Table 1Ranking table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Countvectorizer  Acc | Countvectorizer  Cm | Tfidf-vectorization  Cm | Tf-idf  Accuaracy |
| 1.guassianNb | 73% | array([[55, 42],  [12, 91]] | array([[57, 40],  [16, 87]], | 72% |
| 2.Bernoulli NB | 76% | array([[70, 27],  [20, 83]], | array([[70, 27],  [20, 83]] | 76% |
| 3. Multinomial Nb | 76.5% | array([[72, 25],  [22, 81]], | array([[72, 25],  [22, 81]], | 76.5% |
| 4.XG-boost | 72.5% | array([[84, 13],  [42, 61]], | array([[80, 17],  [45, 58]], | 69% |
| 5.Gradient Boost | 74 | array([[92, 5],  [47, 56]], | array([[91, 6],  [45, 58]], | 74.5% |
| 6.Ada Boost | 71,HPT(n=100)=72% | array([[90, 7],  [51, 52]], | array([[81, 16],  [47, 56]], | 68 ,HPT(n\_est=100)=69.5% |
| 7.Logistic Regression | 71%,  HPT(l1,solver=liblinear)=72% | array([[76, 21],  [37, 66]], | array([[85,12],  [37, 66]], | 75.5% |
| 8. SVC | 73% | array([[90, 7],  [47, 56]], | array([[83,14],  [35, 68]], | 75.5% |
| 9.KNN(HPT) | 63 | array([[80, 17],  [57, 46]], | array([[74,23],  [24, 79]], | 76.5 |
| 10.Decision tree(HPT) | 70% | array([[74, 23],  [35, 68]], | array([[81,16],  [34, 69]], | 75% |
|  |  |  |  |  |

Ranking Table :

KNN (TFIDF) > Bernoulli NB (TFIDF) = Bernoulli NB (CV) > Multinomial (TFIDF) = Multinomial NB (CV) > Logistic Regression (TFIDF) = SVC (TFIDF) > Gradient Boost (TFIDF) = Gradient Boost (CV) > SVC (CV) = Gaussian NB (CV) = decision Tree (CV) > Ada Boost (CV) = Gaussian NB (TFIDF) > Logistic Regression (CV) = Ada Boost (TFIDF) > XG Boost (CV) > KNN (CV) = XG Boost (TFIDF)