

**Q1. I developed a system that detects tweets about about different candidates in a general election. I have classified a gold-standard set of 200 tweets, 100 of which I identified to be about the election and 100 I classed as being about non-election things.**

**When I vary the similarity threshold of the system from 1 – 50 I get different numbers of correct and incorrect answers, that is different numbers of True Positives (TP), False Negative (FN), False Positives (FP)**

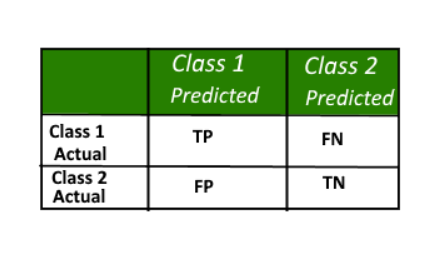
**and True Negatives (TN) tweets. For example, when my system correctly identifies the tweets as being about the election and it was indeed about the election, it’s a True Positive. When my system says that**

**the tweet is about the election and it is not, then I have got a False Positive.**

**Taking this data, can you compute the Precision and Recall for the system at each threshold and identify the threshold values at which it does best,according to the F1 measure?**

**Ans1.**

**Confusion Matrix:** Confusion matrix is defined as a table used to describe the performance of the model. It allows visualization of the algorithm’s performance. It consists both Predicted as well as Actual classes. Example is shown below:



**True Positive (TP):** A true positive is an outcome where the model correctly predicts the positive class. Observation is positive, and is predicted to be positive also.

**False Positive (FP):** A false positive is an outcome where the model incorrectly predicts the positive class.

Observation is negative, but is predicted positive.

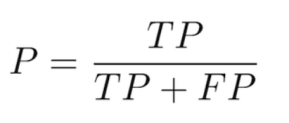
**True Negative (TN):** A true negative is an outcome where the model correctly predicts the negative class.

Observation is negative, and is predicted to be negative.

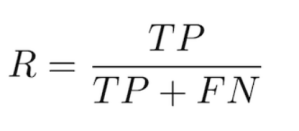
**False Negative (FN):** A false negative is an outcome where the model incorrectly predicts the negative class.

Observation is positive, but is predicted negative.

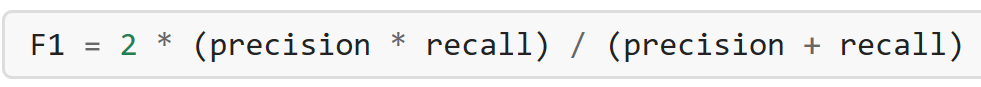
**Precision:** It defines what proportion of all the predictions that we made with the model are actually true. It is calculated as the ratio of correct positive predictions to the total predicted positive.



**Recall/True Positive Rate:** It is a measure that gives the fraction of all relevant documents that are successfully retrieved. It is calculated as the ratio of correct positive predictions to the total positive examples.

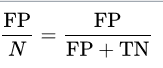


**F1 Measure:** It is a way of collaborating both Precision and Recall measures of any model with the help of Harmonic Mean of both of them. For uneven class distribution, where accuracy is not an apt measure, F1 measure works there. Higher the value of F1 measure, higher accuracy for the system (mostly). It is calculated as follows:

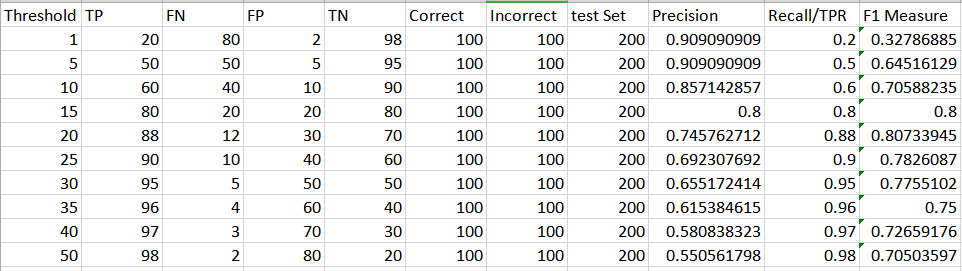


**False Positive Rate:** The false positive rate is an accuracy metric, calculated as the ratio between the number of negative events wrongly categorized as positive (false positives) and the total number of actual negative events regardless of classification.

The false positive rate infers to the expectancy of the false positive ratio.



Recall and Precision values for every threshold value is calculated. According to F1 measure, Threshold Number 20 and 15 does their best as compared to the rest threshold values.



