Unsupervised learning concluding project

Part 1, project proposal

Erik Lager 209522234 and Dor Polovinchik 200973568

The paper we chose is “Data-Driven Approach for Credit Card Fraud Detection with Autoencoder and One-Class Classification Techniques” by Abdoul-Fatao Ouedraogo, C´edric Heuchenne, Quoc-Thong Nguyen, and Hien Tran

Paper summary:

The paper deals with the problem of credit card fraud. This is a problem which concerns everybody since Fraudulent credit card transactions pose a significant threat to both the cardholder and the financial institution that issued the card. Such transactions result in financial losses and compromise sensitive personal and financial information, which can lead to identity theft and other crimes. Additionally, fraudulent activities can damage the reputation of financial institutions, leading to decreased consumer trust and stricter security measures. The traditional methods for the problem of credit card fraud detection include statistical methods such as rule-based systems, decision trees, and logistic regression.

The paper presents a data-driven approach for the problem using autoencoder and one-class classification techniques. The authors aim to improve upon the traditional fraud detection methods. As far as we saw, this approach was not too popular up until this paper and the results they got were impressive compared to the existing ones.

The approach involves using an autoencoder to reduce the dimensionality of the transaction data and extract relevant features, which are then used as input to a one-class classification model. The one-class classification model is trained to identify normal transactions, and any instances that fall outside of the learned normal behavior are flagged as potential fraud. We can see the workflow in figure 1:

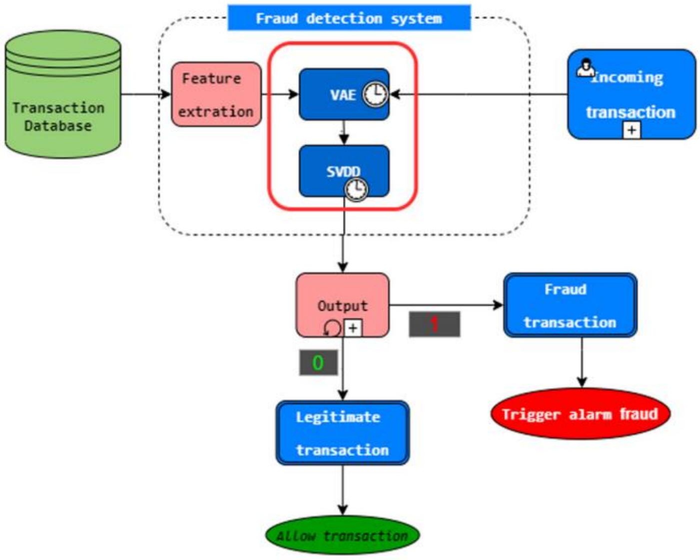
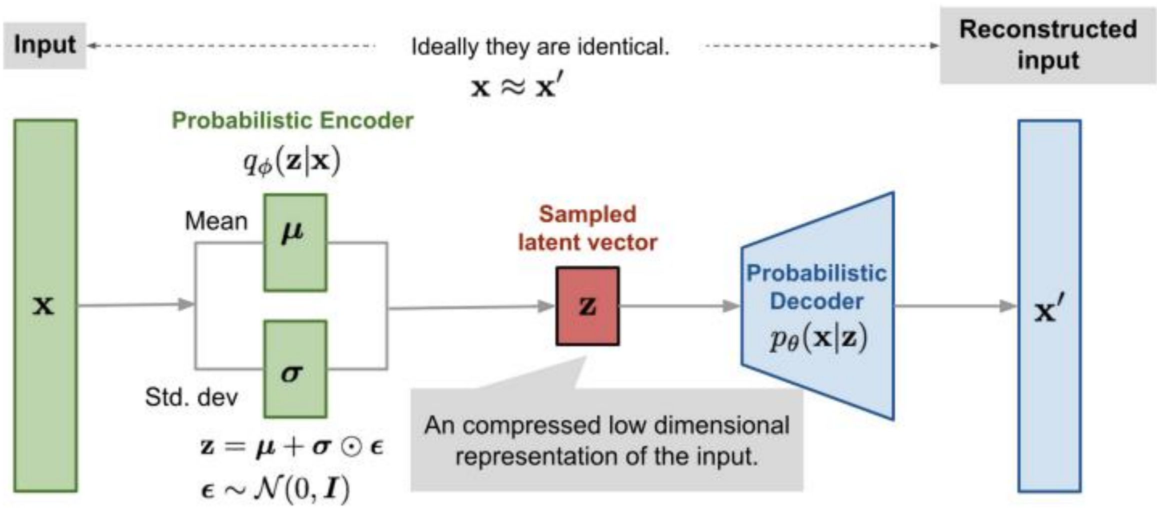


Fig. 1. Approach to the proposed solution

The results of the experiments indicate that the proposed approach is effective in detecting credit card fraud and outperforms traditional methods in terms of accuracy. The authors also report that the proposed approach can handle high-dimensional data and identify complex and non-linear relationships in the data, which can be useful in detecting more sophisticated forms of fraud. We can see the illustration of the variational auto-encoder in figure 2:



Overall, the paper provides a promising solution for credit card fraud detection, leveraging the strengths of autoencoder and one-class classification techniques to improve accuracy and robustness. However, as with any data-driven approach, the performance is dependent on the quality and representativeness of the training data, as well as the choice of hyperparameters and model architecture.

Since we are familiar with the field of fraud detection, we know that the main problem comes from lack of data and the fact we don't get feedback when we decline a "good" transaction. Therefore, we want to use this approach to create a larger training set for different approaches and see if the mixture between the dataset with new data with different models (which have also been tested in recent years) will give better results.

We intend to split the data in two (train and test) and use the training set to get the generated fraud examples.

The upsides and downsides of the approach are like the ones that auto encoders and one-class classification have. The upsides include:

1) Improved accuracy compared to traditional methods.

2) The ability to identify complex and non-linear relationships in the data.

3) The ability to handle high-dimensional and unstructured data.

The downsides of such approaches might be:

1) The potential for overfitting and poor generalization to unseen data

2) The computational cost and complexity of training deep neural networks

3) The need for large amounts of labeled data to train the model effectively

4) One must be careful when testing on generating Data. Since the new examples should only be drawn from a model which only saw the training set.

A more in-depth explanation of what we intend to do:

The field of fraud detection suffers greatly from the problem of imbalanced data. Therefore, an approach we thought could be interesting is using generative functions to create more examples of fraudulent transactions.

Instead of using the full approach from the paper, we will mainly focus on the VAE, and use it to create fraudulent data.

After creating more data, we will use some different approaches on the it without the additional fraud examples and on the combined data. We expect to see higher accuracy for the results which were trained on the combined data.