**Machine Learning – Technical test**

**Objective:**

The purpose of this technical test is to test your ability to code, present analysis and write scalable solution in code repository which can be shared with us. In this role, you would be expected to not only model the algorithm but also to productionize it in an efficient and scalable way and present insights to stakeholders.

Primary objectives of this test are as follows:

1. Full Functioning SDK for the model, Please refer to the packaging from <https://the-hitchhikers-guide-to-packaging.readthedocs.io/en/latest/quickstart.html>
2. Version Controlled (GIT) code in Github or Bitbucket.

Predictions using the SDK if any.

**Tasks:**

Given data was extracted from a marketing platform comprising of anonymous 10K records. This training dataset is a sample from a much larger training set for Machine Learning model training. To solve this technical test, you are free to use any Python packages from available ML toolkits such as Sklearn, H2O, Tensorflow etc. or Java, Scala.

This is unlabelled/unclassified data and your task is to provide an approach to solve a particular problem using a combination of clustering and classification/anomaly detection techniques. Data extract comprises of 3 main categories of features:

1. Account Features
2. Personal ID Features
3. Claims Activity

From all the features on training data, you are expected to perform the clustering and subsequently classification/anomaly detection to visualize the consumer group and classify consumers that are bad i.e. consumers that are doing fraud transactions. Please bear in mind that this is not a statistical analysis test and hence you are expected to present a working model using SDK. Summary of technical test tasks are as follows:

1. Clustering
2. Classification
3. Model metrics and performance measures
4. What, Why and How of the work presented

**Summary:**

Please submit your completed test by 5.00pm Wednesday 1st December .Expected time frame to complete this work is 8 Hours. Please send us the completed code repository with GIT history and if possible, presentation slides for the findings. Bonus points if you can provide a test coverage, writing good solid piece of code which is PEP compliant and with README.md.