

INF 553: YelpProject_UPDATE 2

Accomplishments

- Filtered out only restaurants from the business dataset by using the categories column "restaurats", "food", Bar
- Found out cities with most number of restaurants to build a model around it - Toronto, Las Vegas, Pheonix, Montreal
- Inspected the use of Is_open column in the business dataset, which can be used as a label further in machine learning algorithms.

Plans

- Perform prediction using a simple naïve bayes model
- Decide which feature set to use as an input to the algorithm

Challenges

- Thinking what factors are affect the survival of a restaurant and what features can be used from the available dataset.