|  |
| --- |
| Summative Assessment Report |
| Applied AI |
| 07/03/2022 |

Total Number of Words = 3999

Table of Contents

**Executive Summary2**

**Introduction3**

**Literature Review4**

Current Business Problem and AI Based Solutions4

Techniques to Address Business Problems and their Evaluation5

Critical Evaluation of Solutions Presented in the literature6

**Research Design6**

Assumptions6

Data Pre-processing6

Optimisation Technique7

Supervised Learning Technique7

Validation Technique8

Evaluation Measure 9

Algorithmic parameters9

Sampling Technique10

**Experimental Results and Analysis11**

**Conclusion12**

**References13**

**Executive Summary**

The telecommunication sector is facing hurdles in preserving customers, necessitating the use of churn prediction model to predict customer turnover and take appropriate measures ahead of time. My proposed technique is divided into 6 stages. Data cleaning and data pre-processing are carried out in the first two steps. The third phase is to build a supervised learning model on the full dataset and evaluate its performance. The supervised learning model I have used is Gaussian Naïve Bayes. In the fourth step, I have performed Feature Selection (FS) by implementing Genetic Algorithm (GA), built a supervised learning model on the derived subset of features and evaluated the robustness of the generated model using several evaluation measures, like precision, recall, and F-measure. In the last two phases, I have balanced the imbalanced dataset by using a hybrid technique to generate synthetic samples for minority class and remove irrelevant instances from both classes[10]. After balancing dataset, I have then repeated steps 3 and 4. To balance the dataset, I have used a hybrid sampling approach called Synthetic Minority Oversampling Technique with Edited Nearest Neighbor (SMOTEENN). I have used a validation technique called K-fold cross-validation to evaluate the model’s accuracy. Additionally, I have used 2 stopping criteria to terminate the GA early. The first is to check if the fitness score of any chromosome in the population is greater than the optimal score then stop the program. The second is by fixing the number of generations, I have placed an upper limit [19]. Furthermore, I have taken advantage of elitism and tournament selection strategies and used both to create a population, hence ensuring the best chromosomes are chosen for each generation. Crossover and mutation are used to create a hybrid in the next generation. At the end I have compared all the models and derived conclusion by giving empirical evidence.

The comparison results shows that the Genetic Algorithm based FS has improve the performance of Naïve Bayes model significantly. In future, it would be worth considering improving the accuracy by including more algorithmic parameters and evaluation metrics. It would be valuable to experiment other optimization algorithms in future, to ensure the global optimum is reached.

**Introduction**

One of the common problems across industries around the world is the customer attrition[9]. According to Statista, the customer churn rate in the United States in 2020 is around 25% for general, 22% for online, and 21% for telecom/wireless retailers [8],[9]. Customer churn is the rate at which a business loss its customers within a certain timeframe[2]. The telecommunication industry is encountering challenges in customer retention. Moreover, obtaining fresh customers is far more expensive than retaining existing ones[2]. This necessitates the adoption of a churn prediction model to forecast customer attrition and take relevant actions ahead of time[2]. A customer churn predictive algorithm can accurately identify at-risk customers so that a retention strategy can be offered to them[1]. However, the imbalanced nature of telecommunication datasets are the major impediments in achieving the accurate churn prediction performance[18].

To improve computer applications, most of the businesses around the world make use of Artificial Intelligence (AI) to achieve high accuracy and adaptability over time[3]. AI is the discipline that aims to create artificial animals or artificial people like systems that think, reason, make decisions and perform actions based on their percepts[4]. These systems make use of Machine Learning (ML) to learn and improve itself more and more by time[3]. Its ever-ending journey of improvement and adaptation allows businesses to forecast their customers’ behaviour and devise strategies to hit their targets.

The customer churn problem is one of the major problems faced by most of the businesses. In order to retain customers, companies have to predict in advance who are most expected to leave their services[5]. This can be forecasted by first exploring customer’s information regarding services they used, monthly/yearly bills they have paid, and some of their personal data, for example occupation, age, etc., and then model the data using appropriate AI techniques, for example, supervised learning and feature selection optimization. These techniques are used to remove irrelevant features from the complex dataset and help resolve the class imbalance issues by selecting appropriate parameters. This eventually can help the company to take immediate actions and plan different strategies on how to keep customers attracted to their services[9]. This can result in improving other areas, for example, sending focused marketing emails/messages about the new offers to the existing customer’s or providing discounts. AI techniques also helps improve the efficiency of churn prediction model to make accurate predictions and reduce computation cost[2].

