

Part 3

Setup:

Cross validation is applied to KNN, SVM, Decision Tree, and Random Forest, their accuracy and f1 score is calculated and then they are compared to find the best one.

Accuracy:

KNN:

N_neighbours	Metric	Mean	Std	Confidence Interval	
Manhattan	10.00	0.74	0.03	0.71	0.76
Manhattan	10.00	0.73	0.03	0.70	0.75
Cosine	10.00	0.71	0.03	0.68	0.74
Manhattan	15.00	0.72	0.02	0.70	0.75
Manhattan	10.00	0.71	0.04	0.68	0.75
Manhattan	15.00	0.72	0.03	0.70	0.75
Cosine	15.00	0.72	0.02	0.70	0.74
Manhattan	10.00	0.73	0.02	0.71	0.75
Manhattan	10.00	0.73	0.04	0.70	0.77
Cosine	15.00	0.73	0.02	0.71	0.74
Cosine	10.00	0.73	0.03	0.71	0.76
Cosine	15.00	0.72	0.03	0.69	0.75
Cosine	10.00	0.72	0.03	0.70	0.75
Manhattan	10.00	0.74	0.03	0.71	0.77
Manhattan	15.00	0.71	0.03	0.69	0.74

SVM:

C	Kernel	Mean	STD	Confidence Interval	
10.00	RBF	0.73	0.03	0.70	0.76
10.00	RBF	0.73	0.03	0.70	0.76
10.00	RBF	0.70	0.04	0.67	0.74
10.00	RBF	0.73	0.03	0.71	0.76
10.00	RBF	0.71	0.03	0.68	0.74
10.00	RBF	0.74	0.03	0.71	0.77
10.00	RBF	0.71	0.03	0.68	0.74
10.00	RBF	0.73	0.04	0.69	0.77
10.00	RBF	0.74	0.04	0.70	0.77
15.00	RBF	0.74	0.04	0.70	0.77
10.00	RBF	0.72	0.03	0.70	0.75
10.00	RBF	0.71	0.04	0.67	0.75
10.00	RBF	0.72	0.04	0.69	0.76
10.00	RBF	0.74	0.03	0.72	0.77
10.00	RBF	0.71	0.04	0.68	0.75

Tree:

Ccp_alpha	Criterion	Mean	STD	Confidence Interval	
0.02	Entropy	0.72	0.03	0.69	0.75
0.02	Gini	0.70	0.01	0.70	0.70
0.02	Gini	0.70	0.01	0.68	0.71
0.01	Entropy	0.72	0.02	0.70	0.74
0.02	Gini	0.70	0.00	0.70	0.70
0.01	Entropy	0.71	0.04	0.66	0.75
0.01	Entropy	0.71	0.04	0.67	0.75
0.01	Gini	0.71	0.02	0.69	0.71
0.02	Entropy	0.70	0.01	0.69	0.71
0.01	Entropy	0.72	0.03	0.69	0.75
0.02	Gini	0.70	0.00	0.70	0.70
0.01	Entropy	0.70	0.04	0.67	0.74
0.01	Entropy	0.70	0.03	0.67	0.73
0.01	Gini	0.72	0.02	0.69	0.74
0.01	Gini	0.71	0.03	0.68	0.74

Random Forest:

N_estimators	Criterion	Mean	STD	Confidence Interval	
150.00	Entropy	0.77	0.03	0.76	0.78
150.00	Entropy	0.75	0.02	0.74	0.75
150.00	Entropy	0.75	0.02	0.74	0.76
150.00	Entropy	0.77	0.03	0.76	0.78
150.00	Gini	0.74	0.03	0.72	0.75
100.00	Entropy	0.75	0.03	0.74	0.76
150.00	Gini	0.74	0.02	0.74	0.75
150.00	Entropy	0.75	0.03	0.74	0.76
150.00	Entropy	0.76	0.03	0.75	0.77
150.00	Gini	0.76	0.02	0.75	0.77
100.00	Entropy	0.75	0.02	0.74	0.76
150.00	Entropy	0.75	0.03	0.74	0.76
100.00	Entropy	0.75	0.22	0.74	0.76
150.00	Entropy	0.76	0.03	0.75	0.77
100.00	Entropy	0.74	0.02	0.73	0.75

Best Method:

KNN	0.72	0.01	0.71	0.73
SVM	0.74	0.03	0.73	0.75
Decision Tree	0.70	0.02	0.70	0.71
Random Forest	0.76	0.01	0.76	0.77

Best: Random Forest.