**Text classification using logistic regression**

Using the given spam the dataset, the goal is to implement a classification model to predict if a given text is spam or not.

1. Apply the appropriate preprocessing to your text.
2. Use different vectorization techniques: bow, tfidf, word2vec and doc2vec.
3. Evaluate the performance of each model (confusion matrix, accuracy, etc)
4. Discuss results.
5. How can you improve the performance of your model?