

REDDIT:

Predicting men's and women's fashion post with Natural Language Processing.



01

Reddit & Process

What is Reddit and what was our process

02

Models

Which predictive models were used and their performance.

03

Graphs

Graphs representing top words in posts.

04

Conclusions

Based on the data science process, what can we conclude.

TABLE OF CONTENTS

01 What is Reddit?

"Reddit is a network of communities based on people's interests." -reddit.com

"A subreddit is a specific online community, and the posts associated with it, on the social media website Reddit." -dictionary.com/e/slang/subreddit/

Selected subreddits:

1. mensfashionadvice - 2.3 million
2. womensfashionadvice - 1.1 million





- Gathered data from Reddit using a pushshift API.
- Utilized from over 10,000 posts, and extracting 'title' and 'selftext' for analysis.
- After data cleaning and preprocessing ran & evaluated Random Forest and Multinomial Naive Bayes models.
- Determined the top words in 'title' and 'selftext'

02 MODELS

- **Random Forest Classifier:**
An ensemble algorithm.
- **Multinomial Naive Bayes:**
A classification model commonly used with text/counting words.





PERFORMANCE OF **TITLE** FOR PREDICTIVE ANALYSIS

- The baseline accuracy of the model was 50.1%
- The RandomForest model's accuracy score was 72.6%
- The Multinomial NB model's accuracy score was 72.7%

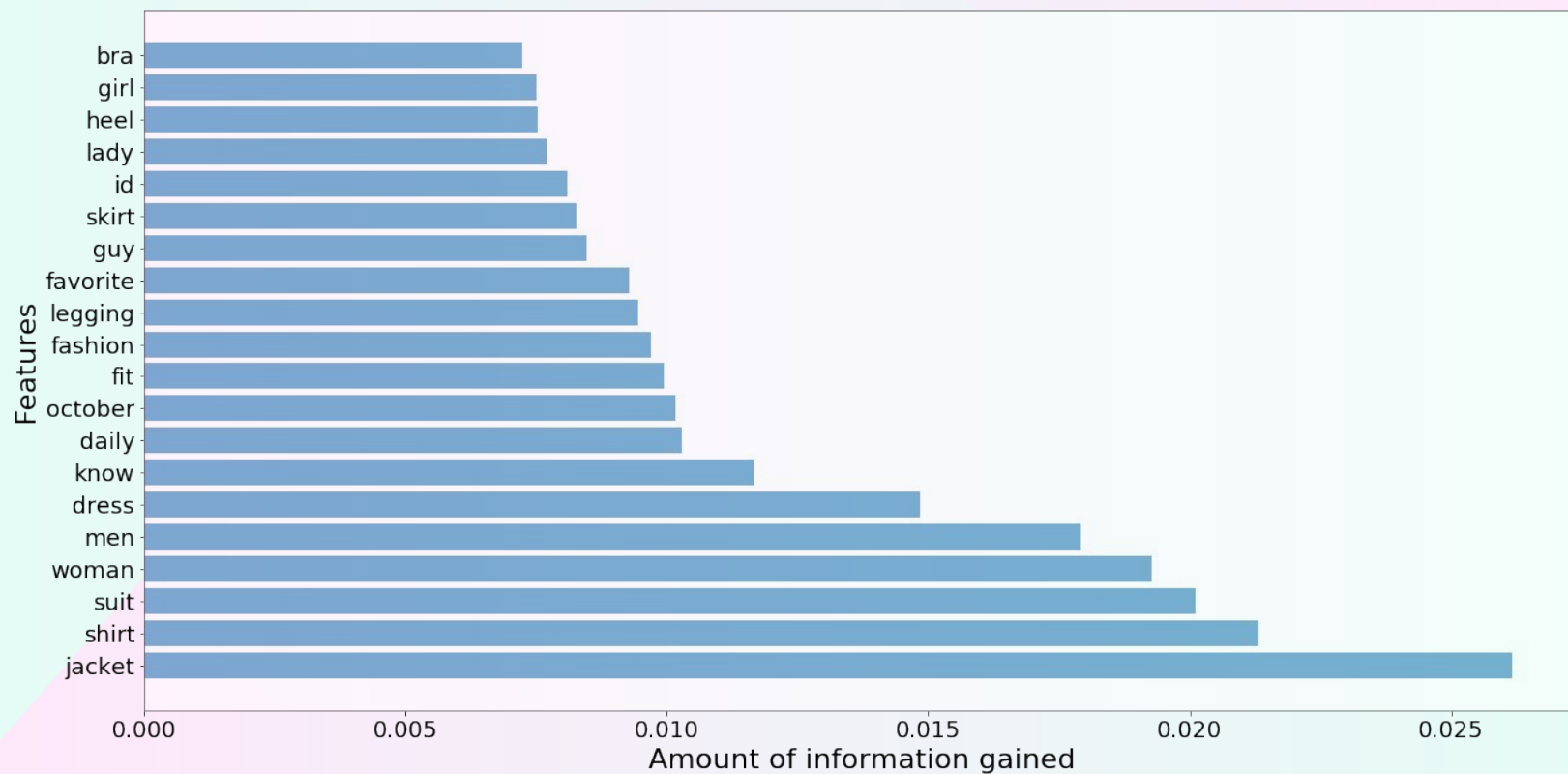
PERFORMANCE OF **SELFTXT** FOR PREDICTIVE ANALYSIS

