**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**

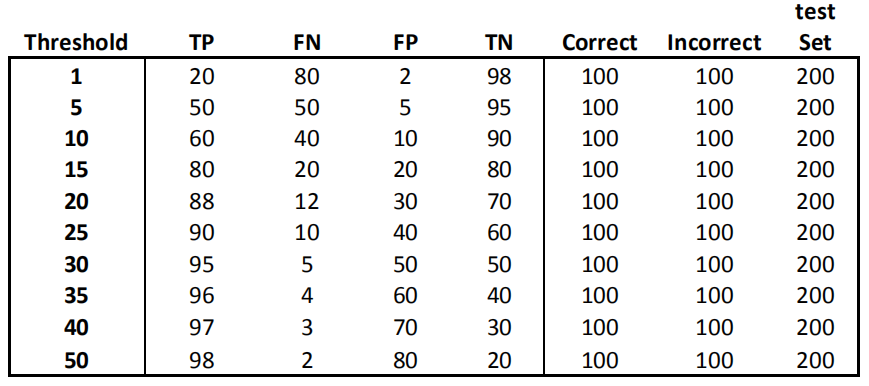
**numbers of True Positives (TP), False Negative (FN), False Positives (FP) and True Negarives (TN) tweets. For example, when my system**

**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?**



**Ans.**

Confusion Matrix: It is a quality metric for classification of machine learning problems where two or more categories can be generated. It is a table of 4 different predicted and real values combinations.

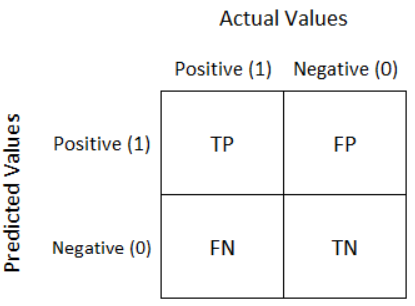


Fig: Confusion Matrix

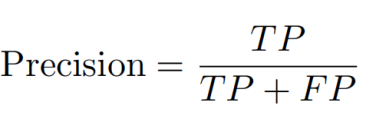
**True Positive(TP)**: It is an outcome where the model ***correctly*** predicts the positive class.

**True Negative(TN):** It isoutcome where the model ***correctly*** predicts the negative class.

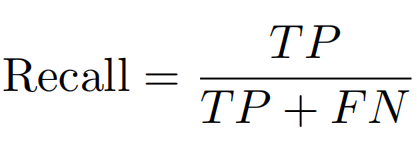
**False Positive(FP):** It isan outcome where the model ***incorrectly*** predicts the positive class.

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

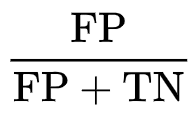
**Precision:** It is the fraction of retrieved documents that are relevant to the query.



**Recall:** Recall is the fraction of the relevant documents that are successfully retrieved. It is also called as Sensitivity.



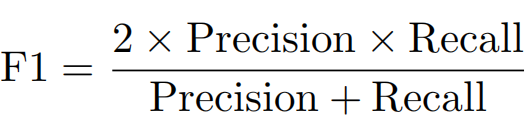
**False Positive Rate(FPR):** The false positive rate is calculated as the proportion between the number of falsely classified negative events as positive (false positives) and the overall number of actual negative events (irrespective of classification).



