

Project - Spam Detection

Design a **End to End NLP Project** on the data-set given below:

Download the dataset from this link. [CLICK HERE TO DOWNLOAD DATASET](#)

The detailed description about the dataset can be found here. [Dataset Description](#)

For your reference, kindly have a look at this jupyter notebook. [CLICK HERE](#)

If you are facing any difficulty, follow the steps mentioned below:

Step - 1 - Getting things ready for the project

- Import the required libraries
- Load the data

Step - 2 - Perform the Exploratory Data Analysis on the data set

Step - 3 - Preprocess the data

- Remove special characters
- Conversion to lowercase
- Tokenization
- Stemming or Lemmatization

Step - 4 - Data Preparation

- Train Test Split
- Identify the target variable
- Use Bag of Words to convert the Text data to Numerical Vectors

Step - 5 - Train the Model. Use the following algorithms:

- Logistic Regression
- Decision Trees
- SVM

Step - 6 - Evaluation. For each Model, evaluate using following metrics:

- Check accuracy score
- Build confusion metric
- Print Classification_report