**INFT 5603 Principals of Data Science, Fall 2023**

**Graduate College, Information Technology, Arkansas Tech University**

**Assignment 13 / Due date: Wednesday, November 29, 2023 before midnight. Total points: 100.**

1. (*30 points*) Explain support vector machine (SVM) classification algorithm in a few sentences.

* SVM is a robust classification technique particularly effective in high-dimensional spaces, with its capability to find complex decision boundaries using kernels.
  + **Objective**: SVM aims to find a linear hyperplane (decision boundary) that separates the data into different classes. The core idea is to maximize the margin between different classes.
  + **Hyperplane and Margin**: The hyperplane is a line (in two dimensions) or a plane (in higher dimensions) that divides the data points. The margin refers to the distance between the hyperplane and the nearest data points of each class, known as support vectors.
  + **Support Vectors**: These are the data points closest to the hyperplane. The position of the hyperplane is determined based on these support vectors, hence the name of the method.
  + **Maximizing the Margin**: The best hyperplane is the one that has the maximum margin, meaning it is farthest from the support vectors of each class.
  + **Linear and Nonlinear Classification**: SVM can be used for both linear and nonlinear classification. For linear classification, it finds a straight line (or plane) as the decision boundary. In cases where the data is not linearly separable, SVM uses the kernel trick to transform the data into a higher-dimensional space where a hyperplane can be found.
  + **Kernel Trick**: This is used in nonlinear SVM. It involves transforming the data into a higher-dimensional space to make it possible to separate the classes linearly. Common kernel functions include polynomial, radial basis function (RBF), and sigmoid.
  + **Optimization Problem**: Learning in SVM is an optimization problem, often solved using methods like the Lagrange multiplier method. It involves maximizing the margin while minimizing a penalty term related to misclassification.

1. (*70 points*) Implement the support vector machine (SVM) method to the Iris data set in Python. Use the polynomial kernel function for SVM. Also, show the confusion matrix and normalized confusion matrix. *Note*: You can see SVM pre-defined function and it’s all parameters from the following official Scikit Learn library link: https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html



A blue squares with white text

Description automatically generated

A diagram of a diagram

Description automatically generated with medium confidence

Page **1** of **1**