Predict & Prevent Customer Churn



***Double Click on video*** *to Play*

*Video illustration of problem & solutioning using data science*

Vidyasagar Kummarikunta

Bellevue University

Bellevue, NE 68005, USA

vkummarikunta@my365.bellevue.edu

Julia Cuellar

Bellevue University

Bellevue, NE 68005, USA

jucuellar@my365.bellevue.edu

Jaimie Cain

Bellevue University

Bellevue, NE 68005, USA

jcaingreer@my365.bellevue.edu

Manish Shekar

Bellevue University

Bellevue, NE 68005, USA

mshekhar@my365.bellevue.edu

Jedediah Brough

Bellevue University

Bellevue, NE 68005, USA  
jebrough@my365.bellevue.edu

Paste the appropriate copyright/license statement here. ACM now supports three different publication options:

• ACM copyright: ACM holds the copyright on the work. This is the historical approach.

• License: The author(s) retain copyright, but ACM receives an exclusive publication license.

• Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single-spaced in Verdana 7 point font. Please do not change the size of this text box.

Each submission will be assigned a unique DOI string to be included here.

# Abstract

The purpose of the project is to study customer's behavior and data to predict and possibly prevent customer churn using data science, business intelligence and business process integration. Customer churn occurs when customers or subscribers stop doing business with a company or service. Also known as customer attrition or turnover, it is a critical metric due to the cost savings for a company to retain a customer in comparison to acquiring new ones. Awareness about customer churn helps an organization define customer retention processes, project goal success rates and identify strategies for improvement.

# Author Keywords

Customer churn, Retention, Attrition, Prediction, Logistic Regression, Classification, Data Science, Machine Learning, EDA, feature engineering, data cleansing, bias, outlier, training, validation, test, feature scaling, data skew, balancing, modeling, KNN, SVM, decision tree, K-fold, parameter tuning, correlation, NLP, Stemming, Lemmatization, TF-IDF, Sentiment Analysis, Report, Dashboard, Business Intelligence, CRISP-DM.

# Current State (Case Studies):

1. <https://www.vttresearch.com/sites/default/files/julkaisut/muut/2006/customer_churn_case_study.pdf>
2. <https://www.elderresearch.com/hubfs/Resources/Elder_Research_Case_Study_Predicting_Account_Churn_in_Retail_Banking.pdf>

# ACM Classification Keywords

1. Computing methodologies~Artificial intelligence;500
2. Computing methodologies~Machine learning~Machine learning approaches~Classification and regression trees;500
3. Computing methodologies~Machine learning~Learning paradigms~Supervised learning~Supervised learning by classification;500
4. Computing methodologies~Artificial intelligence~Natural language processing~Information extraction;300
5. Computing methodologies~Machine learning~Machine learning algorithms~Feature selection;100

# Introduction

One of the key aspects of a successful business model is how to establish and maintain a loyal customer base. If a company can keep their customers in a subscription-based revenue system, they can maintain a strong financial foundation and increase their return on investments. The cost savings can then be used to adopt better product technology and to create a more diverse marketing strategy to provide better support & decrease customer churn. Through a stable customer base, the company can establish a strong customer review system and increase the likelihood of customer referrals as a source of increased product subscriptions & acquire new customers. Major reasons for customer attrition include bad quality product, the low cost for customers to switch to competitors, and bad customer service. Companies should take advantage of their data resources, apply data analytics & data science techniques to discover insights that can help predict & prevent customer churn. Next we have discussed few such approaches utilizing data science.

*Figure 2 – SVM*

*Figure 1 – Logistic Regression*

# Data Science Approach – Predict Churn

# A picture containing rain Description automatically generated

A close up of a map

Description automatically generated

A close up of a map

Description automatically generated

A screenshot of a cell phone

Description automatically generated

To predict customer churn, we will apply data science techniques & core principles of CRISP-DM life cycle as follows:

*Figure 3 – KNN*

1. **Business Understanding**: Conduct sessions to understand the problem, discuss possible causes, and understand known impacts. Considering retail domain.
2. **Data Understanding**:
   1. **Data Acquisition**: Identify & source relevant data i.e. considered, rewards & offer used, surveys, online reviews, complaints etc.
   2. **EDA (Exploratory data analysis)**: Clean data & identify valuable attributes as features. **Data cleansing** involves null value handling, duplicate removal, standardization & formatting. **Feature engineering** involves visualizing data using histograms, pie charts, correlation matrix, & scatterplots to identify relevant features, understand trends, & **bias** within the data.

*Figure 4 – Decision Trees*

1. **Data Preparation**: Split data into **training** (50%), **validation** (20%) & **test** (30%) sets.

