**Name** : Sheshank Bandakunta

**Title** : Credit card fraud detection analysis

**Abstract** : The datasets contains transactions made by credit cards in September 2013 by European cardholders. This dataset presents transactions that occurred in two days, where we have 492 frauds out of 284,807 transactions. The dataset is highly unbalanced, the positive class (frauds) account for 0.172% of all transactions. Since the dataset is unbalanced I will use under sampling to make the data balanced, Then I will train my model with the balanced data and then I will use the original data set as testing dataset on my model. I will be using 3 or 4 algorithms as my models depending on time. I will show visually representation of the dataset before and after balancing as a part of data exploration and then after modeling I will show the accuracy of the algorithms on this dataset.

**Algorithms** – tenserflow, Decision tree, Logistic regression, SVM.

**Data Sources** : <https://www.kaggle.com/dalpozz/creditcardfraud/data>

**References** : <https://www.analyticsvidhya.com/learning-paths-data-science-business-analytics-business-intelligence-big-data/learning-path-data-science-python/>

<https://www.kaggle.com/dalpozz/creditcardfraud>

<https://stackoverflow.com/questions/29204005/how-to-perform-under-sampling-in-scikit-learn>

<https://stackoverflow.com/questions/34831676/how-to-perform-undersampling-the-right-way-with-python-scikit-learn>

<https://www.kaggle.com/joparga3/in-depth-skewed-data-classif-93-recall-acc-now>

<https://gallery.cortanaintelligence.com/Experiment/Evaluating-and-Parameter-Tuning-a-Decision-Tree-Model-1>

<https://www.tensorflow.org/get_started/graph_viz>