Ta’Quawn Watts

Plotting, Salting, and Smoothing using Apache and JFreeCharts

For Project 2 of CSCi-3327 Probability and Applied Statistics course at Stockton University, my task was to create a program implementing Apache Commons Math Library and JFreeCharts to produce a variety of graphs.

Apache Commons Math is a Java library of mathematics and statistics functions that is not available in Java programming. It contains functions related to numerical analysis, linear algebra, optimizations and more. JFreeCharts is a Java chart library that allows users to create a variety of charts like bar charts, pie charts, scatter plots and more. For this project, I implemented Apache Commons Math Library version 3.61 and JFreeCharts version 1.0.19.

Using Apache Commons Math Library and JFreeCharts, the program should produce three charts. The first chart should be the graph of a function. The second graph should be a graph of the function with “salted” output values. Salting, or adding noise, to data means to add randomness to input data. Salting data is very common with areas like machine learning and neural networks as it benefits neural networks and models performance with variations in data. The third graph should be a graph after “smoothing” the previously salted values. Smoothing reduces the randomness from data set typically by using an algorithm. Smoothed data is usually used to find and predict trends within the data set.

A significant part of this task was learning how to implement Apache Commons Math Library and JFreeCharts into it. Before this project, I had no experience with implementing external libraries into Java. First, I researched how to import external libraries into Eclipse, my Java IDE of choice. Afterwards, I downloaded the Apache Commons Math Library and JFreeCharts jar files and imported it into my library. After testing to see if my classes could use the two libraries’ methods, I read through the documentation. After getting a grasp of what the two libraries had to offer, I began to research how to actually implement them into my own project. Majority of the learning process involved researching uses for the libraries and how to implement them into my own work.

The function that I chose to use for this program was . I chose this function because it was a parabola. I was interested to see how drastically the salted graph would change and how accurate the smoothed graph would be to the original. It was also the same function I used in another Java program I created that plotted, salted, and smoothed data using only Java. functionSolver method has a parameter of integer x representing the variable x in the function and returns the output of .

The method functionXY uses XYSeries, a class that stores x and y values in a data set as data points, to store the x and y values of the function. functionXY has two parameters, integer xMin, and integer xMax. These parameters represents the range of input values input into the function. Using functionSolver, functionXY stores the input and output values of the function in arrays and uses XYSeries to store them as x and y points. It then returns the XYSeries.

The method saltXY uses XYSeries to store the x and salted y points of the inputted XYSeries. functionXY has three parameters: XYSeries functionXY, integer saltMin, and integer saltMax. The functionXY parameter is the XYSeries file that will get salted values. The saltMin and saltMax parameter is the range of the randomized numbers to be added to the y points. After adding the y points to an array, saltXY adds or subtracts a random number, generated by Apache Common Statistics Library’s RandomDataGenerator class, within the specified range to each y point. Afterwards, it adds the original x values and salted y values to a XYSeries and then returns the XYSeries.

The method smoothXY uses XYSeries to store the x and smoothed y points of the inputted XYSeries. smoothXY takes two parameters: XYSeries saltedXY and integer windowValue. The saltedXY parameter is XYSeries that will get smoothed values. The windowValue is the number of points being taken right and left of the index point. After adding the salted y points to an array, smoothXY uses Apache Common Math Library’s DescritpiveStatistics to add the points to a set and takes the average of the values around the initial value set by windowValue. After smoothing each point, it adds the original x values and smoothed y values to a XYSeries and then returns the XYSeries.

The method displayGraphs uses JFreeCharts to plot and display graphs using data sets generated by functionXY, saltXY, and smoothXY. displayGraphs takes five parameters: xMin, xMax, saltMin, saltMax, and windowValue. xMin and xMax represents the range of input values input into the function. The saltMin and saltMax parameter is the range of the randomized numbers to be added to the y points. windowValue is the number of points being taken right and left of the index point. Using various JFreeCharts, displayGraphs creates a function graph, a salted graph, and a smoothed graph. With the parameters -25, 25, 0, 100, 5, displayGraphs outputted the following graphs:

![Chart, line chart

Description automatically generated]()

![Chart, scatter chart

Description automatically generated]()

![Chart, line chart

Description automatically generated]()

The original function is a parabola. However, after salting the data, the smooth curves of the parabola turn into random jagged lines. After smoothing the data, the graphs begins to look very similar to the original function’s graph.

Works Cited

Dhir, R. (2022, November 20). *Data smoothing: Definition, uses, and methods*. Investopedia. Retrieved April 24, 2023, from https://www.investopedia.com/terms/d/data-smoothing.asp

The Apache Commons Mathematics Library. (n.d.). Retrieved April 24, 2023, from https://commons.apache.org/proper/commons-math/

TutorMaster, A. I. (2023, February 2). *What is gaussian noise in deep learning? how and why it is used?* Medium. Retrieved April 24, 2023, from https://ai.plainenglish.io/what-is-gaussian-noise-in-deep-learning-how-and-why-it-is-used-af3730449e3a

www.jfree.org. (n.d.). Retrieved April 24, 2023, from https://www.jfree.org/