility of linking personal information(age, length of customer relationship, gender, population). Customer's city) Sales time series in this study. Ranks are divided into groups, each with groups identify customers and are formatted in two dimensions line. Outliers like customers with two active profiles above Identical names and addresses have been removed.

B. A highly imbalanced dataset

Different traits can be seen depending on the species data. Any of these characteristics can make it difficult a data mining algorithm for extracting useful patterns. One of the main issues of the dataset used for this study are class imbalance. If there is an imbalance between In general, learning algorithms tend to only make predictions. Majority class to minimize errors. For example, there were only 60,000 active customers 2 thousand charners. That's 96% of active consumers Only 4% of people are positive about churn, representing typical churn cases class imbalance. Resampling is a widely used addressing technique.The problem of class imbalance. There are two ways to do this:

DOI:10.46647/ijetms.2023.v07i01.052 ISSN: 2581-4621

Either oversampling or undersampling can be used. When Under sampling, only a subset of the majority class is used train your model. In this study, a random sample of inputs was created excluded from the set of active clients, Equal number of churns and non-churns.

C. Data Augmentation by Scaling

It consists in artificially generating other time series to increase the number of class samples. Scale of the original time series. From a customer who always takes care of me you can buy around 100 euros per week and generate cryptocurrency one who buys about 50 euros every week, and another People who buy about 200 euros every week. the latter The two dynamics are exactly the same as the first dynamic. Idea is to artificially create new customers with the same behaviour Same as the original, but with a different scale.

IV. MODELS

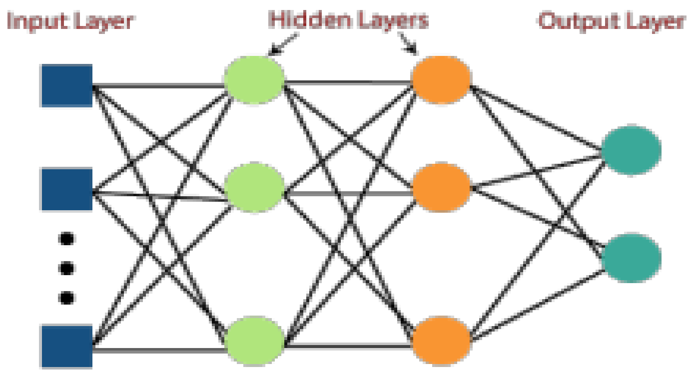
In the following, a linear regression model was compared to some machine learning approaches that have been proven to work for churn prediction in general. Efficiency and reliability were considered during the comparison.

A. Linear regression

Linear regression is known as the linear approach in statistics a scalar response and One or more explanatory variables (also called dependent and independent variable) linear regression is light weight. Statistical tools that help you gain insight into customer behavior and understand your business and influencers Profitability. Linear regression has limited applicability in the business world, Although the dependent variable is continuous in nature, it is still a known technique in situations where it can be used [11].

B. MLPs

Multilayer Perceptron (MLP) is a variant of the original Perceptron model proposed by Rosenblatt in 1950 . MLPs With the advent of Deep, it has proven itself in many areas
learning . One or more hidden layers in between its input and output layers. neurons are structured layer, connections are always directed from the lower layer Neurons in the same layer as the upper layer are not connected.



. A graphical representation of a Multilayer Perceptron.

C. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is an open source library that provides an efficient and effective implementation of the gradient boosting algorithm. XGBoost has become's go-to method and often a key component in winning solutions to classification and regression problems in 's machine learning competitions. As this study is subject to classification problems, it was considered useful to test the results for comparison with other approaches.

D. LSTM

DOI:10.46647/ijetms.2023.v07i01.052 ISSN: 2581-4621

long short-term memory (LSTM) networks are a class of recurrent neural networks (RNNs) that can learn order dependence in sequence prediction problems was introduced in his late 1990s, but it is only recently that LSTM models have become viable and powerful time series forecasting techniques. LSTM solves a big problem that plagues recurrent neural networks. H. Failure to learn dependencies from long sequences. Using a series of "gates", each with its own RNN, LSTM keeps, forgets, or ignores data points based on probabilistic models. See Figure 6 for the LSTM plot.

