Algorithm steps:
1) 2 scores are calculated for each term-class pair
1st score measures the degree of belongingness or probability that a term belongs to a particular class.
2nd score measures the degree to which a term contributes in discriminating the class.

2) Score 1:
Probability of belongingness of a term depends on the number of documents in which a term appears and number of unique terms appeared in that class. Algorithm for calculating PPD value of a term is given below.

Input: Document corpus (D) with labels (C) positive or negative,
Output: Ranked term list based on the scores calculated

Step 1 Preprocessing
t ← Extract Unique Terms(D)
F ← Total Unique Terms(D)
Wp ← Total Terms In PositiveClass(D,C)
Wn ← Total Terms In NegativeClass(D,C)

Step 2 Main Feature Selection loop
for each t in F
 Ntp =CountPositiveDocumentsInwhichTermAppears(D,t)
 Ntn =CountNegativeDocumentsInwhichTermAppears(D,t)

[Negative documents refers to all which belong to any other classes, One against rest case]

end

for each t in F

$$score1 = \frac{Ntp}{W_p+F} - \frac{Ntn}{W_n+F}$$

end

Logic:
If a term has almost equal probability of belongingness to both the categories, in that case the
term is not useful in discriminating the class. The score of a term is calculated by computing the difference of probabilities that a term will belong to positive class or negative class. Thus, if a term has high score, it indicates that the term is important for classification.

Score 2:
This score of a term is computed by finding the ratio of the difference between the number of documents of a category in which it

appears and the number of documents in which it appears of another category, to the total number of documents in which that term appears.

$$score2 = |\frac{posD - negD}{posD + negD}|$$

posD is the number of positive document in which a term appears, and negD is the number of negative documents in which that term appears.

LOGIC: If this score of a feature is close to 1 it means that this feature is occurring dominantly in only one category of documents. For example if "Excellent" word is occurring in 150 positive documents and in 2 negative documents, then value of this feature will be (150-2)/(150+2)= 0.97 , its value is near to 1 indicates that this term is useful in identifying the class of unknown document. It indicates that if a new document is having "excellent" word, there is a high chance that this document belongs to positive category.Similarly if a word occurs in same number of positive and negative documents, then Score 2 will be 0, which indicates that this term is not useful for classification.

The two scores are calculated for each term class pair.The Feature selection method employed takes a Harmonic Mean of both the above scores to get a final term rank list for each class.
Harmonic mean ensures that if atleast one of the two scores are small, the term be ranked lower.
Logic behind rejecting a term if even one of the scores is smaller:
The 1st score can remove the terms with less document frequency, which is not important for sentiment classification like rare terms.
2nd score can eliminate the terms with high document frequency but are not important like stop words.

If the system has to be trained with K% features then, K/(number of classes) % terms are taken from the ranked term list of each class, so that no class is under-represented.

Results:

PRECISION

X	10	20	40	50	70	80	100
SVM	21	86	93	95	95	94	94
KNN	79	84	70	85	38	32	85
NB	94	97	95	95	95	95	95

RECALL

X	10	20	40	50	70	80	100
SVM	23	53	90	93	93	93	93
KNN	63	77	57	40	27	40	80
NB	93	93	97	93	93	93	93

FSCORE

X	10	20	40	50	70	80	100
SVM	12	55	90	93	93	93	93
KNN	59	74	54	40	18	27	80
NB	93	97	94	94	94	94	94

Accuracy

X	10	20	40	50	70	80	100
SVM	23	53	90	93	93	93	93
KNN	63	77	57	40	27	40	80
NB	93	97	93	93	93	93	93