Capstone: What makes a Yelp review useful?

Problem Statement

Yelp is a platform that allows users to make reservations, leave reviews, and find businesses. In addition, users can mark which reviews they find useful. Reviews can be both very helpful and very harmful for businesses. Being able to determine what makes a useful review can help businesses create better, more targeted listings.
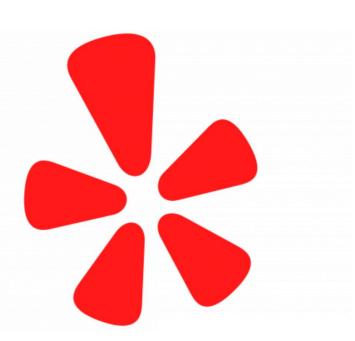
What is a "Yelp"?

- Find and connect with businesses
  - Appointments
  - Waitlists
  - Reservations
- Business info
  - Hours of operation
  - Location
- Reviews, photos
  - Funny/Cool/Useful

The Data!

- Available from https://www.yelp.com/dataset
- 6,990,280 reviews (resampled to approx 20k)
- 150,346 businesses
- 200,100 pictures

Prepping the Data

- New lines and punctuation cleaned up
- Each review is converted in preparation for modeling
  - Each word gets a real world number
  - Allows us to compare words importances

Modeling

- Model accuracy for predicting if a review is useful or not
  - Model 1: 62%
  - Model 2: 64%
  - Model 3: 62%
  - Model 4: 63%

Themes

- ## Location!
  - Shop
  - Place
  - Store
- ## Praise
  - Great
  - Good
  - Like
- ## Temporal
  - New
  - Time

Conclusion

- Somewhat limited in how accurate we can get the models
  - Accuracy around 62-64%
- All models performed comparably
  - No clear winners based on multiple metrics