This report uses a Churn dataset that is imbalanced as there are substantially more preserved customers than churning ones. We are interested in predicting the attributes that influence “Churn” the most. Sometimes even by choosing the appropriate parameters to minimise the class imbalance issues, still the algorithm is more skewed towards predicting the majority class. To resolve this issue, it is worth trying to balance the dataset by using appropriate sampling techniques, for example a hybrid method[6]. We can then compare the results of balanced and imbalanced datasets.

Apart from choosing the appropriate parameters and balancing the class labels, it is also important to choose suitable performance metrics to evaluate the supervised learning model. The metrics chosen should be the one that minimises the effect of data imbalance issue on the score. For example, F1 macro-averaged score is a reliable metric in imbalanced cases as it is insensitive to the imbalance of the classes and considers them all as equal[7]. It is best suitable for the cases when the minority class is more important, and you want to eliminate the bias towards the majority class[7].

So, an accurate churn prediction performance can be achieved by choosing appropriate algorithmic parameters, using suitable supervised learning model, applying feature selection optimization, balancing the dataset, and selecting suitable evaluation metrics. This benefits businesses to accurately forecast customer attrition and understand customers’ behaviour to minimise the churn rate.

**Literature Review**

1. Current Business Problem and AI Based Solutions:

Researchers are currently working to reduce the obstacles associated with forecasting client attrition using standard methods[12]. Researchers are looking to investigate different ways to improve the accuracy of the customer churn prediction model[11].

Current churn predictive models have limited characteristics generated generally from demographic data, transactional data, social network data, and other behaviour data[15]. The focus is now on investigating new features which could significantly affect customer churn and hence enhance the prediction accuracy[12],[13].

Researchers are now considering thinking outside the box and experimenting with new ways to increase the efficiency of churn prediction. In the past, churn prediction classifiers were domain specific and there was a scarcity of study focusing on customer-company interaction features such as email messages, phone conversations, and customer care chats[10]. Due to the scarcity of information, a single source of data cannot correctly anticipate e-commerce consumer attrition[12].

Client data in telecommunication sector is multidimensional and imbalanced, making traditional models even less efficient[12]. Using appropriate Machine Learning methods and feature extraction techniques can help minimize this issue. Over-sampling of minority class/under-sampling of majority class can help make the class balance[10]. However, oversampling allows redundant instances to be replicated, hence causes overfitting, and under-sampling causes loss of information[10]. A hybrid sampling strategy is more robust sampling technique that seeks to combine the advantages of the previously stated sample strategies while attempting to address their limitations[10].

1. Techniques to Address Business Problems and their Evaluation:

Recent research has demonstrated that studying customer contacts by assessing the social proximity of recent churners might increase churn prediction accuracy[12]. [12] is the recent study attempted to skip traditional service usage-based parameters and choose to focus on customer behaviour through modelling the interactions between customers. They have presented a unique algorithm called Group-First Churn Prediction, which eliminates the need to know in advance who has freshly churned by using the information based on customer interactions[12]. This research uses only the past call data records and did not use the financial or demographic information. Their results showed that this approach can accurately predict churn among a large population[12].

Another study has considered to model customer company interaction and found that the customer-initiated contacts to service providers through multiple channels show an overall higher customer loyalty[13]. Moreover, in many recent studies text analytics is being used with churn prediction model to extract useful features from raw data and anticipate churn [14]. By incorporating textual information in a customer churn prediction (CCP) model and extracting features using Convolution Neural Network (CNN) from the raw textual data can significantly boost the predictive performance[14]. Content mining linked together with Natural Language Processing (e.g., sentimental analysis) on raw data has enhanced existing predictive modelling in recent years, yielding extremely remarkable results[15].

To reduce the class imbalance issues, Synthetic Minority Oversampling Technique (SMOTE) is a popular over-sampling approach used to generate new samples for the minority class [10]. Another technique called Synthetic Minority Oversampling Technique with Edited Nearest Neighbor (SMOTEENN) is a hybrid sampling technique, which is a combination of over-sampling and under-sampling approaches[10]. Moreover, in [18], an efficient strategy for oversampling minority classes termed HEOMGA is provided by integrating Heterogeneous Euclidean-Overlap Metric (HEOM) and Genetic Algorithm (GA). The HEOM is used to determine the GA's fitness function[18].

