BLG 454E Project Proposal

Interactive Regression and Clustering Tool

Team:
- Ahmet Enes Topçu, 150210310
- Alihan Esen, 150220338
- Bedirhan Öztürk, 150210321

Abstract:
Our goal is to create an interactive window that allows users to enter datapoints into given textboxes or by clicking at any point on the graph and choose whether to apply regression or clustering to those datapoints.

Problem Statement:
Our motive was to make an intuitive tool that could be used by anyone with little to no coding and math knowledge. There are a couple of tools that achieve this on the internet, however we couldn't find a simple all-in-one tool for regression and clustering.

Methods - Data - Expected Results:
We plan to use Python for the code, NumPy for the math and the Tkinter library for the interactive window. We plan to implement a couple of clustering methods and add polynomial regression.

We will add a couple of datasets and apply clustering and regression to those datasets to provide examples. We will use datasets from github.com/deric[1] and NIST[2]

At the end of this project, we expect to achieve an all-in-one clustering and regression tool that is intuitive and simple, the tool will provide multiple clustering methods, linear regression and polynomial regression.

Users will be able to add datapoints in a couple of ways:
- By clicking on the graph

- By entering the x and y values on a textbox
- By importing a CSV file (with a specified layout)

[1] https://github.com/deric/clustering-benchmark
[2] https://www.itl.nist.gov/div898/strd/lls/lls.shtml