IRSv2 Group B

Jared Raiola, Adam Chau, Seungju Jason Lee

Introductions

- Jared Raiola
  - Computer Science - Third Year
  - Programming Experience: Python, Java, SQL, JavaScript, HTML, CSS, PHP, C, C++
- Adam Chau
  - Computer Science - Third Year
  - Programming Experience: Python, Java, SQL, JavaScript, HTML, CSS, C#, C
- Seungju Jason Lee
  - Electrical Engineering - Third Year
  - Programming Experience: Python, Java, SQL, C++

Objective

To improve data visualization, question difficulty scaling and gear class focus towards more difficult topics

Motivation

Provide a tool that informs the user which topics covered in the class of Digital Signal Processing is determined to be the most difficult.

With this information, the user can determine which topics to focus on more heavily and those topics that can be focused on to a lesser degree.

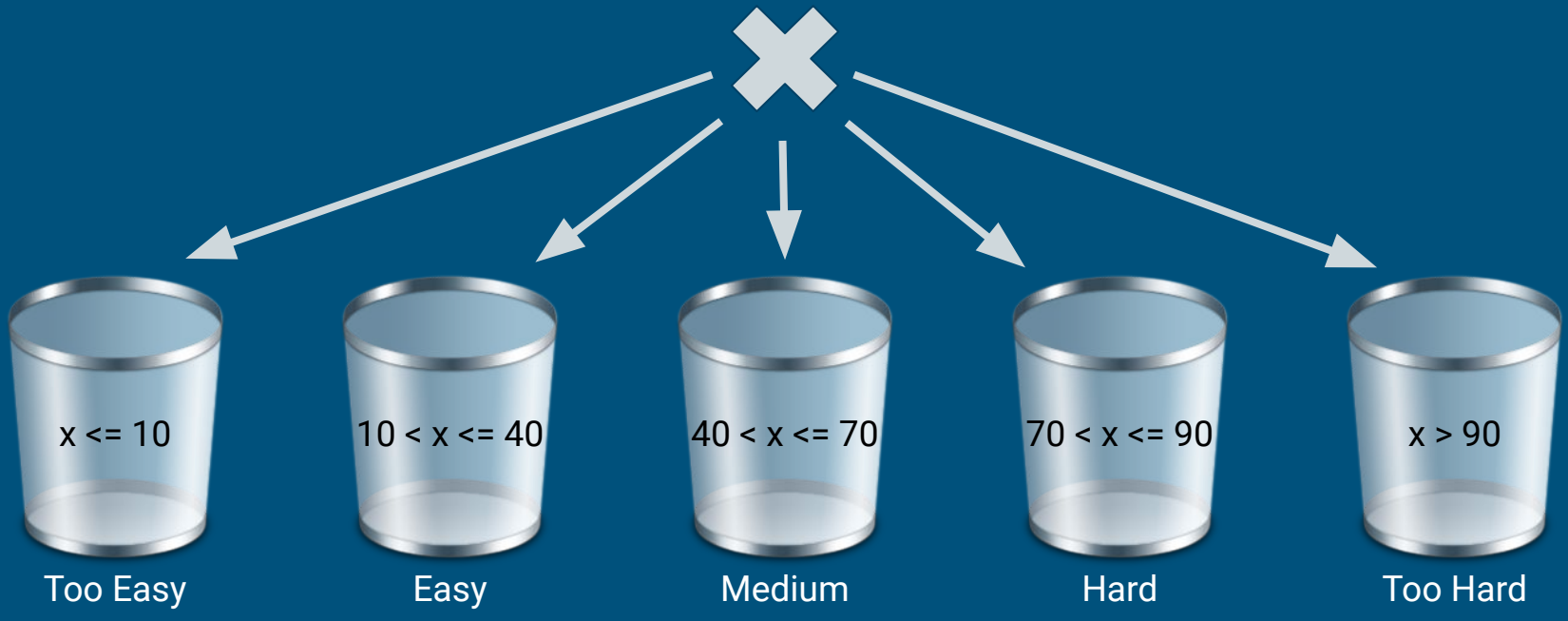
Tools Used

Python Libraries

Other Tools

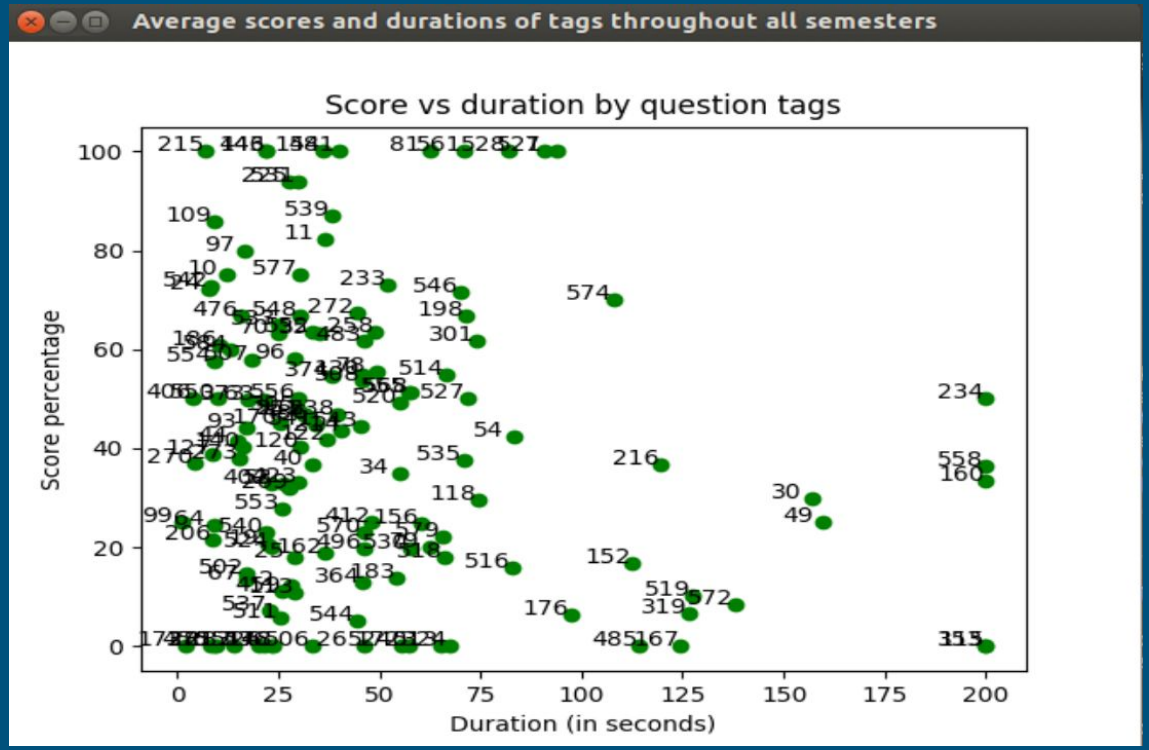
Difficulty Metric

Difficulty = x = $100 - \left[\frac{\frac{300 - duration}{60} + 15\left(\frac{score}{100}\right)}{20}\right] * 100$

