Practical Business Applications of Sarcastic Tweet Detection Using Natural Language Processing and Machine Learning

William C. Webb

The Sarcasm Problem

- Sarcasm implies the opposite of the literal meaning
- Speakers show sarcastic intent through body language or tone of voice
- Written communication lacks cues for indicating sarcasm

Sentiment Analysis

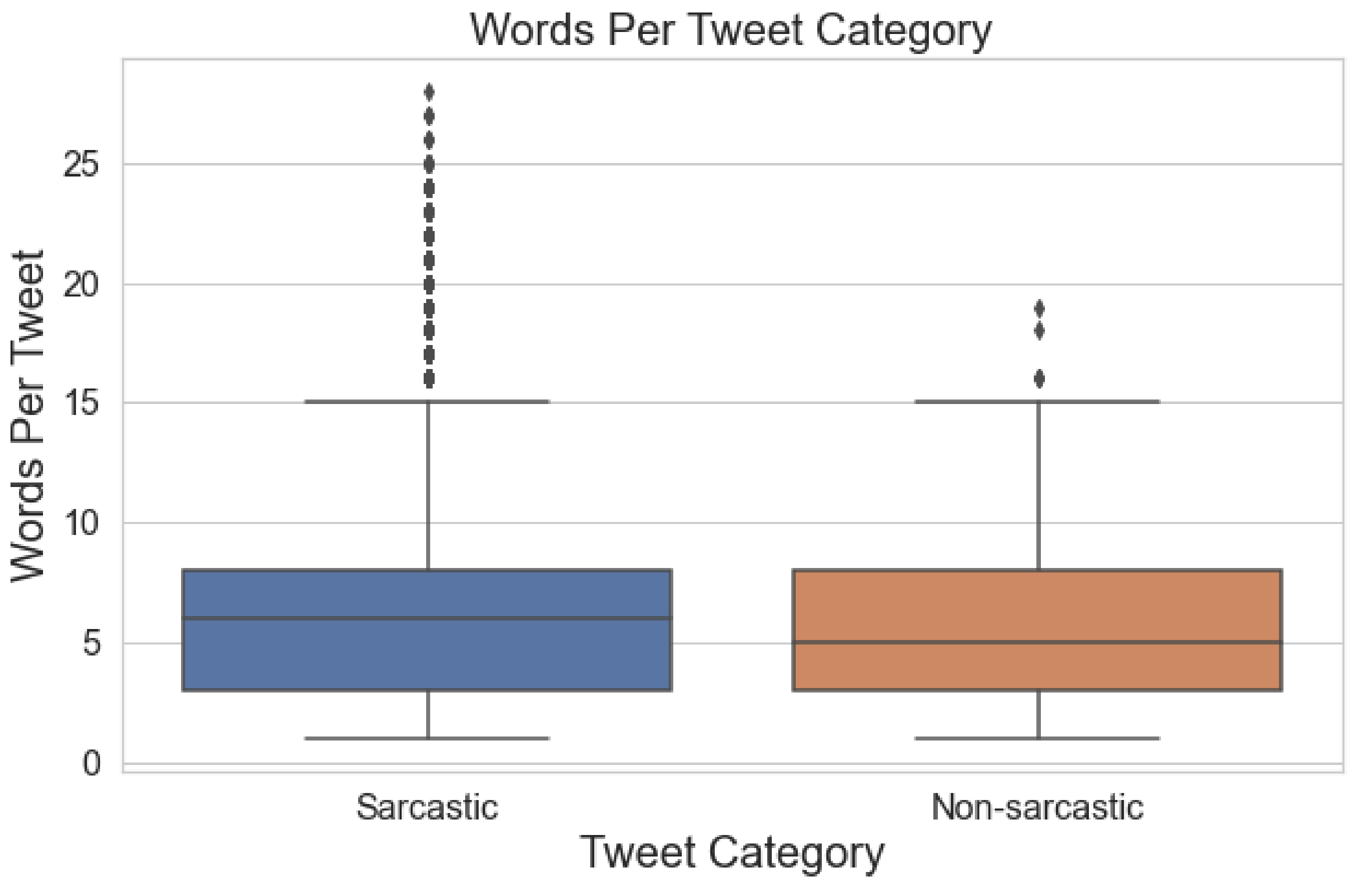
- Gauges public attitudes towards a particular topic
- Important business tool
- Social media often used to measure attitudes towards products
- Sarcastic posts represent potential large source of error

