Program #4: Machine learning

CISC 3410

In this problem you will use classifiers in the python sklearn package to learn a classification model for a chessboard-like dataset.

1. Open the dataset chessboard.csv in python. Make a scatter plot of the dataset showing the two classes with two different patterns.
2. Train SVM, k-nearest-neighbors, and decision tree classifiers for this data.
   (a) For SVM, try linear, sigmoid, RBF, and poly kernels.
   (b) For KNN, try k values of 1, 5, 10, and 15.
   (c) For the decision tree, try max_depth of 1, 2, 4, 8, and 16.
3. Find the classification accuracy for each classifier with 10-fold cross-validation. Report the mean and standard deviation for each classifier's accuracy.
4. Plot the decisions made by each SVM classifier on the training data.
5. Discuss your results.

Submit:

1. Code for your classification experiments (python)
2. Table with classification accuracies for each classifier (4 + 4 + 5).
3. Paragraph discussing which classifiers and parameters are most appropriate for this task.

Submit through Blackboard.