Most ★★★★★ Cafes

Valentina Rizzati

Opportunity

Objective: rebuild Yelp's reputation w/ businesses

Scope: Yelp reviews on Portland cafes

Impact Hyp: by providing visibility on what customers care about, Yelp will allow cafe owners to target their investments and expand their customer base

Billion Dollar Bully Highlights Why Yelp Feels Unfair

Yelp compares itself to Wikipedia. But there's a major difference.

BY **STEPHEN HARRISON**   JUNE 06, 2019 • 5:45 AM

Methodology

Data Extraction
MongoDB with 5 collections

Topic Modeling
TF-IDF & NMF with 9 topics

Cross Validation & Model Selection
Random Forest

Data Cleaning & EDA
Business and Reviews dataframes

Baselining
Logistic Regression with Y = sentiment & X = topics

UNSUPERVISED LEARNING
Topic Modeling

Topic Modeling - Positive Sentiment Clouds

Coffee / Place / Time

Main elements in positive reviews

Food

Sweet and savoury food for breakfast and lunch

Service

Elevated customer experience

Topic Modeling - Negative Sentiment Clouds

Coffee / Place / Time

As for positive sentiment reviews, customers care a lot about these elements
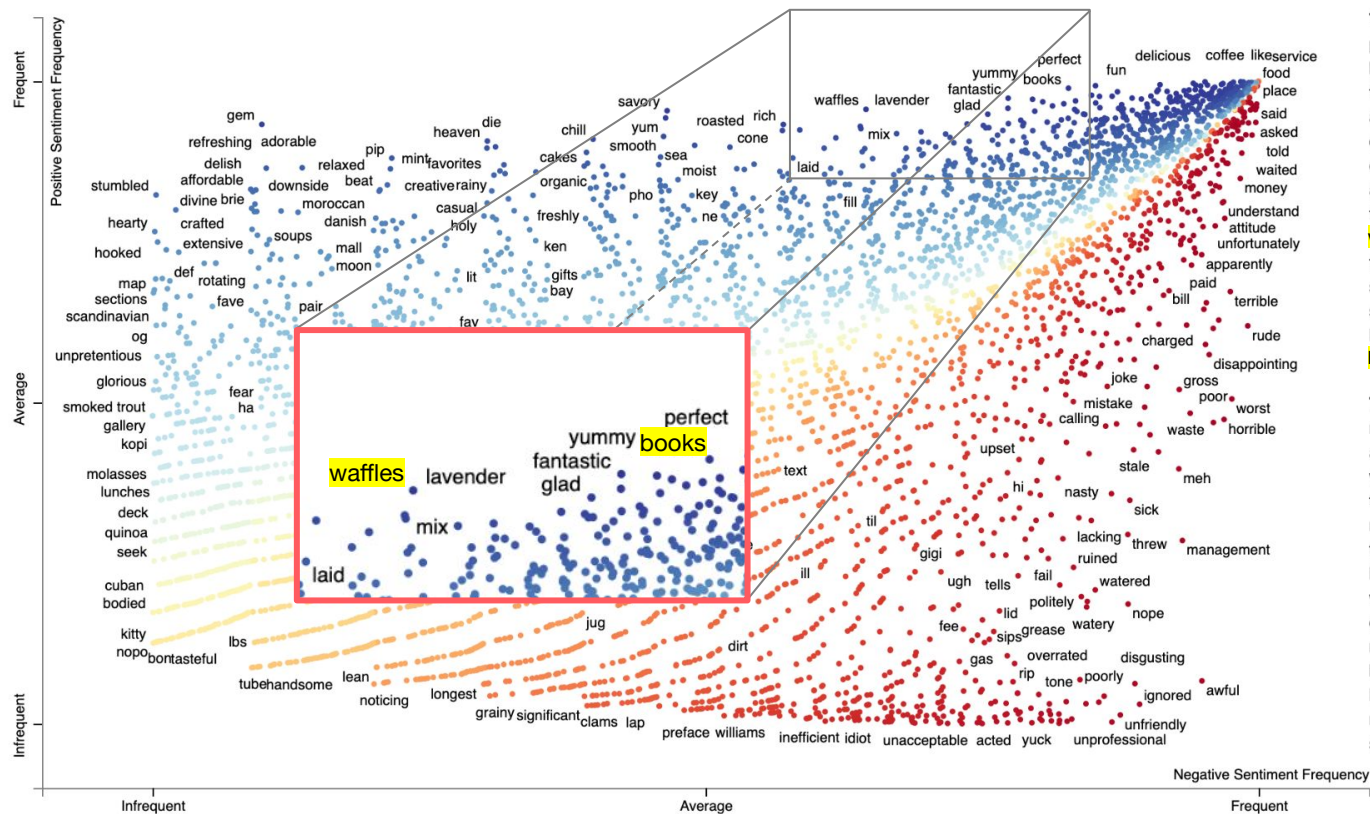
Service

Three out of nine topics on service when topic modeling on negative sentiment reviews only

Food

Less prominent than service for negative sentiment reviews

Scattertext

Positive Sentiment document count: 12,890; word count: 1,102,769
Negative Sentiment document count: 2,110; word count: 249,225

