

FLATIRON ANALYTICS

SXSW TWEET ANALYSIS

ANALYSTS:
AUNG SI
JF ROBERTS
MICHAEL ROMANSKI

The SXSW 2023 logo is a vibrant orange diamond shape with a thick green border. The text "SXSW 2023" is written in bold black letters, rotated diagonally across the center. A black arrow points downwards from the top right corner of the diamond. The background features abstract green and grey shapes.

OUR TEAM



AUNG SI



JF ROBERTS



MICHAEL ROMANSKI

OVERVIEW

◆ **BUSINESS PROBLEM**

◆ **DATA &
PROCESSING**

◆ **EVALUATION METRIC**

◆ **PREDICTIVE MODELS**

◆ **MODEL
EVALUATION**

◆ **RECCOMENDATIONS
& NEXT STEPS**



BUSINESS PROBLEM



SXSW DEVELOPING A NEW APP
FOR THE FESTIVAL:
APPLE STORE vs GOOGLE STORE



SENTIMENT ANALYSIS ON TWEETS
FROM FESTIVAL



RECCOMEND SPECIFIC OPERATING
SYSTEM FOR THE APP



THE DATA

ROUGHLY 9,000 TWEETS FROM
DATA WORLD

CROWDSOURCED SENTIMENTS
FROM FESTIVAL GOERS

DATA IMBALANCE:
+60% OF DATA IS NEUTRAL



DATA PROCESSING - NLP

◆ DROP IRRELEVANT CHARACTERS

EX. '#', '@', '12345'

◆ DROP IRRELEVANT WORDS

EX. 'SXSW' - APPEARS IN ALL TWEETS

◆ DROP "NEUTRALS"

ASK OUR AUDIENCE !

**"FORGOT MY IPHONE FOR SXSW. ANDROID ONLY.
KNIFE TO A GUN FIGHT"**

POSITIVE OR NEGATIVE

DATA PROCESSING - NLP

◆ DROP IRRELEVANT CHARACTERS

EX. '#', '@', '12345'

◆ DROP IRRELEVANT WORDS

EX. 'SXSW' - APPEARS IN ALL TWEETS

◆ DROP "NEUTRALS"

ASK OUR AUDIENCE !

"FORGOT MY IPHONE FOR SXSW. ANDROID ONLY.
KNIFE TO A GUN FIGHT"

POSITIVE OR NEGATIVE

NEUTRAL!