Detecting Sarcasm

- Developed a model to detect sarcastic tweets
- Used NLP and machine learning
- Tuned the model for hypothetical business applications
- Tested the model on a novel sentiment analysis of tweets

Sarcastic Posts Contained More Words

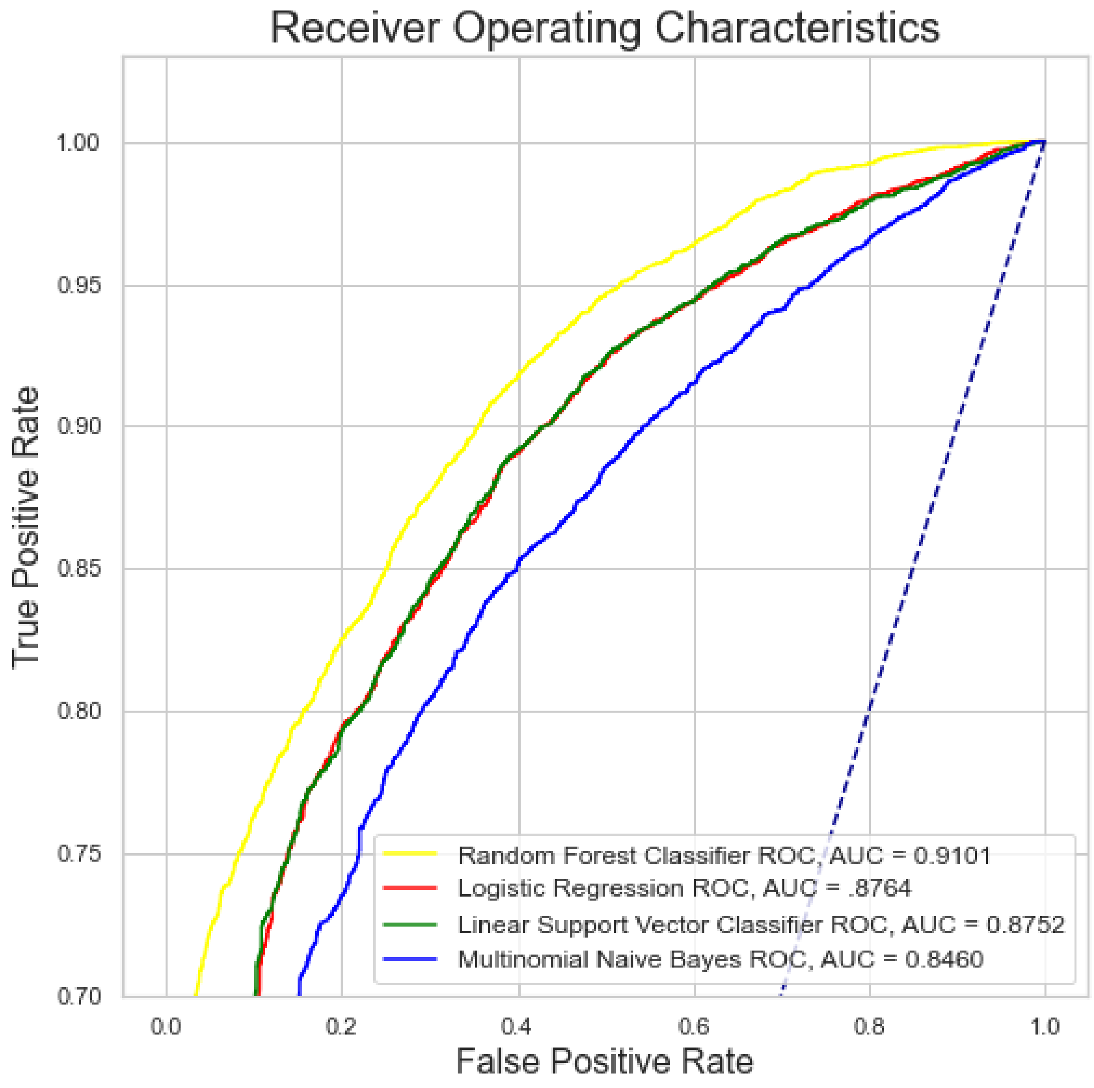
- Collected > 100,000 non-sarcastic & sarcastic tweets
- Processing steps including removing stop words, etc.
- Processed tweets contained 292,000 total words



