**UNSUPERVISED LEARNING EXERCISE REPORT**

**Report submitted by: Priyank Namdev (21250124)**

* *Introduction:*

In unsupervised learning, a dataset is provided without labels, and a model learns useful properties of the structure of the dataset. We do not tell the model what it must learn, but allow it to find patterns and draw conclusions from the unlabelled data.

Unsupervised machine learning techniques uses ML algorithms to analyse and cluster the unlabelled data. Varying for the supervised learning techniques, algorithms discover the hidden patterns/grouping of data without any human input and then categorise these patterns into the pre-described clusters/groups based on the similarities and differences in the information. Thus, it makes an ideal choice for exploratory of big data where labelling is not possible for a human.

* *Algorithms:*

Clustering: - ML Algorithms cluster/group the data based on the similarities and differences once it is labelled. These algorithms are used to process the raw, unclassified data objects into groups by structures or patterns in the data. I am using LSTM layers to model my algorithm.

* *Methodology: (Using LSTM algorithm in Python)*
* Data: The data is 1149 text files containing 9 different fields
* Steps:
* Step 1: The data was first imported. Few of the features were removed from the data as a part of making the algorithm more effective with consistent ‘prediction’ results. Removed files has decoding issue.
* Step 2: Feature extraction is done using regular expressions. I took out the text between “== RESULTS” and “== ISSUES”.
* Step 3: Created a token named “awsome\_tokens”.
* Step 4: Initialising the list of tokens.
* Step 5: Taking out the max token sequence.
* Step 6: Initialising Predictor label.
* Step 7: Applied LSTM to my model.
  + Epochs: 10
  + Accuracy: 0.10
  + Time took: 3 hours.
* Learnings:
  + I need a better machine (no pun intended).
  + I should learn more about these Unsupervised learning as it is really fun to explore.
* Code Link: https://github.com/priyank001/Unsupervised-Learning---Text-Clus