EVALUATION METRIC

ACCURACY

**HOW ACCURATELY
MODEL PREDICTS
POSITIVE TWEET**

RECALL - FALSE NEGATIVE

**TWEET INCORRECTLY
CLASSIFIED AS NEGATIVE**

PRECISION - FALSE POSITIVE

**TWEET INCORRECTLY
CLASSIFIED AS POSITIVE**

MODELS

BINARY CLASSIFIER
POSITIVE VS NEGATIVE

ACCURACY
85%

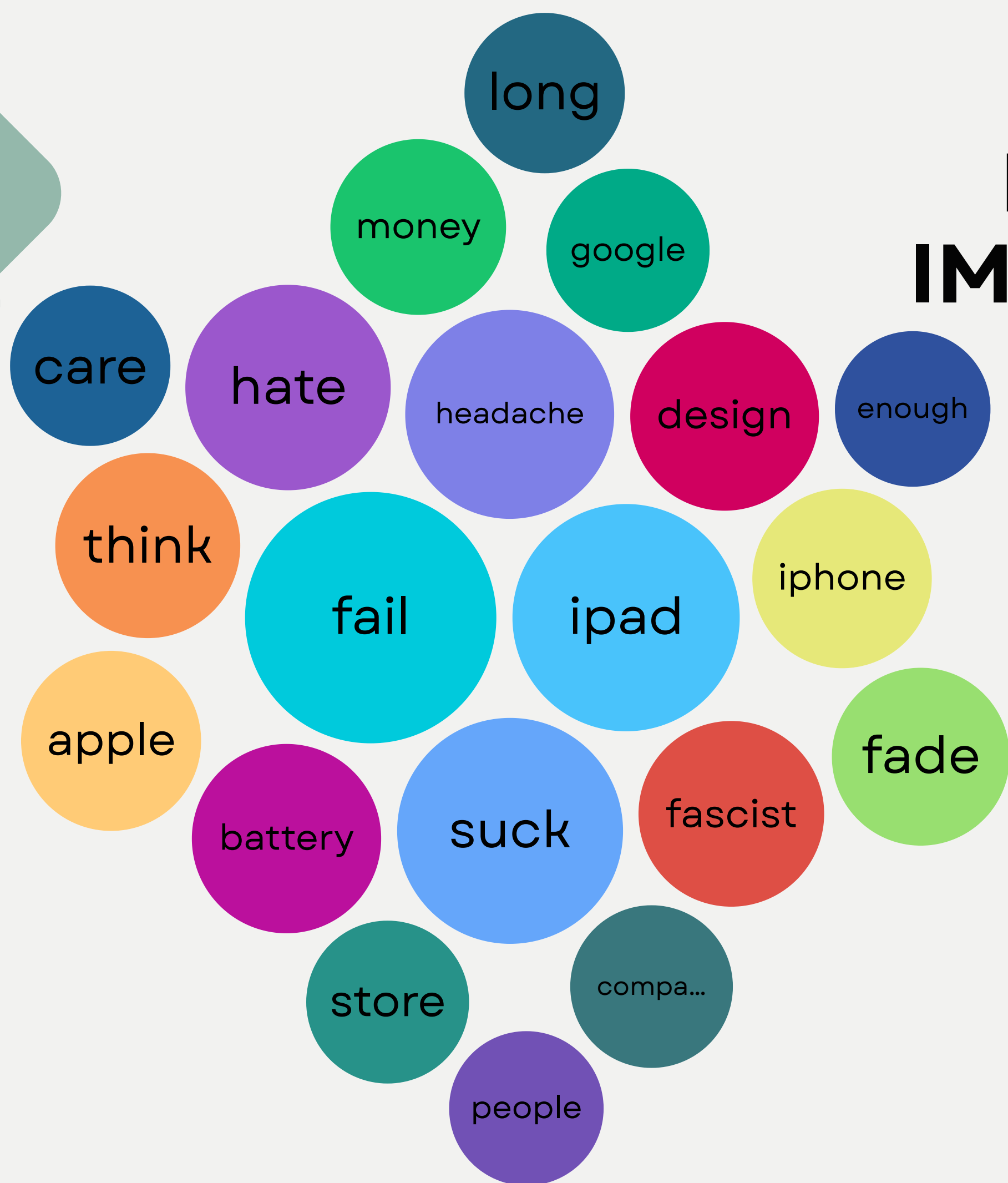
TERNARY CLASSIFIER
POSITIVE VS. NEGATIVE VS. NEUTRAL

ACCURACY
64%

NEURAL TERNARY CLASSIFIER
POSITIVE VS. NEGATIVE VS. NEUTRAL

ACCURACY
68%

FEATURE IMPORTANCE



RECOMMENDATIONS

APPLE OVER GOOGLE

