

Context

It is important that credit card companies are able to recognize fraudulent credit card transactions so that customers are not charged for items that they did not purchase. Credit card Fraud detect system using Convolutional Neural Network
We can consider the project as a Recommendation System

Content

Using tabular data from Kaggle <https://www.kaggle.com/mlg-ulb/creditcardfraud> motivated me to create a classification system for transactions that could detect real or fake transactions. Most of the transaction must be real, only a small amount could be fraud. I can observe that the data is imbalanced. Therefore it would be a factor to take into account.

My dataset contains only numerical input variables which are the result of a PCA transformation. Unfortunately, due to confidentiality issues, they cannot provide the original features and more background information about the data. Features V1, V2, ... V28 are the principal components obtained with PCA. I'll use only one dataset

The only features which have not been transformed with PCA are

- 1)'Time' and 'Amount'. Feature 'Time' contains the seconds elapsed between each transaction and the first transaction in the dataset.
- 2)The feature 'Amount' is the transaction Amount, this feature can be used for example-dependant cost-sensitive learning.
- 3)Feature 'Class' is the response variable and it takes value 1 in case of fraud and 0 otherwise.

Given the class imbalance ratio, I recommend measuring the accuracy using the Area Under the Precision-Recall Curve (AUPRC). Confusion matrix accuracy is not meaningful for unbalanced classification.