MIM JOURNAL 2023 illuminate your work
Made in Millersville Journal | www.mimjournal.com

Combining Unsupervised and Supervised Learning for Credit Card Fraud Detection

Briar Sauble

Abstract

Due to the rapid growth of e-commerce, finding a proper method to detect credit card fraud has become more important than ever. In machine learning, the two main methods of detecting credit card fraud are through supervised and unsupervised learning. Supervised learning methods are designed to find patterns of known credit card transactions but are flawed due to missing novel patterns of fraud alongside missing patterns on which they haven't yet been trained. Unsupervised learning, in contrast, can be used to find hidden patterns in unclassified data but is dependent on human intervention to understand how each feature within a dataset creates its patterns. In this project, I developed a way to improve upon credit card fraud detection by applying elements of a supervised learning model to an unsupervised model. Using the decision tree classified learning algorithm, I found the importance of different components of a credit card transaction in relation to fraud. This was used to apply to a newly designed weighted k-means algorithm that uses the importance of each feature as a weight, creating a hybrid model. I tested my method by using a dataset consisting of synthetic credit card transactions containing over one million transactions. By using this algorithm, the chance to accurately classify a transaction as fraud is improved to 98.51%, compared to the normal k-means clustering model accuracy of 87.92% – suggesting value in using a hybrid model over one singular method alone.

Credit card use has exponentially increased in the modern age due to its ease of use in transactions in digital and physical retail spaces. However, this corresponding increase has subsequently led to similarly large losses due to credit card fraud. Out of 320k registered reports on fraud, as noted by the Federal Trade Commission (FTC), 133k of these were related to credit card fraud, as noted in previous fraud research (Dornadula & Geetha, 2009). The consequences of immense loss because of this issue have increased demand for efficient ways of detecting credit card fraud.

Machine learning algorithms provide a way to automate the process of data

analysis to calculate the likelihood of a given transaction being credit card fraud. Two primary paradigms exist: unsupervised and supervised learning. Unsupervised learning is used on unlabeled datasets, clustering information based on the patterns it finds without any information on which transactions are fraudulent or not. Supervised learning, in contrast, uses labeled datasets to train models that classify data based on its attributes to predict outcomes accurately. Both supervised learning and unsupervised learning, however, hold significant limitations. Unsupervised learning does not define the background behind patterns and requires intervention to recognize them, while supervised learning depends on the existence of labeled data as well as the accuracy of the data it is trained on, which may be limited by the synthetic nature of the data acquired (Casolino et al., 2019). To attempt to get past both the flaws of unsupervised and supervised learning, this research combined the models of both unsupervised and supervised learning, hypothesizing that this would improve results over use of a singular type of learning alone.

Methodology

To simulate a set of credit card transactions, a previously generated synthetic credit card fraud dataset provided on Kaggle was utilized holding the information of 1000 credit card users performing transactions among 800 merchants (Shenoy, 2020). Each element of the dataset holds information about a single transaction, such as the credit card number used and longitude and latitude values for the user and the merchant. To process this data and to perform the various machine learning algorithms required to generate the model, the scikit-learn library (Pedregosa et al., 2011) for the Python programming language was utilized. This library simplifies the process of machine learning by providing pre-written code for creating programs involving predictive data analysis. Creating the hybrid model is a multi-step process utilizing features of a supervised learning algorithm to calculate feature importance statistics as weights to adjust the results of an unsupervised algorithm, determined by the relevance of a component of the credit card transaction. In order, the steps to create the hybrid model were to create a decision tree, acquire a feature importance value for each attribute within the decision tree, and to apply weights to a k-means clustering algorithm to get a final hybrid model.

A decision tree is a supervised learning algorithm that creates a hierarchy of leaves that split off to designate an order of decisions, with its initial root node and its decisions having the highest importance in deciding a given element (Aggarwal, 2015). In this case, a decision tree was used to organize the data by the level of importance each element of a given transaction has in deciding credit card fraud. I wished to capture this level of importance as a value to properly quantify the relevance of each part of the transaction. From the decision tree, I was able to gather the feature importance, which defined from a scale from zero to one how relevant a given feature of a transaction would be. Out of 22 attributes, only nine attributes held enough significance based on this metric to continue to be used for future classification.

To verify the accuracy of the nine attributes acquired from the feature importance metric of the decision tree, these were applied to an unsupervised learning algorithm. This algorithm, known as the k-means clustering algorithm, organizes each transaction into clusters of data, the count of clusters being decided by the value provided to the parameter k (Aggarwal,

2015). To verify the worth of the feature importance metric, the k-means algorithm was modified to use the importance for each feature to modify how the algorithm would determine which cluster a transaction would belong to. It was hypothesized that attributes that are given a lesser importance would have less of a deciding factor in which cluster a transaction will be assigned to. Instead, those with greater importance would have a larger range of values that would make more of an impact in deciding where the transaction would go. The parameter k, in this case, is set to two—making for two clusters of data, holding either "non-fraud" or "fraud" transactions. As fraudulent credit card transactions can be considered as outliers due to there being a low quantity of them in comparison to non-fraudulent transactions, it was also hypothesized that most transactions would be within the "non-fraud" cluster. Thus, once generated, the cluster holding the greatest number of transactions was determined to be non-fraudulent, and accuracy of the weighted k-means algorithm was then obtained based upon if the clusters accurately held the corresponding type of transaction—either fraudulent or non-fraudulent—within them.

Results

On their own, the initial decision tree classifier that was used to calculate the feature importance statistic through supervised learning acquired an accuracy of 95.78%, and the k-means classifier without any applied weights obtained an accuracy of 87.92%. In contrast, the weighted k-means algorithm obtained an accuracy of 98.51%. A higher accuracy rate from the hybrid model suggests that using it will make it more likely to accurately determine whether something is fraudulent or not. As these are results based on synthetic credit card transactions, it's uncertain how much the synthetic nature of the data impacts these results. Based on these initial observations alone, however, it appears that use of a hybrid model has improved the ability to detect fraudulent transactions.

Conclusion

An improved model for detecting credit card fraud was achieved by using a hybrid model consisting of pre-existing machine learning methods, which provided more accurate results over using a singular form of machine learning on its own. Further research will be needed to see similar levels of improved accuracy are found within other synthetic datasets. The foundation for future research has been made for expanded hybrid models which can be used to future improve the likelihood for detecting credit card fraud.

References

Aggarwal, C. C. (2015*). Data mining: The textbook* (1st ed.). Springer.

Casalino, G., Castellano, G., & Mencar, C. (2019). Credit card fraud detection by dynamic incremental semi-supervised fuzzy clustering. *Proceedings of the 2019 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology* (EUSFLAT 2019). https://doi.org/10.2991/eusflat-19.2019.30

Dornadula, V. N., & Geetha, S. (2019). Credit card fraud detection using machine learning algorithms. *Procedia Computer Science*, 165, 631–641. https://doi.org/10.1016/j.procs.2020.01.057

Shenoy, K. (2020). Credit card transactions fraud detection dataset. *Kaggle*. Retrieved March 28, 2023, from https://www.kaggle.com/datasets/kartik2112/fraud-detection

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.

Recommended Citation
Sauble, B. (2023). Combining unsupervised and supervised learning for credit card fraud detection. *Made in Millersville Journal*, 2023. Retrieved from https://www.mimjournal.com/computer-science-2023