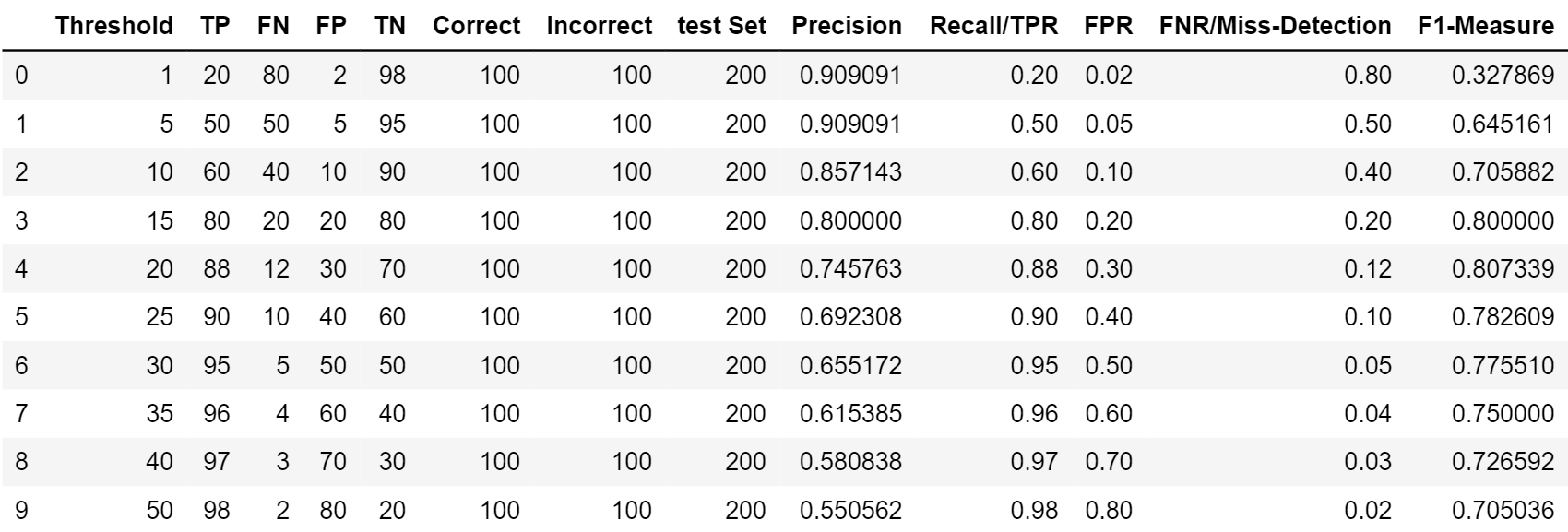
**False Negative Rate(FNR) or Miss-Detection:** The false negative rate is the proportion of the individuals with a known positive condition for which the test result is negative. This rate is sometimes called the miss rate.



**F1-Measure:** It is a measure of a test's accuracy that trades off precision against recall, for a given level of balance. It is the harmonic mean of the precision and recall, where an F1 score reaches its best value at 1 (perfect precision and recall) and worst at 0.



I have calculated all these parameter for all the thresholds that were given in the question. Below is the result after calculating all parameters.