USE OUR MODEL TO CLASSIFY
NEXT YEARS' SENTIMENTS

COMPARE PREDICTIVE VS CROWD
SOURCED

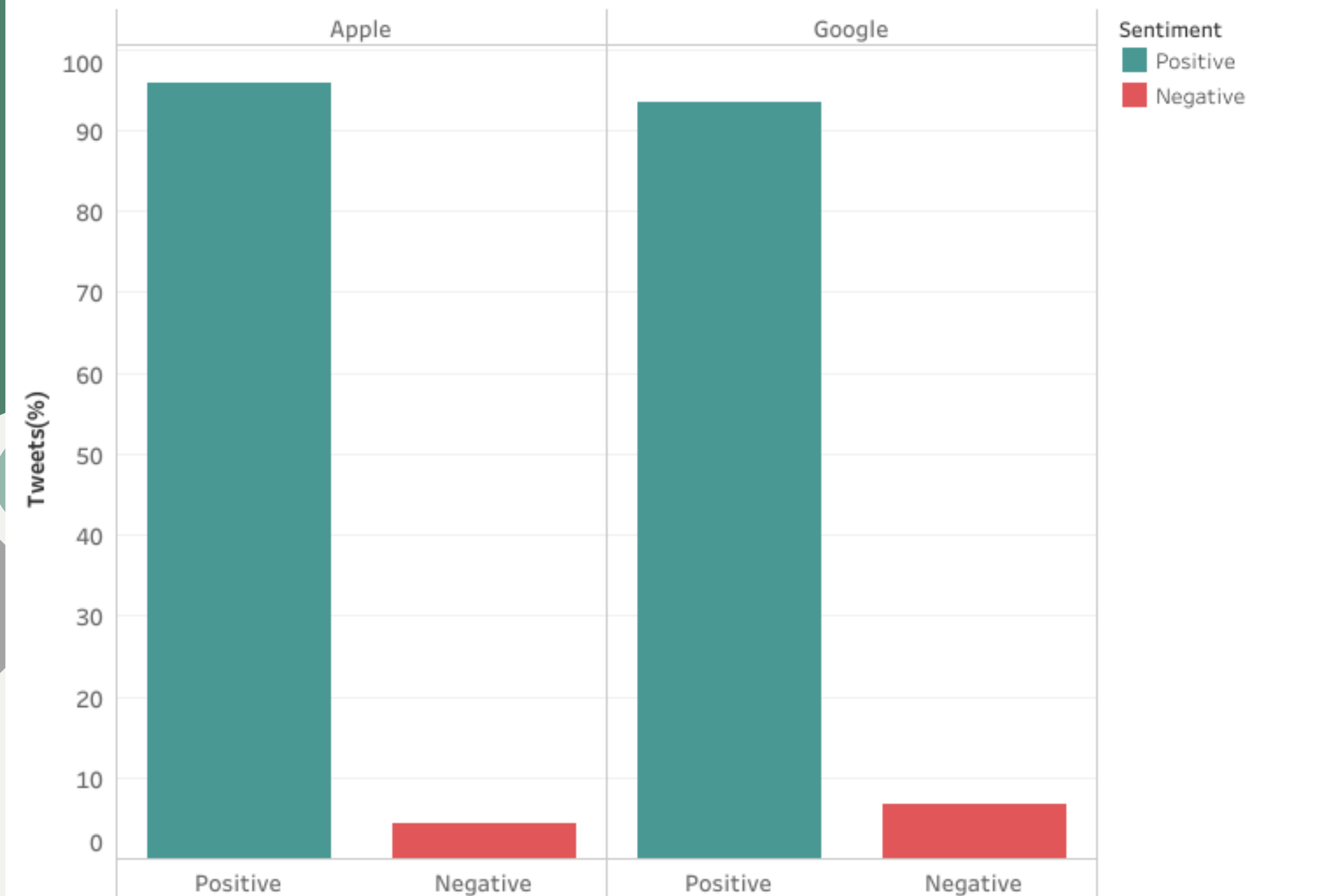
Brand Sentiment



NEXT STEPS

FLATIRON ANALYTICS

Brand Sentiment(Neutral Data Test)



SOURCE MORE RELIABLE DATA &
A LARGER DATA SET

DIG INTO MIS-REPRESENTED
“NEUTRAL” TWEETS

HOW TO HANDLE SENTIMENTS
THAT ARE BOTH POS & NEG

THANK YOU



jfbr1283



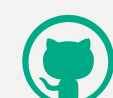
jfwholehealth@gmail.com



m-romanski



michael.romanski4@gmail.com



aungsi



aungsi.99@gmail.com