Topic Modeling

Final Topics

Vectorizer: TF-IDF

Topic Modeler: NMF

Num Topics: 9

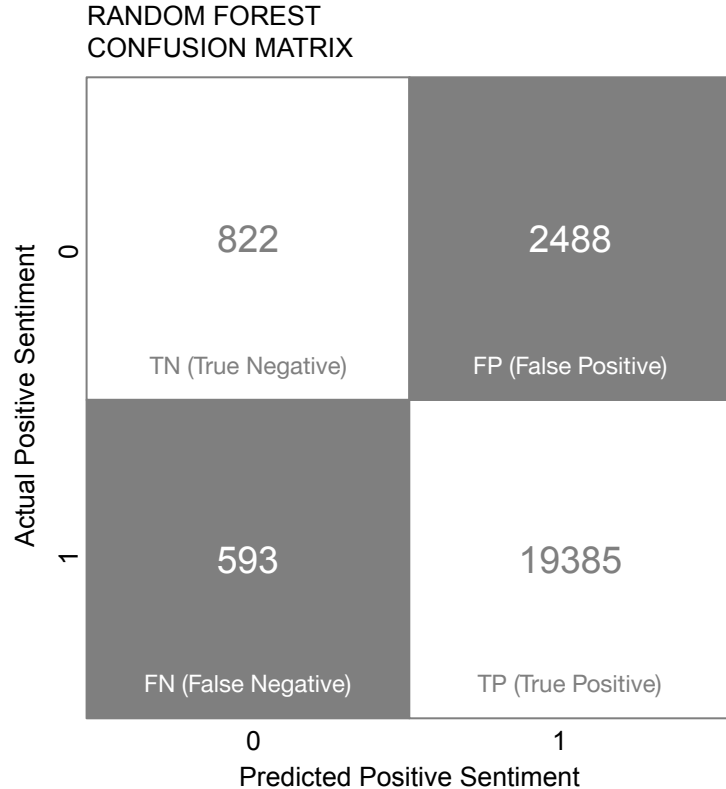| Topic	
	Savoury Food
	Doughnuts/Voodoo
	Coffee/Barista
	Tea
	Ice Cream
	Books/Powell's
	Location/Atmosphere
	Service
	Pastries/Dessert

SUPERVISED LEARNING
Classification

Random Forest

Confusion Matrix

Target: Sentiment (+ or -)

Features: 9 topics

Data: Topic strengths from NMF doc-topic matrix

Output: model is better at assigning positive than negative sentiment reviews

RANDOM FOREST
CONFUSION MATRIX

Random Forest

Feature Importance
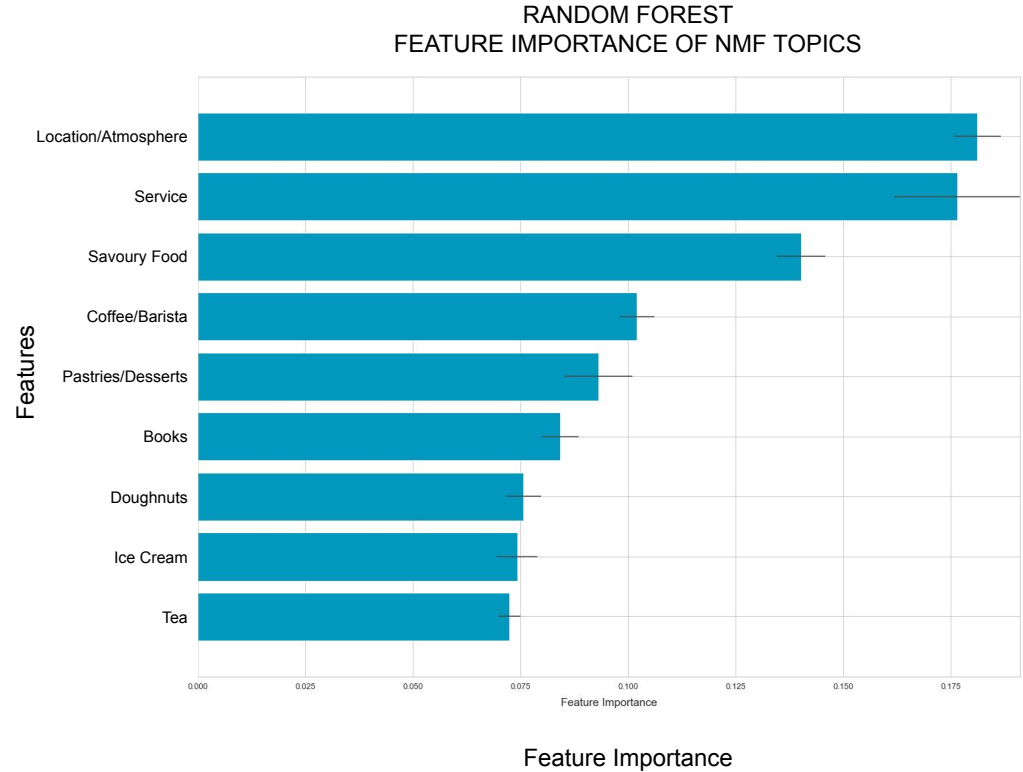
Location/Atmosphere

Setting the right ambiance to enjoy a cup of coffee

Service

Strengthening the core of customer experience

Food

Curating the selection of complementary items

RANDOM FOREST
FEATURE IMPORTANCE OF NMF TOPICS

Feature Importance

Summary Metrics

Random Forest Model

0.93

F1 score

0.89

Precision

0.88

Accuracy

0.97

Recall

Final Recommendations to the Yelp Product team

- **Atmosphere:** support café owners in improving the ambiance and decor by building an Interior Design Playbook.

- **Staff Training**: convey the importance of regular trainings on both the human interaction and the operational elements of service (e.g. waiting time) to maximize customer experience.

- **Coffee:** café owners should maximize the quality of coffee beans and expand the variety of coffee types to build an appealing menu.

- **Food:** prioritize building a diverse supply of savoury food items, possibly by setting up partnerships with other local businesses.

Questions?