Sarcasm Probability Scores

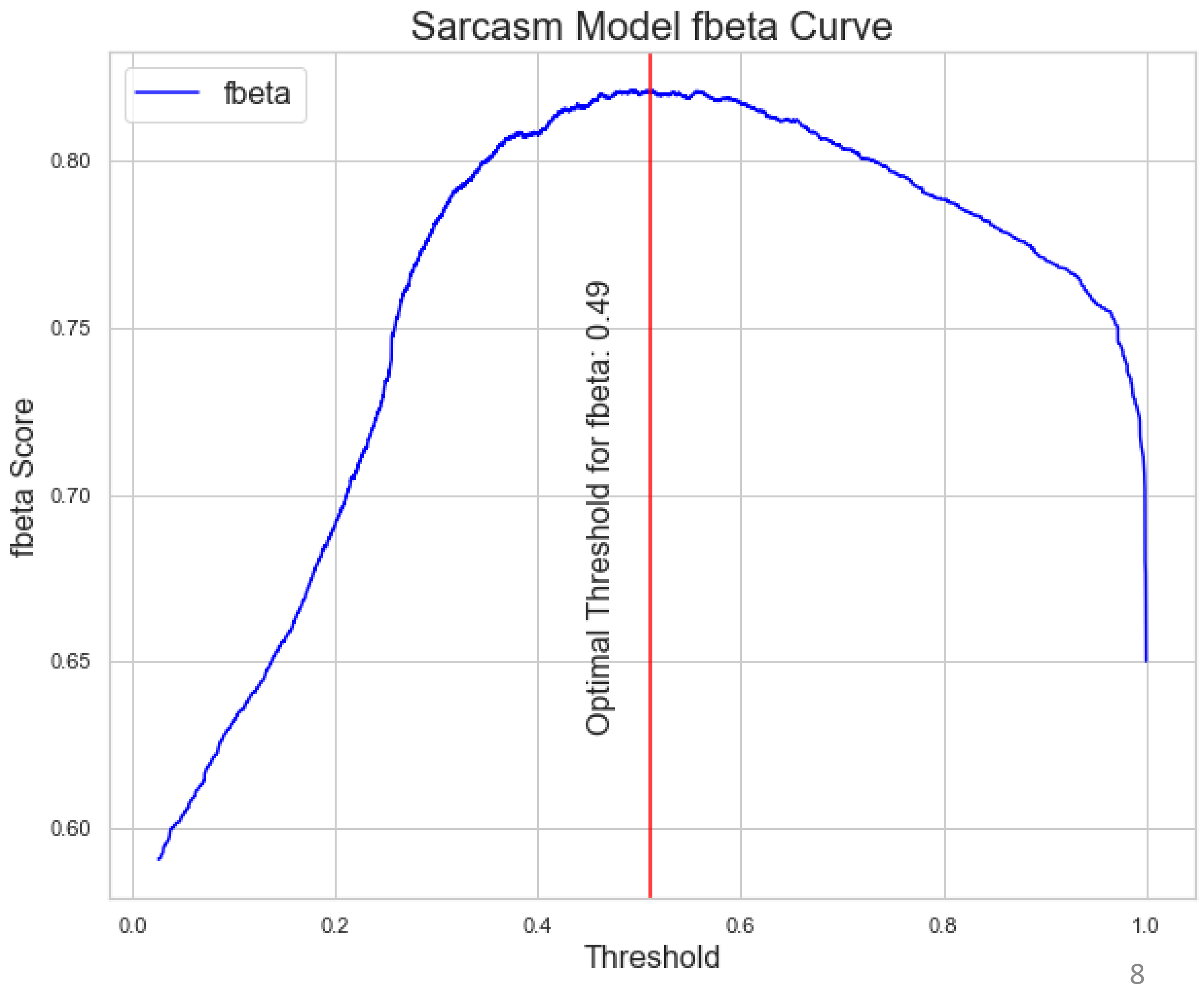
Most sarcastic words		Least sarcastic words	
resist	0.9517	biden	0.0219
belt	0.945	prepared	0.0312
rose	0.933	coronavirus	0.0346

