Progress Report

Ochaun Marshall, Richard Andrews, Christian Simaan, Jeremy Hutton

The NostraDomicile Project has moved forward in the last few weeks. The profile of the system is a web application comprised of a database, backend, frontend, machine learning algorithms, and a cloud hosting service, whose two main functions are to predict whether a house will sell in a specific area based on the home’s attributes, and given a zip code, what are the most important factors leading to a sale in that area.

Our approach was initiated using research to find the most promising path to completing this project. For the database implementation, we decided to use MySQL as our group has more experience with that language, though a NoSQL database such as MongoDB might be faster. The backend will be created using Django, as our machine learning algorithms will run through Python, a disadvantage of this choice is that no one on our team has any experience using Django while some have experience with Ruby on Rails. For the frontend of the web application, we are using Bootstrap and AngularJS, as they appear to be industry standards with lots of documentation and support, and will speed up the process of building the visual piece of the application, while using a combination of HTML/Javascript/CSS would take too long to learn in the timeframe we have. For the hosting service, we chose AWS because we were interested in learning how to use it and it is dependable, as opposed to creating our own server to host the application as that might take up too much of our time.

The next step in our process was ensuring the feasibility of the system based on our ability to complete and unify the subsystems within the semester. We agreed upon a tier system where the functionality we deemed integral to the project was put in the first tier, things we hoped to add but might not have time for in the second tier, and moonshots in the third tier. The main components of the project, the database, machine learning algorithms and creating the front and backend of the website were deemed feasible as well as implementing the core features of the web application: determining most attractive features in a given area, and predicting the sale of a house based on features.

With the system architecture of the project and its feasibility determined, we delegated the responsibilities for creating each subsystem with secondary responsibilities for subsystems that might be finished early in the process.

Responsibilities:

* Christian: Database/Backend
* Ochaun: Machine Learning/Front End
* Richard: Backend
* Jeremy: Frontend

**1st Tier Timeline**

Develop data visualization functionality

Build front end of website

Develop machine learing algorithms for 2 main functions

Create expert blog

Connect database and back end of website and server

Obtain housing info

Build database using housing info

**2nd Tier Timeline**

Create price estimate based on attributes

Rank realtors by sales count

Determine if house will sell at a given price

Email results printout to user

**3rd Tier Timeline**

Automated links to Home Depot, Paint websites. etc

Suggest attributes to increase sale price

**Requirements:**

**1st Tier:**

1. Allow user to specify attributes for house qualities.
2. Obtain housing info using the Zillow API.
3. Look for available houses within zip code.
4. Create data visualizations based on home sales info.
5. A database to store housing info, query, and run data through machine learning algorithms.
6. Server side backend connected to database and frontend of web application
7. A cloud hosting service(AWS).
8. Frontend of website.
9. User able to determine whether a house will sell or not based on attributes and zip code.
10. User able to find the most desirable factors in a zip code.

**2nd Tier:**

1. Determine if house will sell at a given price.
2. Predict a viable sales price for given attributes
3. Rank realtors by sales count

**3rd Tier**

1. Suggest attributes to increase sales price.
2. Email results printout to users
3. Provide suggested product links based on most important factors