Apply **feature scaling** to normalize data to identify any **outliers** or **redundant values** that may **skew** the data & analysis. This in turn may require data **balancing**, to keep sample consistent with the larger data set.

1. **Modeling**: Fit & train the model. As this is prediction (classification) problem **logistic regression** (Fig -1), **KNN** (K nearest neighbors – Fig 3), **SVM** (Support Vector Machine – Fig 2), **decision trees** (Fig - 4) algorithms are considered.Predict from each model.
2. **Evaluation**: Prediction results from each model are evaluated using **confusion matrix** & measures like **accuracy score**, **F1 score**, **precision score** & a **recall score**. Based on evaluation result best model is chosen. If none gives preferred results model **parameter tuning** is tried and model combinations can be tried and re-evaluated. **K-fold cross validation** is applied to select only effective features. Model selected is recreated with chosen features to run light when deployed.
3. **Deploy**: Model inputs & outputs are integrated with data processing workflow & deployed in prod.

Data Science Approach – Sentiments, Tweets & Complaints

Companies run customer surveys to understand what customer thinks about their products & services. Customer feedback usually answers to series of predefined questions & free form text are explored using Natural

A picture containing table

Description automatically generatedA screenshot of a computer

Description automatically generated

*Figure 5– Topic Modeling*

Language Processing **(NLP)**. Preprocessing removes stop words, applies, **stemming**, **lemmatization**, & **TF-IDF**(Term Frequency-Inverse Document Frequency) to convert words into matrix of numbers. Series of binary classification models can then be used to classify text as

positive, negative, neutral (See Fig 6,7). Each word is a feature with weight for these models which can be varied to achieve desired results. Similarly, tweets can also be analyzed to understand feedback & complaint texts (call transcripts) that can be analyzed to understand complaint reasons & uncover product issues. Topic modeling can be considered here to discover words frequently mentioned with respect to certain contexts (see Figure 5).

*Figure 6 – Sentiment Word Cloud*

# Business Intelligence – Presenting Insights

Outputs from the above data science applications are integrated with customer & other base data and dashboards are created for business executives, to understand customer behavior, attrition reasons, other attributes trend over time for churning customers.

# Customer Churn Prevention

Outputs from data science applications and BI insights are integrated with marketing processes & sales screens to provide real time customer insights, early warnings, and custom guidance for retention. In addition, custom offers and campaigns can be run for the purpose as well.

*Figure 7 – Sentiment Analysis*

A screenshot of a cell phone

Description automatically generated

A picture containing drawing

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A close up of a toy

Description automatically generated

# What is the deliverable for this project?

1. Pre-trained (supervised) model to predict customer churn.
2. NLP (Natural Language Processing) model to understand customer sentiment overall and for each product (positive, negative, neutral).

*Figure 8 – Executive Dashboard*

1. Process integration with marketing & sales screens to provide real time customer insights, early warnings, and custom guidance for retention.
2. Executive dashboards to understand customer behavior and attributes trend over time for churning customers.
3. Executive dashboards to analyze customer churn reasons (i.e. product issues, bad customer service, bad onboarding, budding competition, etc.) and predict trends over time for customer churn.

*Figure 9 – Integration with sales*

# Conclusion

# In this paper, a customer churn analysis is presented in a retail domain. The analysis will focus on churn prediction based on logistic regression, KNN (K-nearest neighbor), SVM (Support Vector Machine) and decision trees. We will train and evaluate multiple models and chose the best in terms of performance. New features will be generated from NLP based feedback analytics and complaints analytics models. We will identify and treat outliers to minimize impact to training data and prediction accuracy. We may do extreme value analysis, calculate z-score, or use K-means clustering to identify outliers and to treat them, we may use random sampling, trimming or top/bottom zero encodings. We will run multiple iterations of model creation, re-evaluation & parameter optimization. Performance improvements considered are choosing multiple probability thresholds for a logistic regression model, in case of KNN model consider feature rescaling, i.e., Min-max scaling and standard scaling, in case of SVM model consider Kernel trick (for multiple input features) & optimize C penalty parameter to achieve best possible decision boundary, and in case of decision tree model utilize pruning. We will consider the ensemble model to aggregate prediction from multiple models & create final prediction, done to minimize generalization error of the prediction.

Lastly, we will integrate model results with other data elements to create summaries and trends on BI dashboards for decision making. Models will also be integrated with real time customer facing and customer support applications.

Churn prediction is vital for devising appropriate customer retention strategies. Customer retention = Customer acquisition, it is cheaper and easier to sell to existing customers than to acquire new ones. More customer's company has more critical retention is.

# Acknowledgments

Thanks to work from home situation that inspired us to pursue the MSDS degree. Also, thanks to churning customers everywhere, without whom we will not have this problem and any training data.