Random Forest Performed Best for Predicting Sarcastic Tweets



Business Scenario One: Minimize Mislabeling Sarcastic Tweets

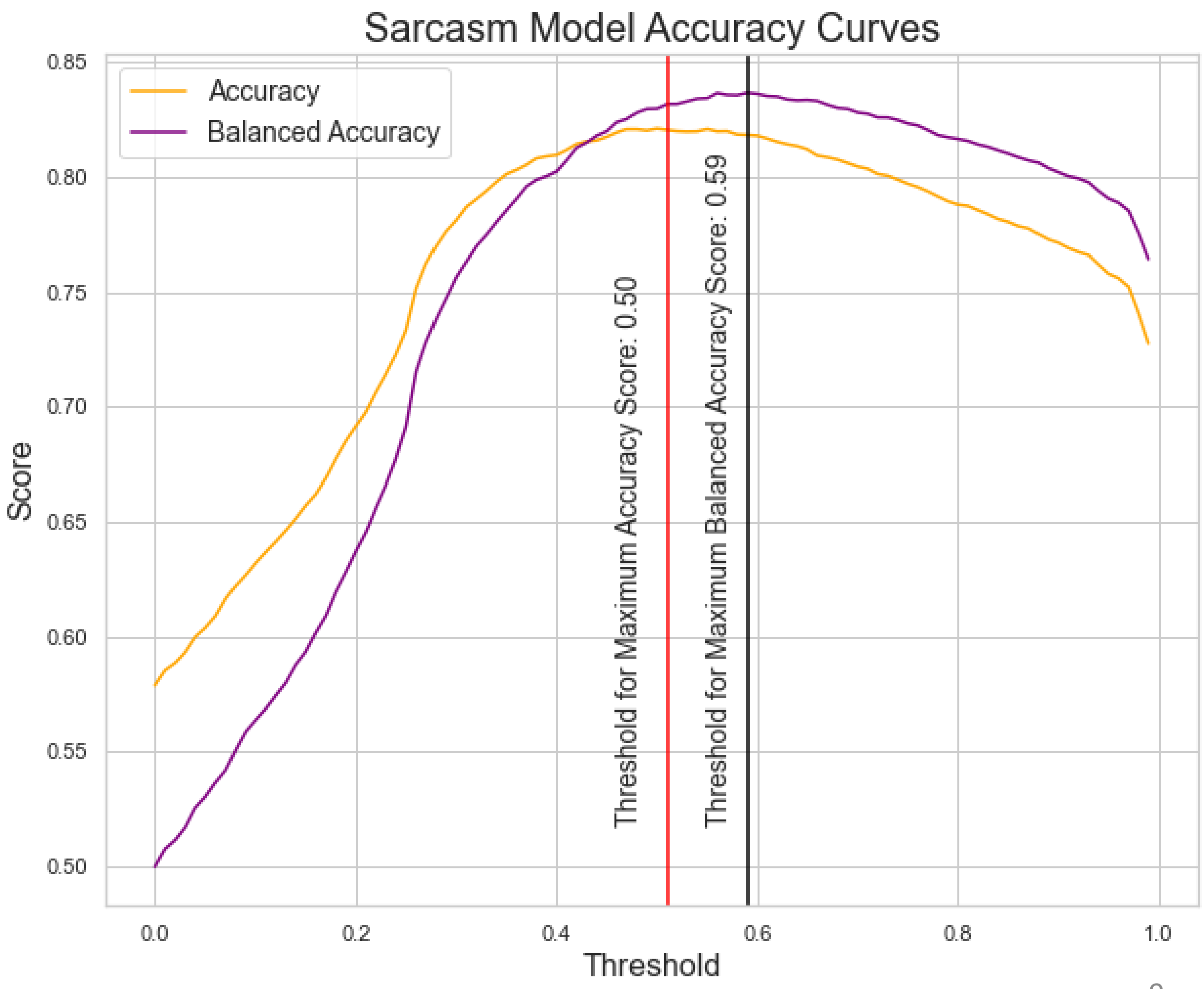
- Adjusted fbeta parameter to emphasize recall
- Determined the threshold for maximizing fbeta
- Reduced false negatives by 834 tweets