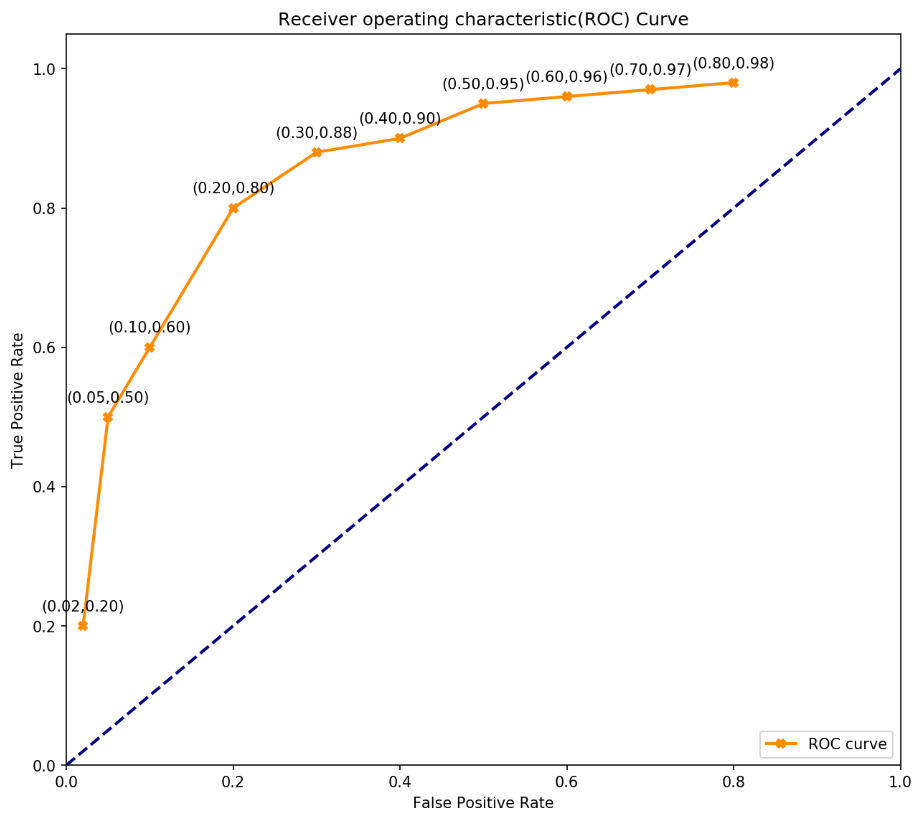


**Q2. Now, can you plot the ROC for this data?**

**Ans:**

ROC curve is a performance measurement for classification problem at various thresholds settings. ROC is a probability curve and AUC represents degree or measure of separability. It tells how much model is capable of distinguishing between classes. Higher the AUC, better the model is at predicting 0s as 0s and 1s as 1s. By analogy, Higher the AUC, better the model is at distinguishing between True and False Hypothesis[1].

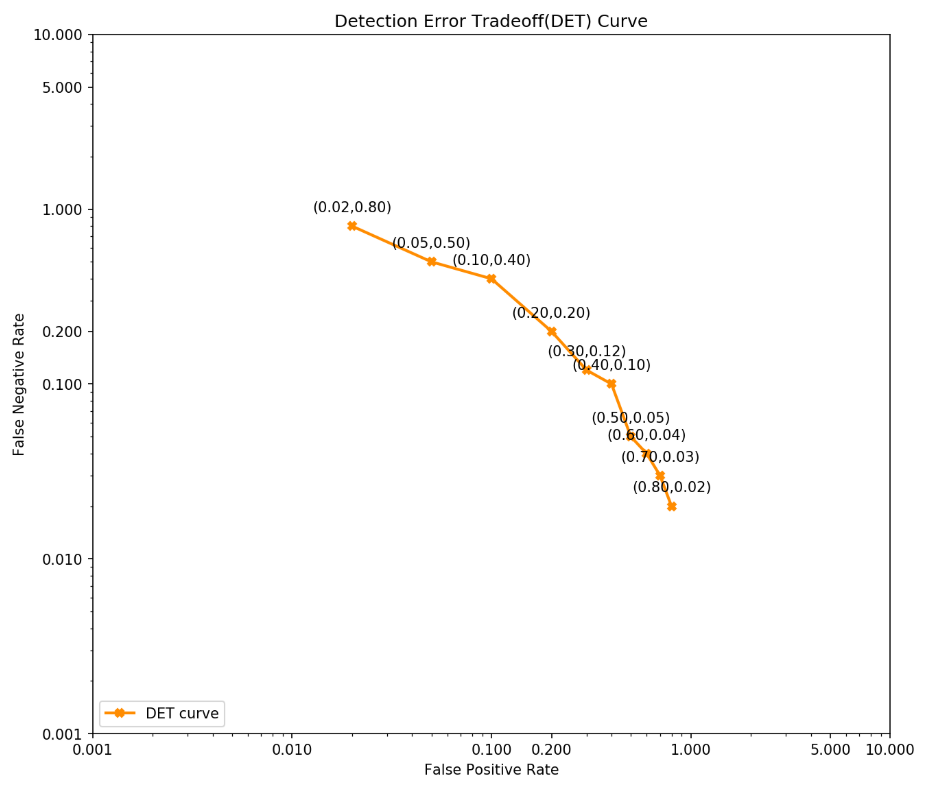
Below is the graph between False Positive Rate Vs True Positive Rate. More the Area under the curve more is the accuracy.



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

**Ans:** A detection error tradeoff (DET) graph is a graphical plot of error rates for binary classification systems, plotting the false rejection rate vs. false acceptance rate. The x- and y-axes are scaled non-linearly by their standard normal deviates (or just by logarithmic transformation), yielding tradeoff curves that are more linear than ROC curves, and use most of the image area to highlight the differences of importance in the critical operating region[2].

Below is the graph between False Positive Rate Vs False Negative Rate.



**References:**

1. S. Narkhede, “Understanding AUC - ROC Curve ”, 26, June 2018 [Online]. Available: <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5> [Accessed Nov. 11, 2019]
2. A. Martin, “Detection error tradeoff”, 4, Sep 2017 [Online]. Available: <https://en.wikipedia.org/wiki/Detection_error_tradeoff> [Accessed Nov. 11, 2019]