# Web References

1. <https://blog.dataiku.com/churn-analytics-marketing-team-best-friend>
2. <https://towardsdatascience.com/how-to-leverage-ai-to-predict-and-prevent-customer-churn-f84d653a76fb>
3. <https://towardsdatascience.com/hands-on-predict-customer-churn-5c2a42806266>
4. <https://towardsdatascience.com/churn-prediction-770d6cb582a5>
5. <https://www.kdnuggets.com/2019/05/churn-prediction-machine-learning.html>
6. <https://www.bizdata.com.au/blogpost.php?p=customer-churn-prediction-with-algorithms>
7. <https://www.zylotech.com/blog/how-to-predict-and-prevent-your-customer-churn>
8. <https://www.comtecinfo.com/rpa/resources/case-study-reduce-customer-churn-using-predictive-analytics/>
9. <https://www.elderresearch.com/hubfs/Resources/Elder_Research_Case_Study_Predicting_Account_Churn_in_Retail_Banking.pdf>

# Academic References

1. Luo, B., Shao, P., Liu, J.: Customer Churn Prediction Based on the Decision Tree in Personal Handy phone System Service. In: International Conference on Service Systems and Service Management, pp. 1—5 (2007)
2. B. Baesens Analytics in a Big Data World: The Essential Guide to Data Science and its Applications. John Wiley & Sons (2014)
3. K. Coussement, K.W. De Bock. Customer churn prediction in the online gambling industry: the beneficial effect of ensemble learning. J. Bus. Res., 66 (9) (2013), pp. 1629-1636
4. W. Verbeke, D. Martens, C. Mues, B. Baesens Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert Syst. Appl., 38 (3) (2011), pp. 2354-2364
5. N.J. Radcliffe, R. Simpson. Identifying who can be saved and who will be driven away by retention activity. J. Telecommun. Manager., 1 (2) (2008)
6. D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, D. Dennison. Hidden technical debt in machine learning systems
7. C. Cortes, N.D. Lawrence, D.D. Lee, M. Sugiyama, R. Garnett (Eds.), Advances in Neural Information Processing Systems 28, Curran Associates, Inc. (2015), pp. 2503-2511
8. Y. Xie, X. Li, E. Ngai and W. Ying, Customer churn prediction using improved balanced random forests, Expert Systems with Applications, vol. 36, no. 3, pp. 5445-5449, 2009
9. Scott A. Neslin, Sunil Gupta Wagner Kamakura Junxiang Lu Charlotte Mason, Defection Detection: Improving Predictive Accuracy of Customer Churn Models
10. Qureshii SA, Rehman AS, Qamar AM, Kamal A, Rehman A. Telecommunication subscribers' churn prediction model using machine learning. In: Eighth international conference on digital information management. 2013. p. 131–6
11. Burez D, den Poel V. Handling class imbalance in customer churn prediction. Expert Syst Appl. 2009;36(3):4626–36
12. Piatesky, G. (2014, October). *CRISP-DM, still the top methodology for analytics, data mining, or data science projects*. Retrieved from KDnuggets: <https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html>
13. Pink, D. (2009). *Drive:*[*The surprising truth about what motivates us*](https://en.wikipedia.org/wiki/Drive:_The_Surprising_Truth_About_What_Motivates_Us)*.*
14. Rich Caruana. Multitask learning. Machine Learning, 28:41–75, 1997. DOI: 10.1007/978-1- 4615-5529-2\_5.
15. Umayaparvathi V, Iyakutti K. A survey on customer churn prediction in telecom industry: datasets, methods and metric. Int Res J Eng Technol. 2016;3(4):1065–70.
16. He Y, He Z, Zhang D. A study on prediction of customer churn in fixed communication network based on data mining. In: Sixth international conference on fuzzy systems and knowledge discovery, vol. 1. 2009. p. 92–4.
17. Ascarza E, Iyengar R, Schleicher M. The perils of proactive churn prevention using plan recommendations: evidence from a field experiment. J Market Res. 2016;53(1):46–60.
18. Trafalis, Theodore B. Support vector machine for regression and applications to financial forecasting, Proceedings of the International Joint Conference on Neural Networks, v 6, (2000) 348-353.
19. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929-1958, 2014.
20. Berry, M. J. A., & Linoff, G. (1999). Data mining techniques: for marketing, sales, and customer support. Morgan Kaufmann Publishers.
21. Burez, J., & Van den Poel, D. (2007). CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. Expert Systems with Applications, 32(2), 277–288.
22. Burez, J., & Van den Poel, D. (in press). Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department. Expert systems with applications. doi:10.1016/ j.eswa.2007.07.036