V. Evaluation Metrics

precision, recall, and F measures are used in this paper as measurement tools to measure the reliability of various predictive models. TP and FP denote true-positive and false-positive samples, respectively, and true-negative and false-negative samples are denoted by TN and FN, respectively. Recall is the percentage of churn customers with correctly identified. Recall is intuitively the classifier's ability to find all positive samples.

VI. Results and Discussion

A. Cross Validation

Cross validation is commonly used to evaluate regression and classification models. Application to time series or other naturally ordered data adds chronological complexity to event. Two techniques are possible.The first is to define a time interval to retrieve data for all customers and apply the k-fold method . To do this, we need to divide the data set into k equal parts called folds. where k-1 folds are the training set and the remaining folds are the test or validation set. Repeat k times with different modulo convolutions as a test set everytimes. The final result is the average of the validation results at the end of each iteration. Several k were tested and the best result of was achieved with k = 4.
The second technique is to choose two time intervals from which to extract data. There are two subsets, each splitting the into two parts. Therefore, four subsets are obtained. These subsets can use the k-fold method is therefore called quadruple. 4711 B. Discussion 4484 We evaluated the performance of the tested classifier using a dataset containing time series and additional information for each customer.

VIII. Conclusion

This post introduces the application of machine learning modeling customer sales data to predict churn. A total of four statistical and machine learning models were compared. This study focuses on predictive performance. With reputation Precedes Machine Learning Models, Easy to suggest that they may be superior to traditional approaches, Companies can use the predictive models proposed there to target future retention marketing campaigns to learn. Future research may use other methods such as Transformers [28]. Since the last machine learning model process more data, External data such as weather, neighbourhood data, and averages income per capita will be very interesting.

REFERENCES

[1] O. O. Aalen. A linear regression model for the analysis of life times. Statistics in medicine, 8(8):907–925, 1989.
[2] S. Abiteboul. Querying semi-structured data. In International Conference on Database Theory, pages 1–18. Springer, 1997.
[3] A. K. Ahmad, A. Jafar, and K. Aljoumaa. Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 6(1):1–24, 2019.
[4] C. Bergmeir and J. M. Ben´ıtez. On the use of cross-validation for time series predictor evaluation. Information Sciences, 191:192–213, 2012.
[5] M. Braun and D. A. Schweidel. Modeling customer lifetimes with multiple causes of churn. Marketing Science, 30(5):881–902, 2011

