# INFT 5603 Principals of Data Science, Fall 2023

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

# Assignment 11/Due date: Sunday, November 12, 2023 before midnight. Total points: 100.

1. (*30 points*) Generate a random data set with the equation y = 5 + 3.x and implement the linear regression for this data set in Python.

A screen shot of a computer

Description automatically generated

1. (*35 points*) Implement the k-nearest neighbor (kNN) algorithm to the Iris data set in Python and set k = 4.

A screenshot of a computer program

Description automatically generated

1. (*35 points*) Apply Gaussian Naïve Bayes classifier for the Iris data set in Python. Use 70% of the data set for training and 30% for testing. Also, find the number of mislabeled points.

A screenshot of a computer code

Description automatically generated