1. Critical Evaluation of Solutions Presented in the literature:

There are many techniques around the world to make the customer churn prediction accurate. Which one to use, depends on the purpose and what you expect as the outcome. It also depends on the task you are working on and the kind of data you have at hand. For example, if you are doing predictive modelling and want to forecast churn by analysing customer’s behaviour, you will most certainly have an imbalanced data at hand. You then will focus on first cleaning the dataset, then balance the labels using a sampling technique, use optimization algorithm to select features, and lastly built your model to predict churners. However, what techniques you will use in each step depends on your experience and wisdom.

**Research Design**

I have worked with the Python language in Jupyter Notebook.

1. Assumptions:

First assumption is that as the dataset contains missing values, it is important to remove them from the dataset so that there are no noise/misleading/NULL values, and the classifier could accurately predict values using appropriate inputs. Any ML algorithm's performance and accuracy suffer as a result of these null values.

Second assumption is that as the dataset is imbalanced, it will be worth trying to balance the dataset using an appropriate sampling technique. Then by comparing the results you could evaluate whether the imbalanced dataset is appropriate to use or balanced dataset is more meaningful for the purpose.

1. Data Pre-processing:

The potential modelling solution would be to first explore the dataset and identify what needs to be fixed, clean the dataset if there are missing values or redundant columns, encode the categorical variables into integer/binary, and normalize and standardize the integer/float variables to rescale them within a range.

What I have done is that, after importing the dataset I have first started with the data exploration to understand the dataset and identify what needs a correction. I have displayed the first 5 rows of the dataset, column name, shape (number of rows and columns), data types of all the columns, printed information about all the columns in the dataset, checked if there is any NULL value in any column and counted the number of NULL values in each column. There were 14 columns that contain NULL values. There is no column that contain more than 60% of the NULL value. Therefore, I have replaced all the NULL values: in the integer columns they were replaced by their mean value, whereas, in the object columns they were replaced by the most frequent value.

Then I proceeded to encode the categorical data into integer/binary as ML models require all input and output variables to be numeric. I have first displayed the unique values in each column, so to check if I need to manually encode certain values. I only needed to manually encode one column called “CreditRating” to split the categorical value from the integer and only keep the integer part. All the other categorical columns were encoded using a technique called “Label Encoder” as it encodes the categorical variable into numeric values between 0 and n classes [20]. Only the column “ServiceArea” was encoded using “cat codes” as it contains too many categorical values with no order.

I have tried normalizing and standardizing the integer/float variables (not the ones I have encoded above) using the MinMaxScalar function. However, my classifier was then working weird, i.e., the GA was selecting only 1 variable rather than a subset of features. Therefore, I skipped that step.

1. Optimisation Technique:

The potential modelling solution would be to choose the one which skips the local optima and possibly reach the global optimum (or at least near the global optima). GA is the algorithm that searches different sections of the parameter space with each successive generation, then steers the search to the fittest regions or at least to those with enhanced performance[21]. This raises the chances of discovering the global optimum[21].

I have used Genetic Algorithm (GA), as my optimization technique, to find the subset of relevant features for my supervised learning model. The randomness in the selection and reproduction makes GA powerful, as in each generation they can simulate any point in the given parameter space, hence skipping from the local optima[21]. Therefore, I have chosen GA to select features for my churn prediction model.

1. Supervised Learning Technique:

The potential modelling solution for the churn classification would be to choose a classifier that does not require much training data (as the dataset provided is not too big) and can handle both continuous and discrete data[22]. As the purpose is to forecast customers’ churn, the classifier should be highly scalable with the number of instances, because in future customers might increase. Classifier also needs to be fast to make quick predictions because an organization doesn’t want to lose their customers due to a slow classifier. Keeping all these requirements in mind, Naïve Bayes is the best classifier for prediction and classification tasks, like medical diagnosis, churn prediction, or news forecast[22]. Therefore, I have used Naïve Bayes for this task.

I have not used any parameter for my model, as there was no need. (Fig. 1) shows how I have built my model to classify churners.

1. Validation Technique:

The validation technique I have used is K-fold Cross-Validation, as the holdout validation approach has a few pitfalls[28]. For holdout validation, it is very important to split the dataset into the right proportion into the train and test sets because we want our model to generalize well to new examples. Our goal is to keep enough samples in the training set that can be sufficient for training the model and enough samples in the test set to avoid the chances of overfitting to the training set[28]. However, it is hard to guess the right proportion for each set. The solution is to use K-fold cross-validation, as the data is separated into k folds[28]. The model is trained on k-1 folds, with 1-fold saved for testing[28]. This method is an iterative process: continued until every fold of the dataset has contributed as a test set[28]. Therefore, this is the best validation-technique to choose for our model.