DOI:10.46647/ijetms.2023.v07i01.052 ISSN: 2581-4621

[6] W. Buckinx and D. Van den Poel. Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual fmcg retail setting. European journal of operational research, 164(1):252–268, 2005.
[7] J. Cao, X. Yu, and Z. Zhang. Integrating owa and data mining for analyzing customers churn in e-commerce. Journal of Systems Science and Complexity, 28(2):381–392, 2015.
[8] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, et al. Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4):1–4,2015.
[9] P. S. Cowpertwait and A. V. Metcalfe. Introductory time series with R.Springer Science & Business Media, 2009.
[10] A. Dingli, V. Marmara, and N. S. Fournier. Comparison of deep learning algorithms to predict customer churn within a local retail industry.International journal of machine learning and computing, 7(5):128–132, 2017.
[11] W. W. Eckerson. Predictive analytics. Extending the Value of Your Data Warehousing Investment. TDWI Best Practices Report, 1:1–36, 2007.
[12] A. Estabrooks, T. Jo, and N. Japkowicz. A multiple resampling method for learning from imbalanced data sets. Computational intelligence, 20(1):18–36, 2004.
[13] J. Gonzalez and W. Yu. Non-linear system modeling using lstm neural networks. IFAC-Papers OnLine, 51(13):485–489, 2018.
[14] I. Goodfellow, Y. Bengio, and A. Courville. Deep learning. MIT press, 2016.
[15] R. A. Hoban. Introducing the slope concept. International Journal of Mathematical Education in Science and Technology, pages 1–17, 2020
[16] Ravindra Parab; Pawan Pandey. "Design of Real-Time Multi-Level Inverter with Remote Monitoring System". International Research Journal on Advanced Science Hub, 3, Special Issue 6S, 2021, 6-10. doi: 10.47392/irjash.2021.157
[17] Mohammed Fasiuddin; Anupama A Deshpande; Mahabood Shaik. "Strengthening WSN Applications in Oil and Gas Industries for Multiple Monitoring Frameworks". International Research Journal on Advanced Science Hub, 2, 6, 2020, 150-153. doi: 10.47392/irjash.2020.54
[18] R. Devi Priya, R. Sivaraj, Ajith Abraham, T. Pravin, P. Sivasankar and N. Anitha, Multi-Objective Particle Swarm Optimization Based Preprocessing of Multi-Class Extremely Imbalanced Datasets", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 30, No. 05, pp. 735-755 (2022), doi: 10.1142/S0218488522500209
[19] Gyanendra Kumar Pal; Sanjeev Gangwar. "Discovery of Approaches by Various Machine learning Ensemble Model and Features Selection Method in Critical Heart Disease Diagnosis". International Research Journal on Advanced Science Hub, 5, 01, 2022, 15-21. doi: 10.47392/irjash.2023.003
[20] Kousik Bhattacharya; Avijit Kumar Chaudhuri; Anirban Das; Dilip K. Banerjee. "A Data Mining based study on Dengue Fever: A Review". International Research Journal on Advanced Science Hub, 4, 04, 2022, 101-107. doi: 10.47392/irjash.2022.025
[21] Nusrath Unnisa A; Manjula Yerva; Kurian M Z. "Review on Intrusion Detection System (IDS) for Network Security using Machine Learning Algorithms". International Research Journal on Advanced Science Hub, 4, 03, 2022, 67-74. doi: 10.47392/irjash.2022.014
[22] Mahalakshmi G.; Shimaali Riyasudeen; Sairam R; Hari Sanjeevi R; Raghupathy B.. "A Survey: Effective Machine Learning Based Classification Algorithm for Medical Dataset". International Research Journal on Advanced Science Hub, 3, Special Issue 9S, 2021, 28-33. doi: 10.47392/irjash.2021.245
[23] Muniyandy Elangovan; Yuvan Siddarth B.; Govindram Uduupa R; Vigneshwar R. "Weapon Detection System in Remotely Operated ASV". International Research Journal on Advanced Science Hub, 3, Special Issue 9S, 2021, 34-37. doi: 10.47392/irjash.2021.246

DOI:10.46647/ijetms.2023.v07i01.052 ISSN: 2581-4621

[24] Dhanya S Karanth; Kumaraswamy H.V; Rajesh Kumar. "Workaround prediction of cloud alarms using machine learning". International Research Journal on Advanced Science Hub, 3, 8, 2021, 164-168. doi: 10.47392/irjash.2021.231

[25] Suraj S. Bhoite; Swapnali K. Londhe. "Aspect Based Online Sentiment Analysis Product Review and Feature Using Machine Learning". International Research Journal on Advanced Science Hub, 3, Special Issue 7S, 2021, 54-59. doi: 10.47392/irjash.2021.209

[26] Logeswari T. "Performance Analysis of Ml Techniques for Spam Filtering". International Research Journal on Advanced Science Hub, 2, Special Issue ICIES 9S, 2020, 64-69. doi: 10.47392/irjash.2020.161

[27] Sona Solanki; Asha D Solanki. "Review of Deployment of Machine Learning in Blockchain Methodology". International Research Journal on Advanced Science Hub, 2, 9, 2020, 14-20. doi: 10.47392/irjash.2020.141

[28] Salini Suresh; Suneetha V; Niharika Sinha; Sabyasachi Prusty. "Latent Approach in Entertainment Industry Using Machine Learning". International Research Journal on Advanced Science Hub, 2, Special Issue ICARD 2020, 2020, 304-307. doi: 10.47392/irjash.2020.106

[29] Salini Suresh; Suneetha V; Niharika Sinha; Sabyasachi Prusty; Sriranga H.A. "Machine Learning: An Intuitive Approach In Healthcare". International Research Journal on Advanced Science Hub, 2, 7, 2020, 67-74. doi: 10.47392/irjash.2020.67