**Classifying movies reviews**

**A potential client for EVS needs to classify reviews in a dataset stored in their records.**

**As a data science analyst, you have been selected to perform an analysis on the data that they have and create a model to automatically classify new records from now on.**

**The dataset is a CSV file which contains 2 columns and 1,000 rows.**

**Your boss requires a python (.py) file that includes:**

**-Pre-processing steps and methods.**

**-The feature engineering process (techniques used, code (comments)).**

**-The model selection process (candidates, selected (why?) and evaluated metrics).**

**Among the following techniques, you’ll need to cover:**

**-Importing proper libraries (5pt)**

**-Importing dataset (5pt)**

**-Stopwords (5pt)**

**-Normalization (Stemming/Lemmatization) (5pt)**

**-String manipulation (5pt)**

**-Tokenization (5pt)**

**-TextBlob features (at least one) (15pt)**

**-Spacy features (at least one) (15pt)**

**-Sparse Matrix (10pt)**

**-Modeling (Bag of words) (10pt)**

**-Performance metrics comparison (20pt)**