I have used 10-folds to split the dataset. I have first specified the number of folds, then built the model, and then computed cross-validation score (Fig. 1).

Graphical user interface, text, application

Description automatically generatedFig.1. K fold cross validation to evaluate performance of gaussian naïve bayes model.

1. Evaluation Measure:

Let’s now have a look into the three-evaluation metrics: F1-score, recall, and precision, and discuss which one will be most appropriate for our churn prediction model.

Precision is the ratio between the True Positive and total positives[23]. For the churn prediction scenario, it would be the measure of customers that we have predicted correctly as churners out of all the customers actually churning. Alternatively, you can say that when the model predicts that a customer will churn then how many percent of the times is it correct?

Recall is a measure of how well our model identifies True Positives[23]. For the churn prediction scenario, for all the customers actually churning, it would show us how many we correctly predicted as churners.

Now let’s see among precision and recall which is more suitable for our scenario. As the client is interested in forecasting the customers who are at risk of churning, the focus of our analysis should then be to detect as many churners as possible, hence it is important to achieve high recall [24]. Now to decide is achieving high precision important for our scenario let’s ask a question to yourself: if we classify some customers as churners, whereas they have not churned in the near future, will the organization be at a risk if they apply strategic plans to the wrong customers, so they stay with the organization instead of leaving? Yes, as the organization will lose a significant amount of money on promotions/offers on extra customers. Also, the organization could lose customers if they might get frustrated due to receiving wrong offers. Therefore, achieving high precision is important for our scenario. It would be harmful to classify a customer as a churner who is actually not churning, and it would be also harmful to miss a customer who was considering to churn[24]. So, we will try to improve both the recall and the precision.

The F1-score is the harmonic mean of precision and recall. It is suitable for the situations where precision and recall both are equally important. In our case both the recall and precision are important, so instead of improving both, we might just strive for a high F1-score[24].

1. Algorithmic Parameters:

The parameters I have used in my validation technique is shown in (Fig. 1). I have not used any parameters for my supervised learning model.

I have used 10 as the number of splits parameter in the KFold-method. For the cross-validation score, I have used a few parameters: estimator (the model I have built to fit the data), X (data I want to fit), y (target variable I want to predict), scoring (the metric applied to the model/estimator), and cv (number of folds). For scoring, I have used F1-score, recall and precision to compare all of them (however, we are only interested in improving F1-score).

In the GA, I have used several parameters to adjust the model. These include features list, chromosome size, population size, number of generations, and optimal score. I have fixed all these parameters at the start of GA to ensure they are consistent and controls the algorithm well. The optimal solution is used to stop the algorithm ones the best fitness score is reached to prevent overfitting. Moreover, I have used a crossover rate of 0.6 and a mutation rate of 0.08 to control the number of times crossover and mutation will occur over generations.

1. Sampling Technique:

I have used SMOTEENN as my sampling technique because it contains the advantages of both: over-sampling and under-sampling approaches[10]. It is the best suitable sampling technique for our churn prediction model, as it overcomes the disadvantages of over-sampling and under-sampling approaches by removing noise and filtering the irrelevant samples[25].

**Experimental Results and Analysis**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Precision** | **Recall** | **F1 -Score** |
| **10 Folds without FS**  **(Using Imbalanced Dataset)** | 29.784% | 81.317% | 40.05% |
| **10 Folds with FS**  **(Using Imbalanced Dataset)** | 36.623% | 32.278% | 55.097% |
| **10 Folds without FS**  **(Using Balanced Dataset)** | 28.981% | 69.277% | 43.020% |
| **10 Folds with FS**  **(Using Balanced Dataset)** | 28.874% | 65.102% | 44.211% |

Tab.1. Evaluation measures for Naïve Bayes.

I have used Python to code my customer churn prediction algorithm. As stated before, our focus is on improving the F1-score. Using (Tab. 1), first let’s compare the F1-scores for imbalanced and balanced datasets separately (comparing only the scores achieved for the model with FS and without FS) and then compare the two best ones among them.

The F1-score obtained for the Naïve Bayes model without feature selection on the imbalance dataset is approximately 15% less than the one obtained with feature selection. Hence feature selection has improved the Naïve Bayes model on imbalanced dataset significantly.