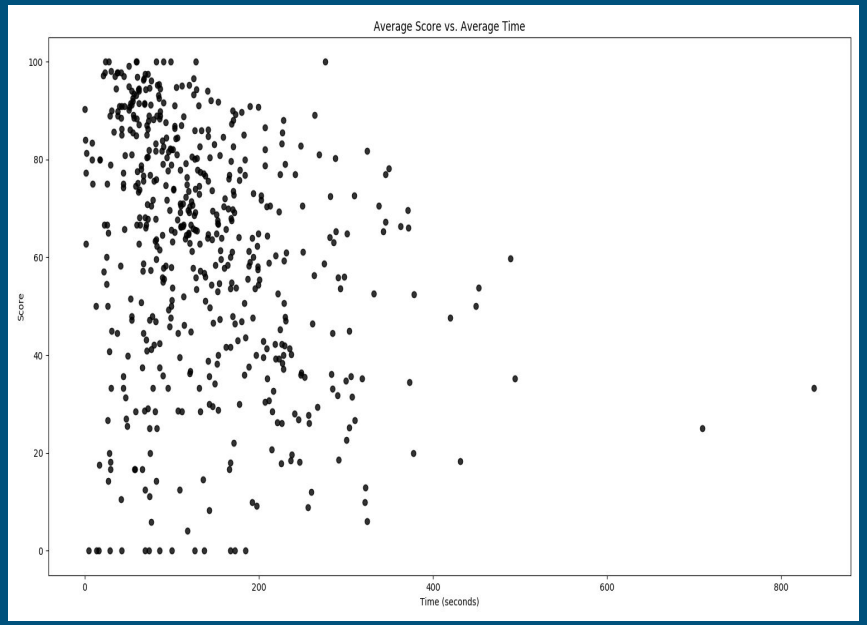


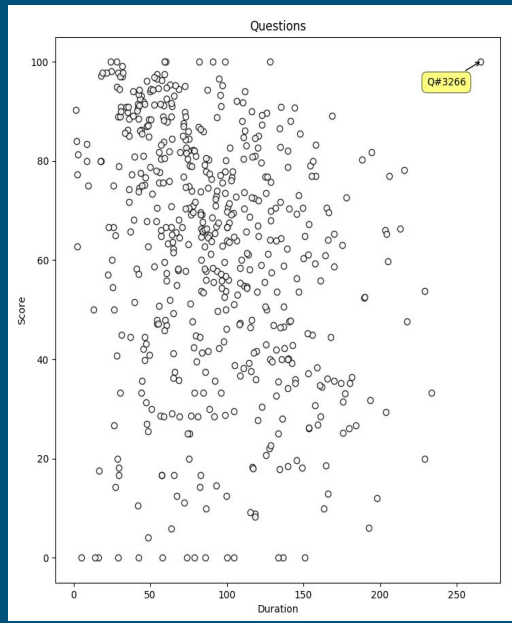
Graphing

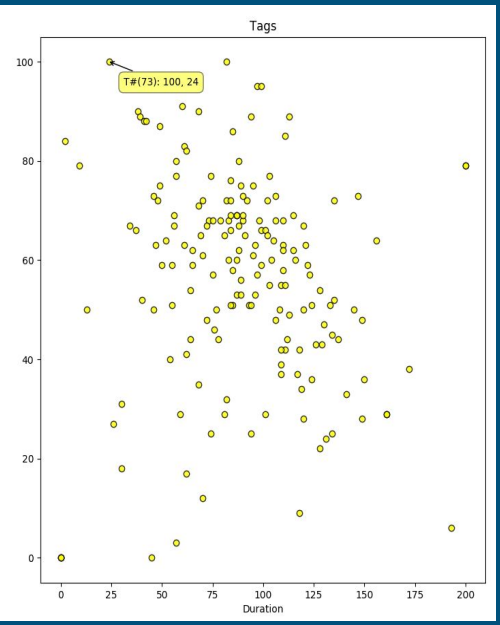
Early Implementation



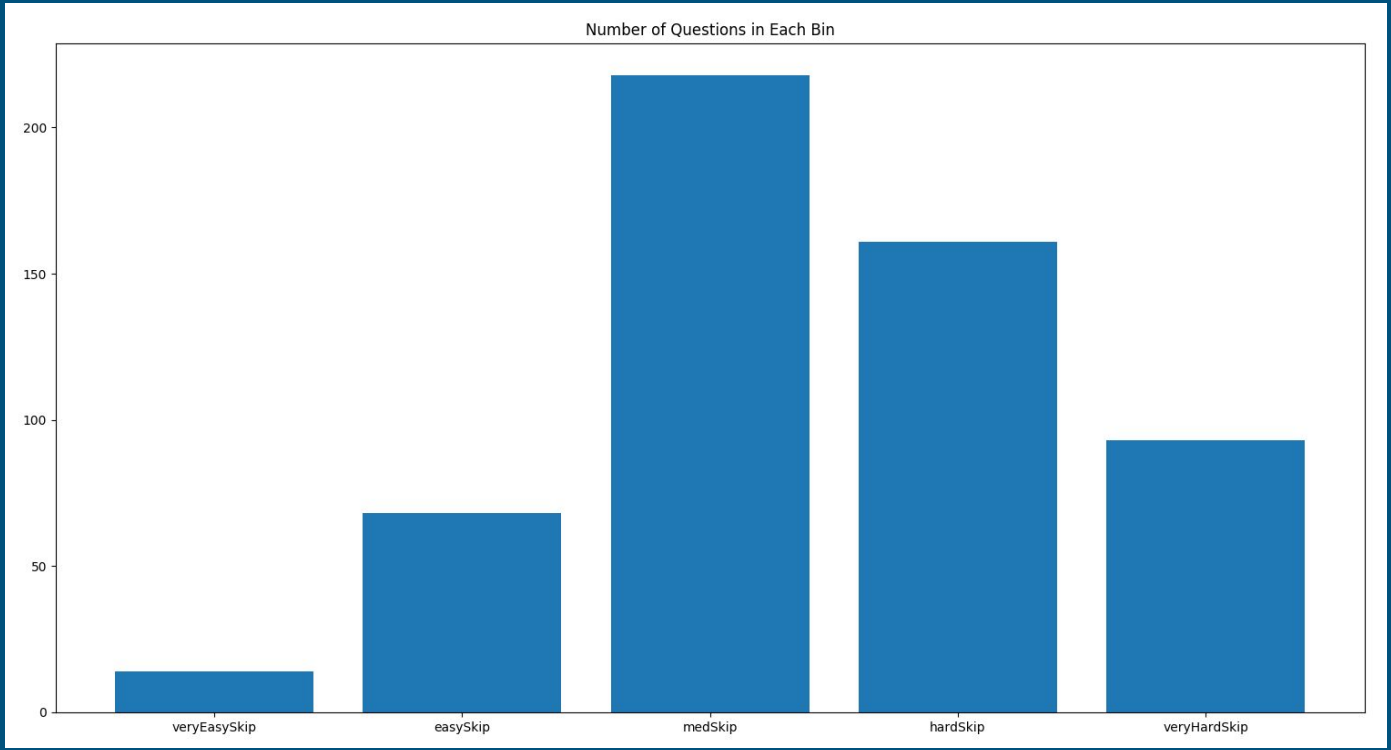
ScatterPlots of Scores vs Duration



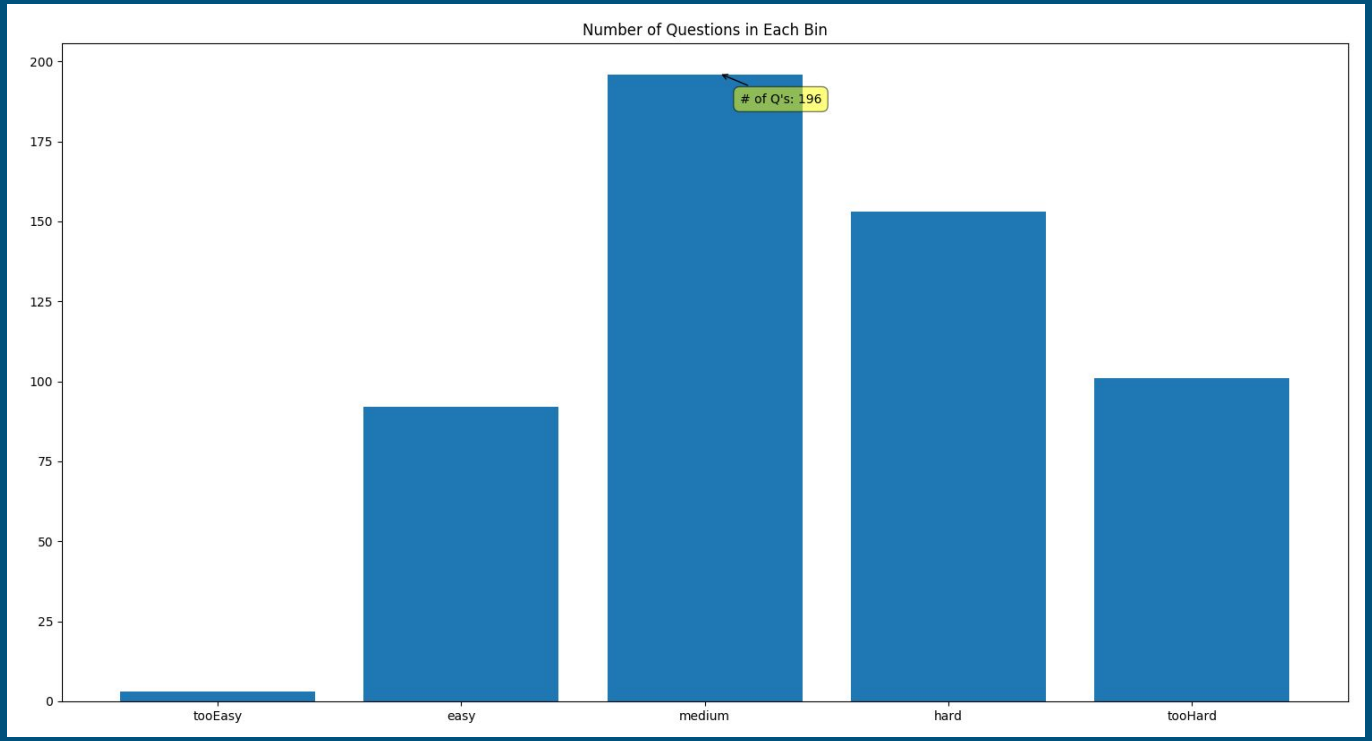




Questions in Skip Bins

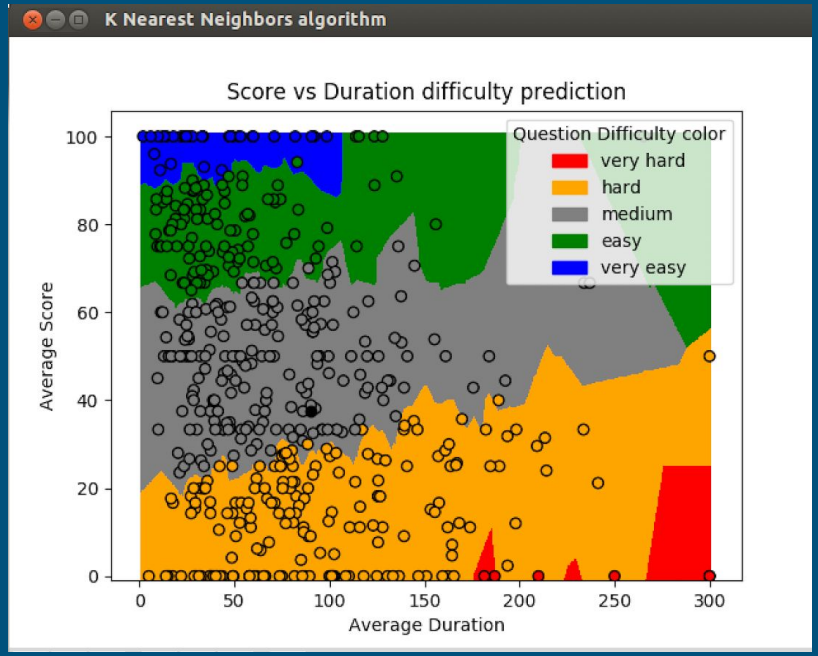


