CS 499 - Team 8
Brandon Doucette

Feature Design

Decision Tree Regression

Design Constraints

Constraints that are beyond the constraints in the architecture

- **Constraint 1**
  - Data that is provided from Papa Johns
    - Each week, we had to request a larger dataset because the current set given was too small for effective measurement of the model
- **Constraint 2**
  - Unclean Data / Inaccurate Reporting
    - Data in the spreadsheet was not cleaned and a lot of data had to be removed to make the file readable for the program
    - Some data was not accurate as human error intervened and caused a lot of outliers that are not normal

Design decision 1

Decision: Cleaning data automatically with filling in empty fields instead of removing them
Because of the limited dataset that this code was run on during evaluation, 1000 rows. I did not want to risk underfitting so instead I allowed the program to fill in any missing values with 0.

Design decision 2

Decision: Regression instead of converting to classification
Implementing the DecisionTreeRegressor from sklearn.tree on the dataset allowed me to

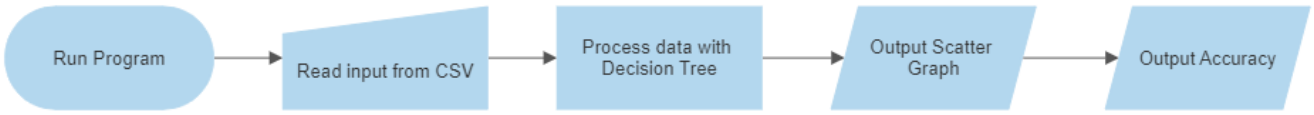
Alternative solutions and justification

Alternative solution: Classification implementation instead of regression implementation
Justification: Because there is a reasonable expectation of how long a pizza and other items could take at the slowest and fastest times, it is possible to normalize the data in expected intervals and change this in to a classification problem

Next page.

CS 499 - Team 8
Brandon Doucette

Diagram that reflects design



CS 499 - Team 8
Brandon Doucette