Final Project: predicting the growth & change of Boston real estate market

Jessica Hu, Annie Li, Bill Zheng

1: What are we working?

This project is specifically oriented to predict the change of Boston's real estate market in the future by learning this region's past. We will incorporate machine learning into the process of predicting housing prices.

2: What are the data sources?

The data source will come from the real estate resources provided by the state government. We have arranged the data to be in csv form, which can be used directly for our data analysis.

3: what are the training algorithms?

For now, we are using generalized linear models such as the least squares to compare the pricing of the houses with different conditions. As time progressed, we will incorporate more and more sophisticated data learning methods to meet our needs.

We mainly used linear regression to perform the training algorithm, as it requires less code to operate, and it can perform at a high efficiency according to the data we acquired. Overall, our training algorithm showed a 70% accuracy in terms of predicting the sales price, which is above over expectation of our program.

4: what will the visualization look like?

The visualization will take in many forms: bars, lines, charts, etc. They can be customized by the users as they wish.

Throughout the course of the project, we have realized that a customized visualization will have two main disadvantages: first being that the programming is too intensive, and the second being

that a lot of the methods, such as, pie chart and circle charts, are ineffective when representing the change of the housing prices of the greater Boston area. In the end, we chose a bar graph and linear chart to represent the change of price, which greatly improved the efficiency, and magnified the function of our website: the change of price over time (especially in the future).

Week 3: starting the basis of the application: making algorithm tests.

Annie: Building the backbone of the program (make sure that the hardware runs properly)

Jessica: Programming the algorithm (using all 3 computers)

Bill: Test the efficacy of different algorithms, making sure that we use the most efficient algorithm (least amount of memory storage for learning and has the max depth).

Week 4: Server development

Annie: Importing the data to the server

Jessica: Developing the server (hardware-wise)

Bill: Building the server environment and making sure that everything works.

Week 5: Web development

Annie: coding the web interface

Jessica: designing the website.

Bill: maintaining the website (make sure that the server and the website fits properly)

Week 6: Finalizing Application

Annie: building the link between app, web, and server

Jessica: importing the data between app, web, and server

Bill: Programming the application and publishing it.