The F1-score obtained for the Naïve Bayes model without feature selection on the balanced dataset is approximately 1% less than the one obtained with feature selection. Hence feature selection has improved the Naïve Bayes model on balanced dataset slightly.

Now let’s compare both the scores for imbalanced and balanced datasets. The F1-score obtained for the Naïve Bayes model on the imbalance dataset is approximately 11% greater than the one obtained on the balanced dataset. It is because our imbalanced dataset contains 72% of negative samples, meaning that True Negatives will be very high making False Positive rate to either decrease or increase very slowly[26]. Therefore, the recall is very low as the classifier have detected a larger number of positives as negative, precision is very high due to low number of False Positives, and hence the F1-score is very high due to high precision for imbalanced dataset[26],[27].

There were six customer churn problems discussed above. First was to achieve an accurate churn prediction performance is a big challenge, second was the imbalance nature of the customer churn dataset, third was to choose appropriate algorithmic parameters, fourth was to select a suitable supervised learning model, fifth was to choose the best optimization algorithm and the sixth was to select suitable evaluation metrics. I have already discussed above the solutions to the last four problems. To solve the second problem, using (Tab. 1) I have selected the model with balanced dataset as it is trained on the sampled instances: same number of samples in both the classes. The above results in (Tab. 1) are very effective in solving the first problem. The Naïve Bayes model (with Genetic Algorithm) trained on the subset of relevant features has the maximum accuracy (F1-score) as compared to the other model (without FS). For all the customers who actually have churned, 65% we correctly predicted as churners.

**Conclusion**

The telecommunication industry is encountering challenges in customer retention, necessitating the deployment of a churn prediction model to forecast customer attrition and take relevant actions to attract churning customers. There are six customer churn issues that I have highlight in my report. The first challenge was achieving an accurate churn prediction performance, the second was the imbalance nature of the customer churn dataset, the third was selecting appropriate algorithmic parameters, the fourth was selecting a suitable supervised learning model, the fifth was selecting the best optimization algorithm, and the sixth was selecting suitable evaluation metrics.

To classify churn customers, I have used Naïve Bayes model. Genetic Algorithm is the optimization algorithm I have used in this study. I have used K-fold Cross-Validation technique for my model. I have also included a hybrid sampling technique called SMOTEENN to balance the labels in my dataset. There were three evaluation metrics I have used: F1-score, precision, and recall, to evaluate my model.

I have first performed data cleaning and data-preprocessing to the original dataset. Then I have built the Naïve Bayes model and evaluated its performance without feature selection. After that, I have performed feature selection using GA and then built the Naïve Bayes model and evaluated its performance. I have then balanced the dataset and repeated the above 2nd and 3rd steps. The results showed that the Naïve Bayes model used with the Genetic Algorithm, performed well on the balanced churn dataset, as it has the maximum accuracy.

Resulting from my analysis, I would like to recommend that the most important thing which should be considered while implementing the customer churn prediction system is to experiment multiple approaches and select the one which yields the highest accuracy for your model. Focus mainly on the tuning parameters as it is used to control the whole algorithm.

**References**

[1] PeerJ Computer Science. An ensemble based approach using a combination of clustering and classification algorithms to enhance customer churn prediction in telecom industry, peerj.com. Available: <https://peerj.com/articles/cs-854/>

[2] IOPSCIENCE. Sequential Feature Selection in Customer Churn Prediction Based on Naive Bayes, iopscience.iop.org. Available: <https://iopscience.iop.org/article/10.1088/1757-899X/879/1/012090/meta>

[3] Cognizant. Applied Artificial Intelligence, cognizant.com. Available: <https://www.cognizant.com/us/en/glossary/applied-ai>

[4] Stanford Encyclopedia of Philosophy. Artificial Intelligence, plato.stanford.edu. Available: <https://plato.stanford.edu/entries/artificial-intelligence/>

[5] Analytics Vidhya. Churn Prediction- Commercial use of Data Science, analyticsvidhya.com. Available: <https://www.analyticsvidhya.com/blog/2021/08/churn-prediction-commercial-use-of-data-science/>

[6] towards datascience. Dealing with Imbalanced Data, towardsdatascience.com. Available: <https://towardsdatascience.com/methods-for-dealing-with-imbalanced-data-5b761be45a18>

[7] Stack Exchange. Macro- or micro-average for imbalanced class problems, datascience.stackexchange.com. Available: <https://datascience.stackexchange.com/questions/36862/macro-or-micro-average-for-imbalanced-class-problems>