- The baseline accuracy of the model was 50.1%
- The RandomForest model's accuracy score was 63.7%
- The Multinomial NB model's accuracy score was 55.7%

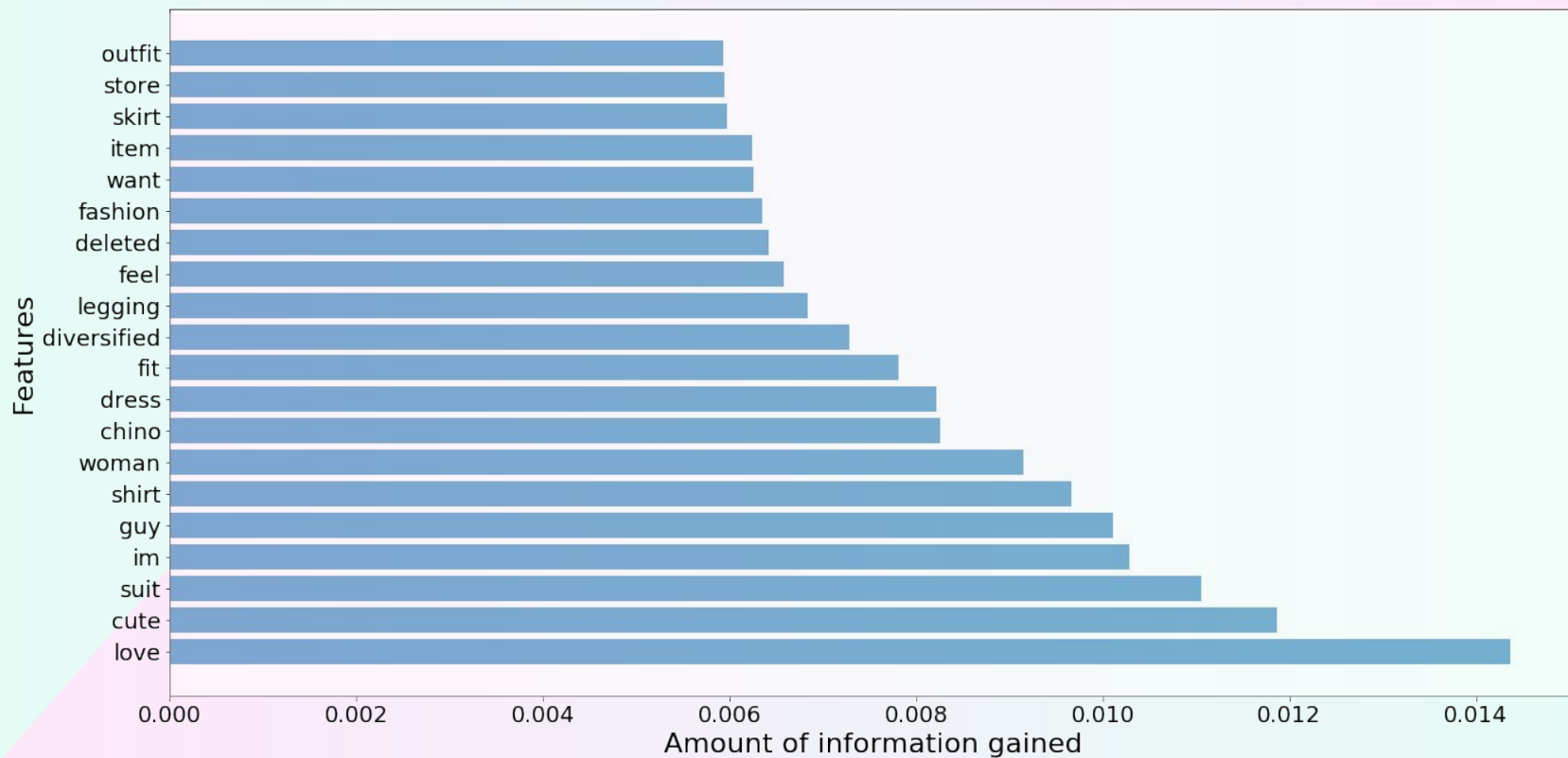


GRAPHS 03

TOP 20 FEATURES FROM TITLE



TOP 20 FEATURES FROM SELFTEXT



04

CONCLUSIONS

Although Reddit has millions of active users, it might not be the target area to invest in advertising due low performing models.

If we do choose to use it, using the word Daily(i.e. OOTD), strong titles

Crossing boundaries by marketing the same items in both genders.



Key words such as:

- jacket
- **shirt**
- **suit**
- **legging**
- october

■ ■ ■

might point us in the right direction.

- Coming out with our main collection in September/October
- Focusing on shirts, suits, jackets and leggings for both genders as well as an androgynous collection.
- People engaging in fashion forums tend to take more fashion risks





THANKS!

Does anyone have any questions?
anapetronijevic56@gmail.com
703-389-6922