Food Finder

Finding you the best burrito

Kevin Wecht

The Problem:

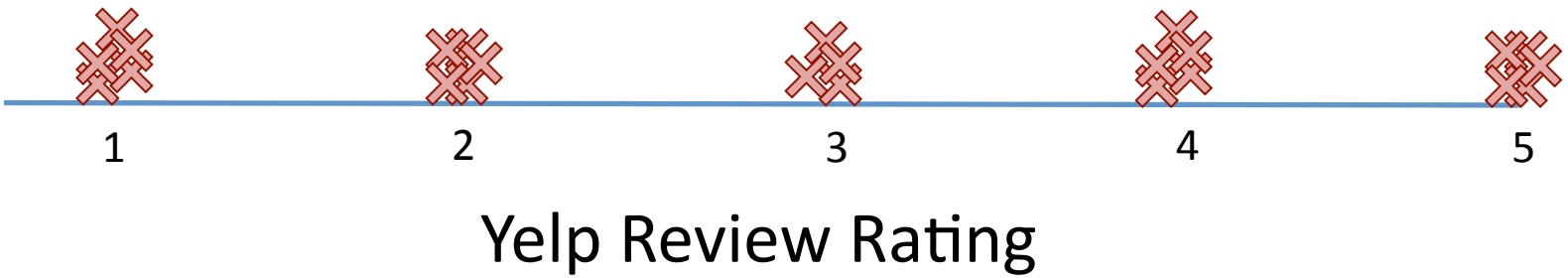
There is no way to evaluate reviews of the best restaurant “dishes” in your area.

Visualizing the Process

800 Mexican restaurants in Phoenix, AZ

40,000 Yelp reviews

Pull apart Yelp reviews for sentiment on the sentence level.

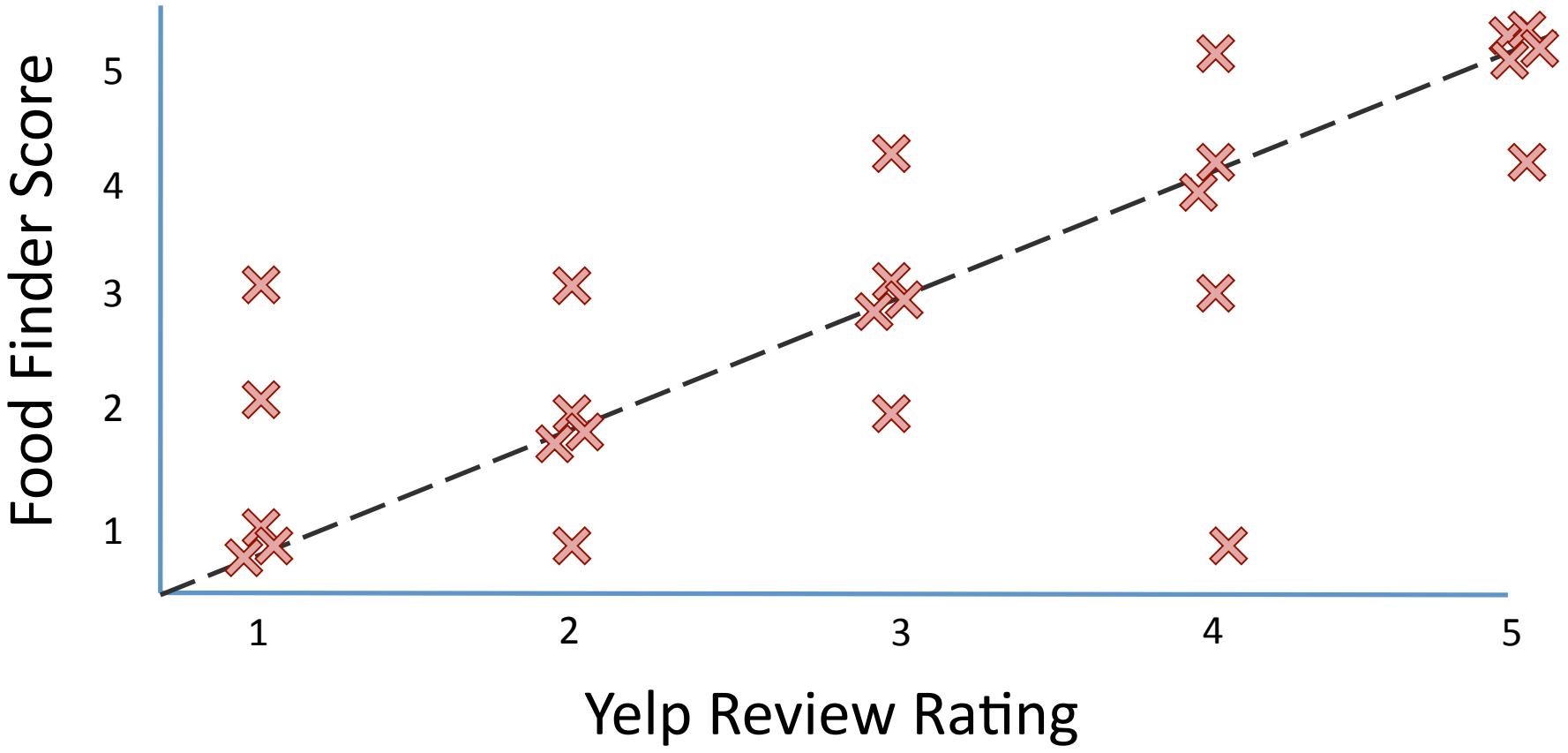


Visualizing the Process

800 Mexican restaurants in Phoenix, AZ

40,000 Yelp reviews

Pull apart Yelp reviews for sentiment on the sentence level.



Visualizing the Data

Reviews grouped by star rating

Each group of ratings has a unique vocabulary that we exploit to classify individual sentences.

1 Star

3 Star

5 Star

Algorithm

1. Natural Language Processing
   - tokenizing, POS tagging, lemmatizing, negation handling, N-grams
2. Classification
   - Multinomial naïve bayes, "logistic regression", SVM

   - What is the probability that your sentence was constructed by randomly picking from each of the 5 distributions? Returns rating with highest probability.

$$\arg\max_{R} P(R|\mathbf{w}) \propto P(R) \cdot P(\mathbf{w}|R)$$

   - Interpretability: Examine most distinguishing features
   - Feature selection: word frequency, tf-idf, doc2vec

3. Calculate "burrito score"

```
SELECT AVG(sentence)
FROM all_sentences
WHERE text LIKE "%burrito%"
GROUP BY Restaurant_ID;
```