**Data Science with Python**

MET CS 677

@2020 Farshid Alizadeh-Shabdiz,

MET CS 555 Assignment KNN

In this assignment you will write a generic script that plots your portfolio growth for several strategies, i.e. ”buy-and-hold” and ”true labels, and also KNN.

Recall that buy-and-hold means that you invest $100 in your stock at the beginning of week 1 and do not trade. As your stock varies in price, so does your portfolio. By contrast, ”true labels” strategy means that you trade according to the labels (”red” or ”green”). You can assign the labels manually or by writing a python script to do that:

* a ”green” week means that it was a good week to be invested for that week - means the opening price on Monday was lower than the closing price on Friday.
* a ”red” week means that it was not a good week to be invested but to keep money in cash – means opening price on Monday is higher than the closing price on Friday.

For each week calculate volatility two different ways as follows:

Parameter 1 - Volatility of reported prices from Monday to Friday of the week

Parameter 2 - Based on five numbers, one number per day of the week, which is the delta difference between the opening price of the day and the closing price of the day.

In this course, we are implementing a number of machine learning classifiers to predict labels. One way to compare our strategies is to compute elements of confusion matrices (e.g. True Positive, False Negative, etc.) Another way is to compare the growth (and volatility) of your (single stock) portfolio.

In this assignment we will focus on comparing portfolio performance vs. buy-and-hold and true-labels portfolios.

Questions:

1. Plot portfolio growth for ”buy-and-hold”, true labels and 3 nearest neighbor strategies. For KNN distance measurement use week number, and volatility parameter 1 and 2 as you it was defined above, and use Minkowski distance with P = 1, 1.5 and 2.

Plot three graphs for KNN with Minkowski distance with P = 1, 1.5 and 2, for year 2.

Each of the 5 plots should be in different color. On X-axis is the week number and on the Y axis is the value of your portfolio for a strategy. In the legend section of the graph, indicate for each strategy an average weekly portfolio value. In other words, if you have 52 weeks 1, . . . , 52 on X-axis, you compute the portfolio values P1,...,P52 at the end of weeks 1, . . . , 52 respectively.

Hint: if you have a Pandas frame where column ’X’ contains week numbers and other columns containing portfolio values, then you can plot all your strategies using Pandas ”plot” method.