Lawrence, Sean

4/30/2023

CECS 550

CECS 550 Final Project Report

Prediction Model Evaluation

Companies show promotional ads to customers to persuade them to buy their products, but they can only spend a limited amount of money on these ads. Therefore, they must spend it on customers who are more likely to buy from them again within a certain period of time, to increase their return on investment (ROI) and reduce promotional costs. So, they would use prediction models to predict which customers are only one-time buyers, and which are repeat buyers who are more likely to be loyal customers, and hence could be marketed to so they do become loyal customers. As with all prediction models, we first need to train it on training data to make sure it can predict labels accurately for labeled customers, so that it can perform well with new data.

We have used different classification models, such as:

1. Gaussian Naïve Bayes (Gaussian) classifier,
2. Logistic Regression (Log) classifier,
3. Linear Discriminant Analysis (LDF) classifier,
4. K Nearest Neighbors (KNN) classifier,
5. Parzen Windows classifier,
6. XGBoost classifier,
7. Neural Network classifier.

We have measured their performance with these metrics:

1. Accuracy of the model, which is how many of the total number of predictions are correct [1].
2. Precision, which is how many of the predicted positive class values are truly positive. This measures a classifier’s exactness [1].
3. Recall, which is how many of the truly positive class values in the test data are predicted as positive. This measures a classifier’s completeness [1].
4. F1 score, which is the harmonic mean of precision and recall [2]. It tells how well a classifier can detect positive values, which is more significant than detecting negative values.
5. ROC-AUC score, which is based on the True Positive Rate (TPR), or how many positive data points are classified as such (same as recall), and the False Positive Rate (FPR), of how many negative data points are mistakenly classified as positive. These rates are plotted on a single graph called the ROC curve, and the area under this curve is the ROC-AUC score, which represents the probability that a randomly chosen positive example ranks above (has a higher probability of being positive than negative) a randomly chosen negative example. High ROC scores mean the probability of a randomly chosen positive example is indeed positive, and that our algorithm does a good job at ranking the test data as negative on one end of a scale and positive on the other [3].

The data has a label imbalance of about 15 non-repeat buyers to 1 repeat buyer, which initially threw off the classifier to mostly classify every user as a non-repeat buyer with too little effort in finding repeat buyers. We fixed this issue by oversampling the minority class using SMOTE, so that repeat buyer data can have a fairer share in the data that gets classified.

Fig. 1 below shows the performance metrics of our classifiers when applied to the oversampled data. It shows that the Parzen Windows classifier has the highest and most consistent metrics amongst all the methods and is therefore the best classifier for this data.

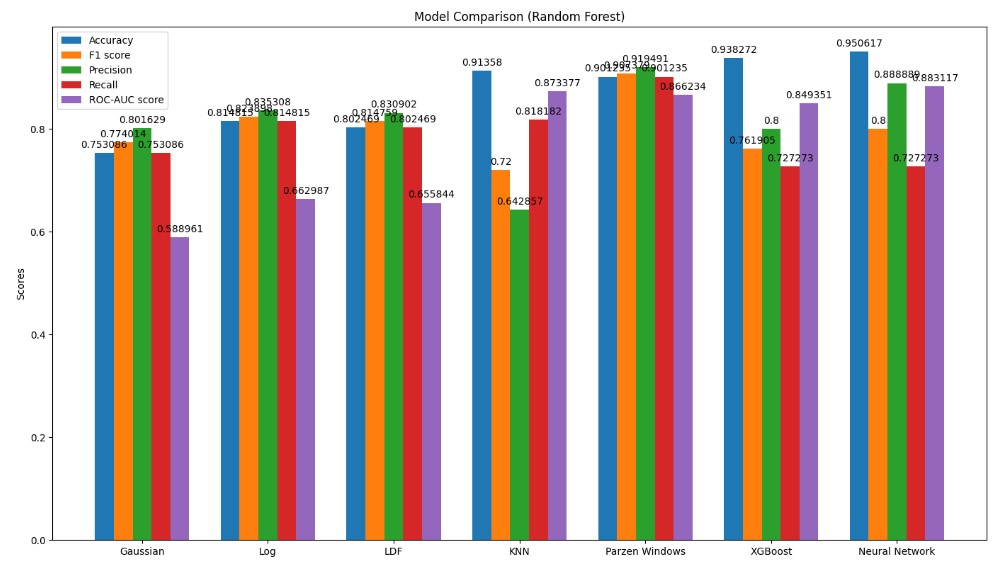


Fig. 1: Model Performance Comparison, with a Random Forest Classifier

Works Cited

1. J. Brownlee, “Classification accuracy is not enough: More performance measures you can use,” *Machine Learning Mastery*, 20-Jun-2019. [Online]. Available: <https://machinelearningmastery.com/classification-accuracy-is-not-enough-more-performance-measures-you-can-use/>. [Accessed: 30-Apr-2023].
2. D. S. Batista, “Evaluation Metrics, ROC-Curves and imbalanced datasets,” *Evaluation Metrics, ROC-curves and imbalanced datasets*, 19-Aug-2018. [Online]. Available: <https://www.davidsbatista.net/blog/2018/08/19/NLP_Metrics/>. [Accessed: 30-Apr-2023].
3. A. Bajaj, “What does your classification metric tell about your data?,” *Medium*, 25-Dec-2019. [Online]. Available: <https://towardsdatascience.com/what-does-your-classification-metric-tell-about-your-data-4a8f35408a8b#:~:text=A%20low%20recall%20score%20(%3C0.5,order%20to%20curb%20FP%2FFN>. [Accessed: 30-Apr-2023].