Business Scenario Two: Capture an Accurate Product Sentiment

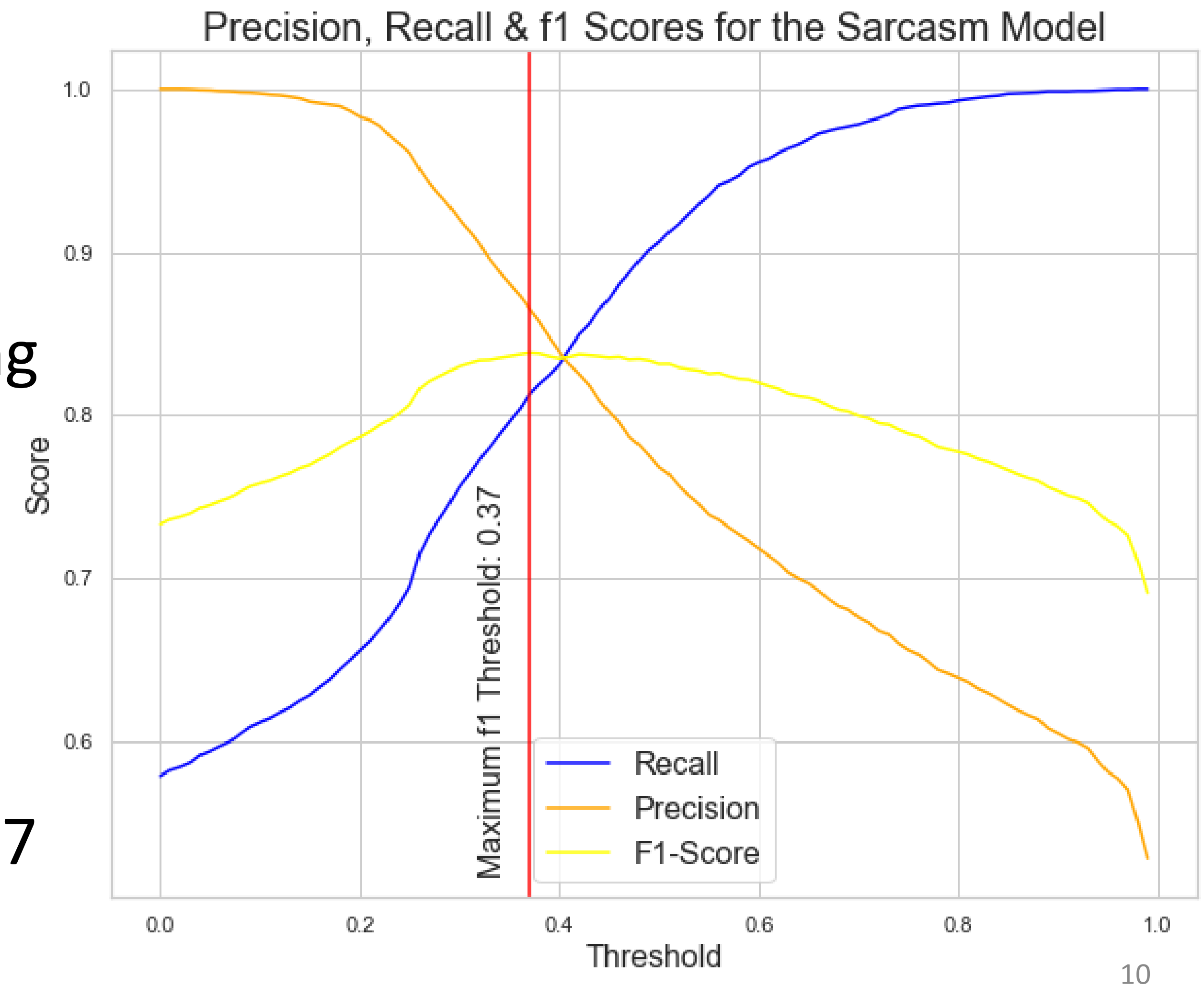
- Calculated accuracy and balanced accuracy scores for a range of threshold values

- Determined thresholds that maximized both scores



Business Scenario Three: Balance False Negatives & False Positives

- Calculated f1 scores for range of threshold values
- Determined threshold maximizing f1 score
- Reduced false negatives by 549
- Increased in false positives by 677



