12012019 – 12012052

WEEK 10 ML ASSIGNMENT

EXPLANATION

LOGISTIC REGRESSION DECISION TREE KNN SVM

ACCURACY 0.72 0.65 0.67 0.73

Based on the provided accuracy values, the Support Vector Machine (SVM) appears to have the highest accuracy (0.73) among the models mentioned but accuracy is just one metric for evaluating model performance, and it may not always be the most appropriate measure, especially if the dataset is imbalanced or if other factors like precision, recall, or F1-score are more critical for the given problem.

Therefore, in terms of accuracy, the SVM seems to perform better than other models for the given dataset and problem.

Now, we have to compare the accuracy of SVM with other ensemble techniques.

RANDOM FOREST ADABOOST SVM

ACCURACY 0.66 0.73 0.73

SVM and Adaboost have the same accuracy of 73%. In order to explore further, we have to compare their precision, recall and F-score.

AdaBoost Classifier:

Precision (Class 1): 0.75

Recall (Class 1): 0.80

F1-Score (Class 1): 0.77

Accuracy: 0.73

SVM Classifier:

Precision (Class 1): 0.74

Recall (Class 1): 0.82

F1-Score (Class 1): 0.78

Accuracy: 0.73

When comparing the precision, recall, and F1-score for Class 1 (which is the positive class in the given problem), SVM performs slightly better in terms of recall (0.82 vs. 0.80) and F1-score (0.78 vs. 0.77) for the positive class. Therefore, based on these metrics, the SVM classifier performs slightly better than the AdaBoost classifier.

However, it is essential to consider the specific requirements of the given problem when

determining which model is truly the best choice. For instance, if we want higher recall

(identifying as many positive cases as possible) over precision, the SVM model might be the

preferred choice due to its higher recall for the positive class. Conversely, if precision is

more critical (minimizing false positives), we might lean towards the AdaBoost model,

although it has a slightly lower recall.