

## **Case study NLP classifier**

The dataset contains 5573 emails where they are labeled as spam and ham, where 4825 are ham (non spam) and 747 spam emails. We need to build a NLP classifier that specially uses word2vec from Google. Divide the dataset into 80 and 20 percent and build 3 types of models

1. CBOW
2. Skipgram
3. Pretrained word2vec model from Google

Please do necessary pre-processing methods before building the NLP classification model.



Please note the following:

Dataset is added along with

Do the assignment in Jupyter notebook/colab

Add the .ipynb file to GitHub and share the link in Paatshala