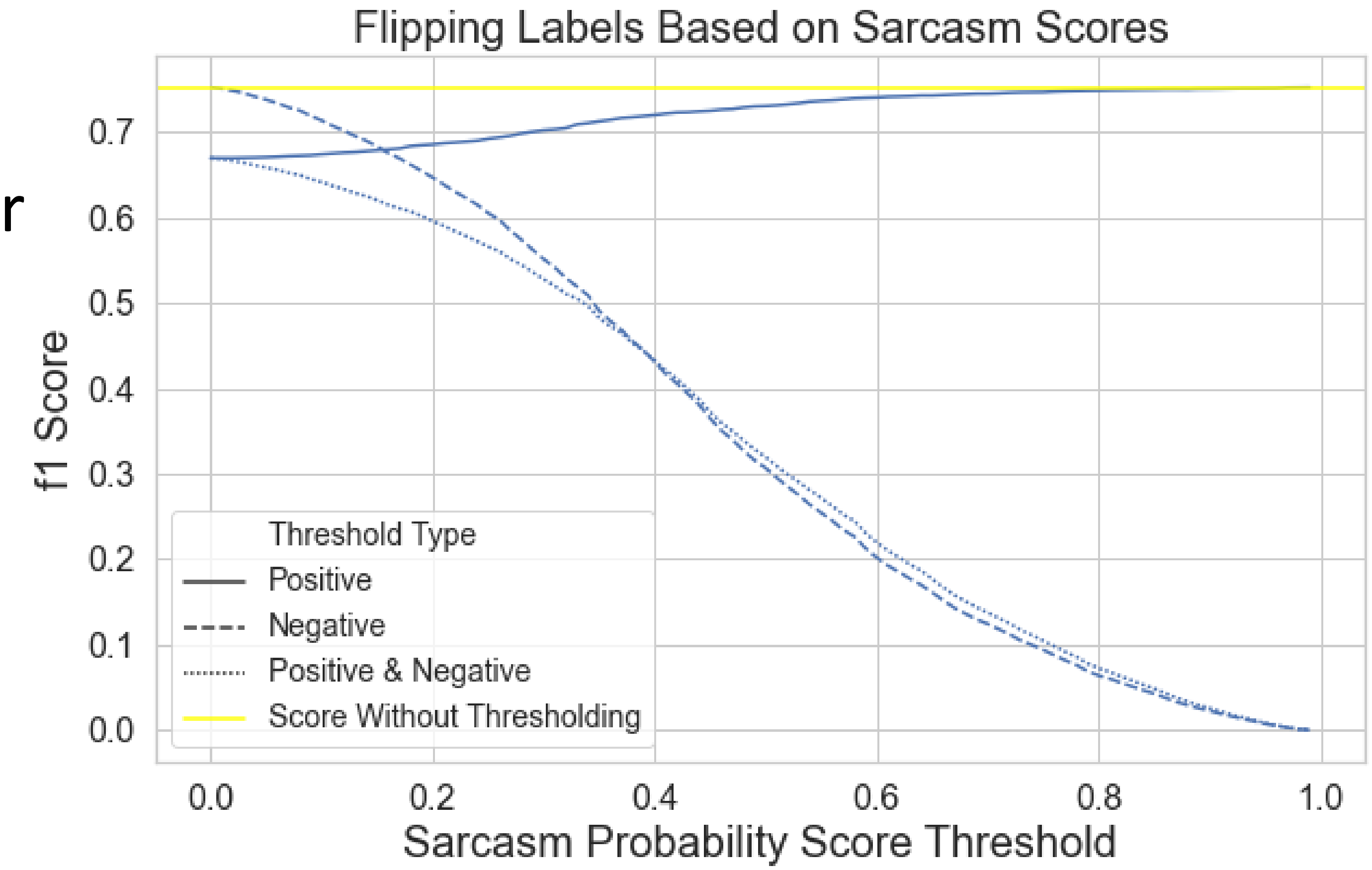
Business Scenario 4: Improve a Novel Sentiment Analysis

- Use sarcasm model to improve an existing sentiment analysis
- Sentiment140 dataset
- Used naïve Bayes to predict positive or negative tweets
- Changed labels based on sarcasm scores for range of thresholds