=========================================================================================

**Q2. Now, can you plot the ROC for this data?**

**Ans2.**

**ROC Curve:** A receiver operating characteristic curve, ROC curve, is a technique widely used in machine learning to evaluate the classifier’s performance. It condenses the classifier’s performance graphically and allow us to compare the performances of various classifiers. ***[1]***

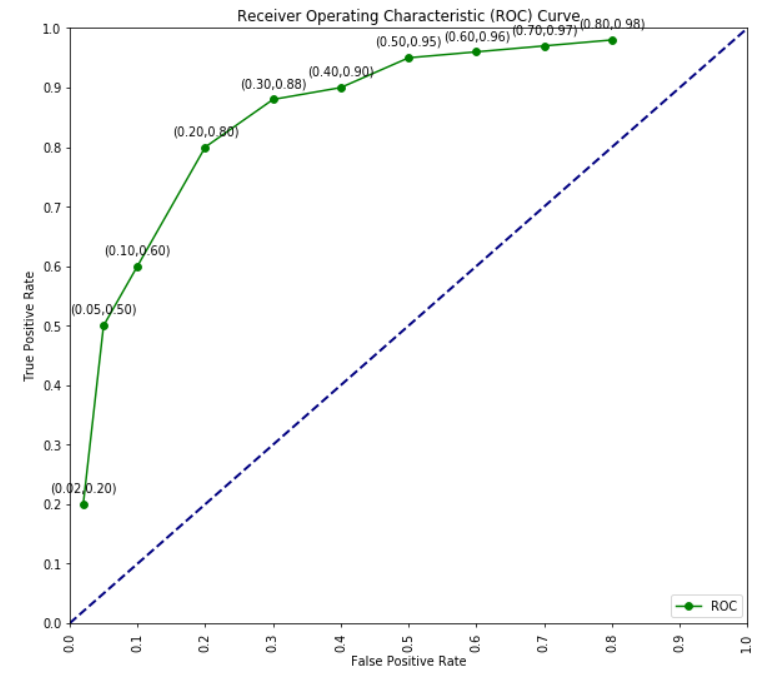
It is a graphical plot that depicts the ability of a binary classifier system as its similarity threshold is varied. It plots the True Positive Rate against the False Positive Rate for every threshold value between 0 and 1.

**Area under ROC Curve:**

To compare the classifiers, Area under Curve (AUC) is useful to summarize the performance of the classifiers in a single measure. Higher the area under curve, model is better at predicting 0s as 0 and 1s as 1.

Each point on the ROC curve shows a sensitivity/specificity (TPR,FPR) pair corresponding to a similarity threshold. AUC is a metric of how well a parameter can differentiate between two classes (election/non-election). ***[2]***

ROC Curve is shown below with equal scaling on both X and Y axis and pair of (TPR,FPR) are displayed on the graph :



=========================================================================================

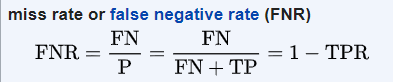
**Q3. Now, can you plot the DET curve for the same data?**

**Ans3.**

**False Negative Rate/Miss Detection Rate:**

The false negative rate is the proportion of the individuals with positive condition for which the test result is predicted negative. This rate is also known as miss rate. ***[3]***

The false negative rate is an accuracy metric, calculated as the ratio between the number of positive events wrongly categorized as negative (false negatives) and the total number of actual positive events regardless of classification.

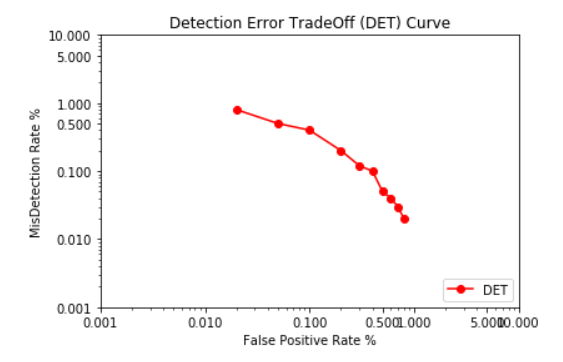


**DET Curve:** Detection Error Trade Off, DET curve focuses on error more and zoom in to the key parts of ROC curve using log axes.

DET curve, is a technique widely used in machine learning to evaluate the classifier’s performance in terms of error. It plots the Miss Detection Rate against the False Positive Rate with log axes.

DET curve is an evaluation metric in binary classification problem with two classes focus is on how it is incorrectly classifying both classes. It is a graphical plot of error rates for the binary classifier.

DET Curve is shown below for the Election-Non-Election tweets classifier with log scaling with False Negative Rate on Y -axis and False Positive rate on the X axis.



=========================================================================================

**REFERENCES**

[1] C.Chan, “What is a ROC Curve and How to Interpret It”, 2019 [Online]. Available:

<https://www.displayr.com/what-is-a-roc-curve-how-to-interpret-it/?utm_referrer=https%3A%2F%2Fwww.bing.com%2Fsearch%3Fq%3Dwhat%2Bis%2Broc%2Bcurve%26FORM%3DAWRE> . [Accessed Nov. 08, 2019]

[2] “MedCalc easy-to-use statistical software”, 2019 [Online]. Available: <https://www.medcalc.org/manual/roc-curves.php> . [Accessed Nov. 08, 2019]

[3] NCSS Statistical Software, “Binary Diagnostic Tests – Single Sample ”, 2019 [Online]. Available:

<https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Binary_Diagnostic_Tests-Single_Sample.pdf> . [Accessed Nov. 08, 2019]