[8] Statista. Customer churn rate in the United States in 2020, by industry, statista.com. Available: <https://www.statista.com/statistics/816735/customer-churn-rate-by-industry-us/>

[9] qualtrics. What is customer churn? How to measure & prevent it, qualtrics.com. Available: <https://www.qualtrics.com/uk/experience-management/customer/customer-churn/>

[10] IEEE Xplore. Effective ML Techniques to Predict Customer Churn, ieeexplore.ieee.org. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9544785>

[11] ScienceDirect. Customer churn prediction in telecommunications, sciencedirect.com. Available: <https://www.sciencedirect.com/science/article/pii/S0957417411011353>

[12] Society for Industrial and Applied Mathematics. Predicting customer churn in mobile networks through analysis of social groups, epubs.siam.org. Available: <https://epubs.siam.org/doi/epdf/10.1137/1.9781611972801.64>

[13] SAGE Journals. Impact of Proactive Postsales Service and Cross-Selling Activities on Customer Churn and Service Calls, journals.sagepub.com. Available: <https://journals.sagepub.com/doi/full/10.1177/1094670519883347>

[14] ScienceDirect. Incorporating textual information in customer churn prediction models based on a convolutional neural network, sciencedirect.com. Available: <https://www.sciencedirect.com/science/article/pii/S0169207019301499>

[15] ScienceDirect. Leveraging unstructured call log data for customer churn prediction, sciencedirect.com. Available: <https://www.sciencedirect.com/science/article/pii/S0950705120307152>

[16] IEEE Xplore. An Enhanced Bank Customers Churn Prediction Model Using A Hybrid Genetic Algorithm And K-Means Filter And Artificial Neural Network, ieeexplore.ieee.org. Available: <https://ieeexplore.ieee.org/abstract/document/9428805>

[17] JOSCEX. Improved Accuracy of Naive Bayes Classifier for Determination of Customer Churn Uses SMOTE and Genetic Algorithms, shmpublisher.com. Available: <https://shmpublisher.com/index.php/joscex/article/view/5>

[18] Springer Link. Anovel HEOMGA Approach for Class Imbalance Problem in the Application of Customer Churn Prediction, link.springer.com. Available: <https://link.springer.com/article/10.1007/s42979-021-00850-y>

[19] medium. Feature Selection using Genetic Algorithm in Python, radhajayaraman11.medium.com. Available: <https://radhajayaraman11.medium.com/feature-selection-using-genetic-algorithm-2f915d1349b0>

[20] Towards Data Science. Choosing the right Encoding method-Label vs OneHot Encoder, towardsdatascience.com. Available: <https://towardsdatascience.com/choosing-the-right-encoding-method-label-vs-onehot-encoder-a4434493149b>

[21] University of York Canvas. 5.2 Lesson 1: Genetic Algorithms (GA's), onlinestudy.york.ac.uk. Available: <https://onlinestudy.york.ac.uk/courses/918/pages/5-dot-2-lesson-1-genetic-algorithms-gas?module_item_id=63334>

[22] simplilearn. Understanding Naive Bayes Classifier, simplilearn.com. Available: <https://www.simplilearn.com/tutorials/machine-learning-tutorial/naive-bayes-classifier>

[23] Analytics Vidhya. Precision vs. Recall – An Intuitive Guide for Every Machine Learning Person, analyticsvidhya.com. Available: <https://www.analyticsvidhya.com/blog/2020/09/precision-recall-machine-learning/>

[24] Jeremy Jordan. Evaluating a machine learning model, jeremyjordan.me. Available: <https://www.jeremyjordan.me/evaluating-a-machine-learning-model/>

[25] ScienceDirect. Class Imbalance Problem, sciencedirect.com. Available: <https://www.sciencedirect.com/topics/computer-science/class-imbalance-problem>

[26] Towards Data Science. What metrics should be used for evaluating a model on an imbalanced data set, towardsdatascience.com. Available: <https://towardsdatascience.com/what-metrics-should-we-use-on-imbalanced-data-set-precision-recall-roc-e2e79252aeba>

[27] medium. Performance metrics for evaluating a model on an imbalanced data set, radhajayaraman11.medium.com. Available: <https://medium.com/datasciencestory/performance-metrics-for-evaluating-a-model-on-an-imbalanced-data-set-1feeab6c36fe>

[28] PLURALSIGHT. Validating Machine Learning Models with scikit-learn, pluralsight.com. Available: <https://www.pluralsight.com/guides/validating-machine-learning-models-scikit-learn>