Sarcasm Model Failed to Improve Some Scores

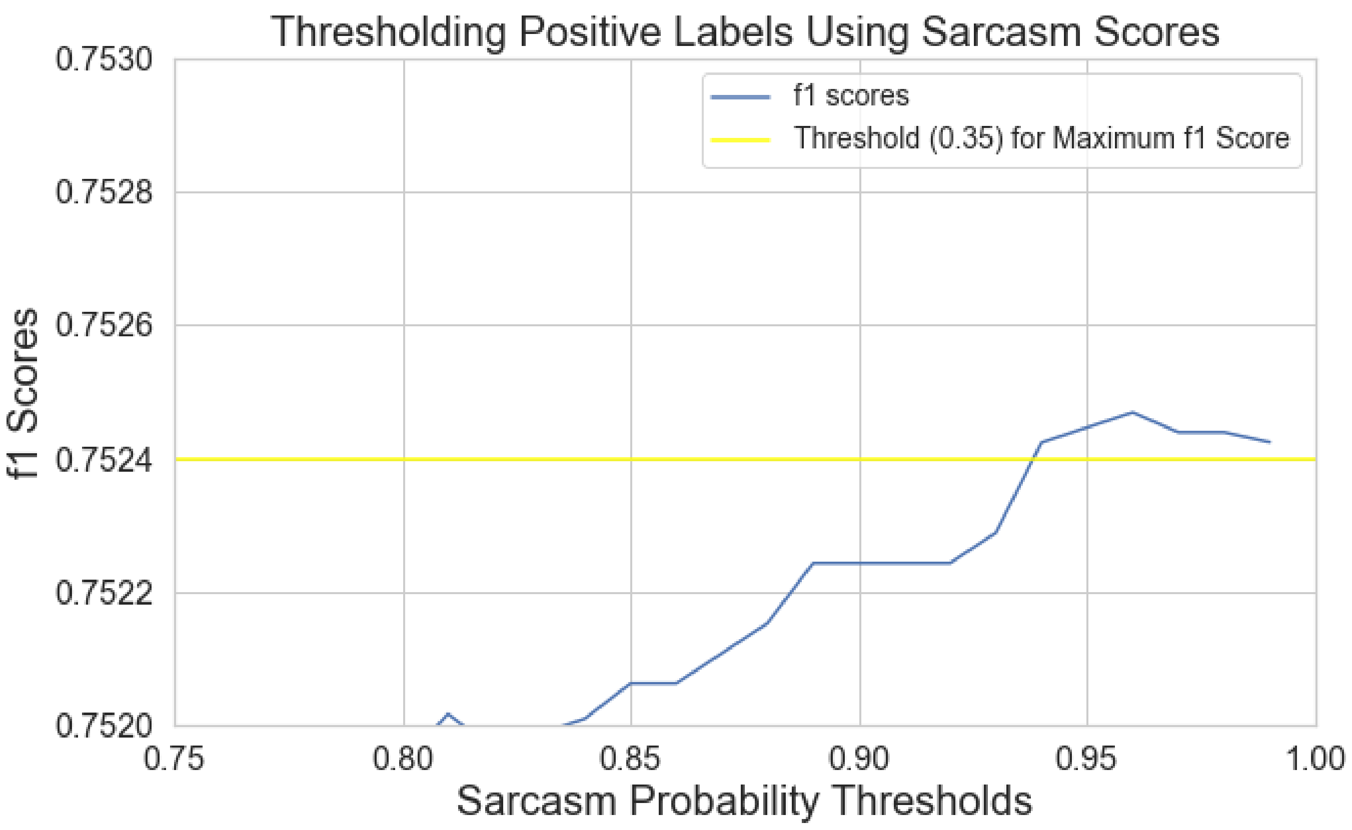
- No improvement in f1 score for flipping negative labels alone

- No improvement in f1 score for flipping both positive & negative labels together



Slight Improvement Flipping Positive Labels

- Very small improvement in f1 score for flipping positive labels



Utility of the Sarcasm Model

- Performs well for identifying sarcastic tweets
- Adjusts to meet different business goals:
  - Minimizing mislabeling sarcastic tweets
  - Maximizing accuracy
  - Balancing the risk of mislabeling sarcastic and non-sarcastic tweets

Can the Sarcasm Model Improve a Sentiment Analysis?

- Probably depends on the dataset

- Sematic (dis)similarity during classification:
  - Negative / Positive (sentiment140) not same as sarcastic/non-sarcastic (sarcasm data)

- Similarity in data collection methods:

  - Sentiment140 classified using emoticons:
  - Sarcasm classified using hashtags: #Sarcastic

  - Sarcasm data filtered for geographic location
  - Sentiment140 data not filtered for location

Web App

- Created a web app
- thesarcometer.com
- Demonstrates sarcasm model
- Compares sarcasm score of a novel username with celebrity Twitter accounts