Questions in Each Difficulty Category

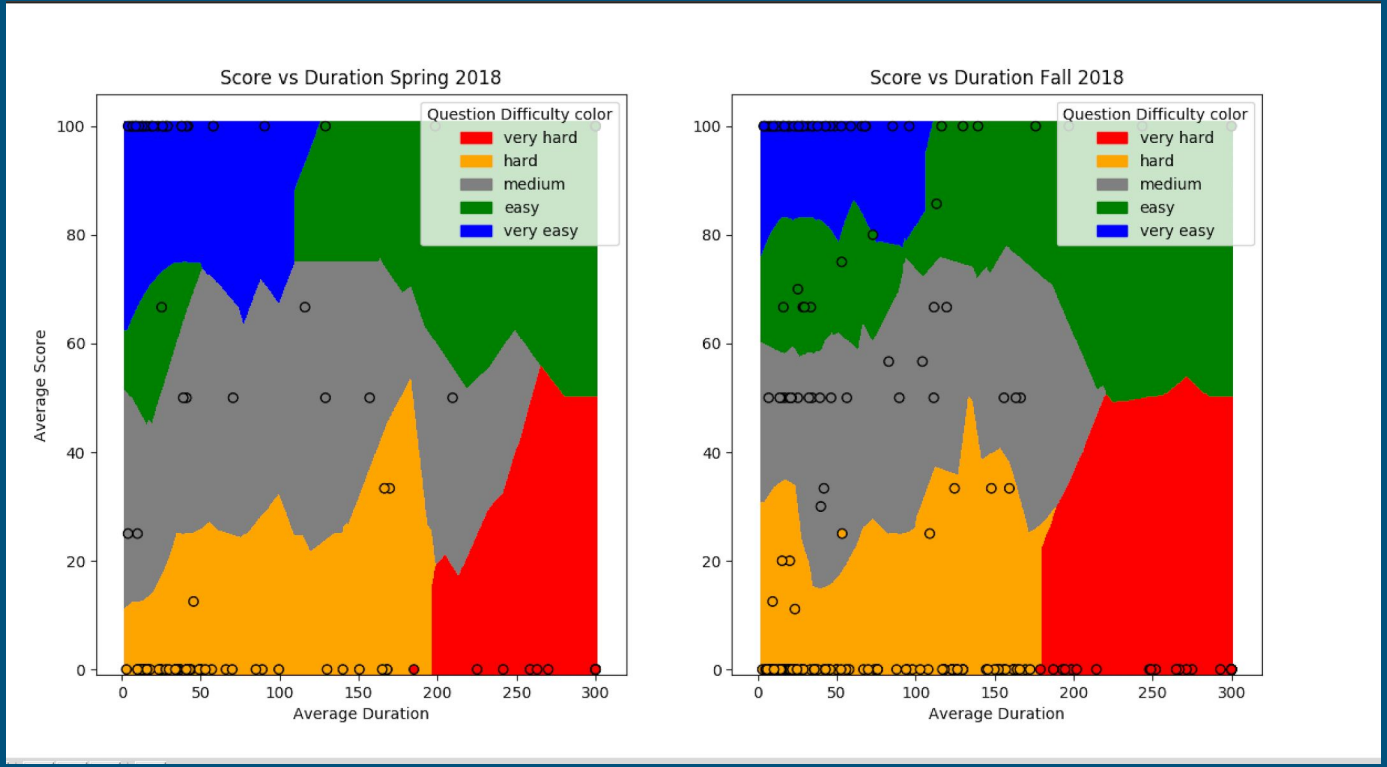


K-Nearest Neighbors Algorithm

- A type of supervised learning in which an input (average score and duration) has a class determined using a certain amount of nearest points (neighbors)
- The k value denoted the closet neighbors the point needs to be to belong in the class
- Knn is very accurate but requires a large training set and prior information
- Can be used to predict difficulty scores for future scores and durations



Machine Learning Graph by semester



Demo

Challenges in the Project

- Downloading and installing Ubuntu on VirtualBox (Jason)
- Processing the database (null values, large data collection)
- Trying to decrease the runtime of the Python programs
- Working with proper graph hover commands
- Virtual Machine space issues

Future Implementations

- Web Application: Results in real time
- Question Scoring Changing: Per difficulty of question
- Question Presentation: Cycling different difficulty based on student performance
- Difficulty Calculation: Implement more factors towards calculation

Questions?