Hello

Amazon Fine Foods Review

DSCI6004 - Eli

Overview

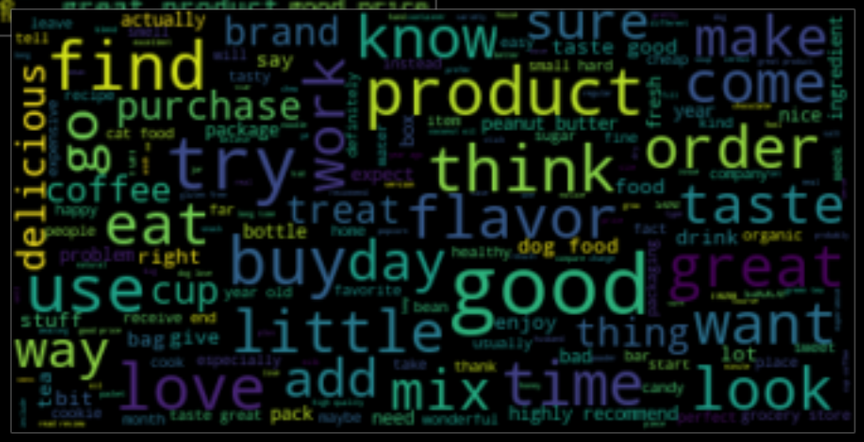
- Data cleaning & exploration
- Data processing
- Evaluation
- Results

Data cleaning

- Limiting sample
- Helpful >= 0.9
- Spacy - astype(str),batch_size=10000,n_threads=20
- Readability - textstat
- Sentiment - (pos,neg,neu,compound) SentimentIntensityAnalyzer
- Lemma- spaces, urls, #s, emails, stop words, and proper nouns

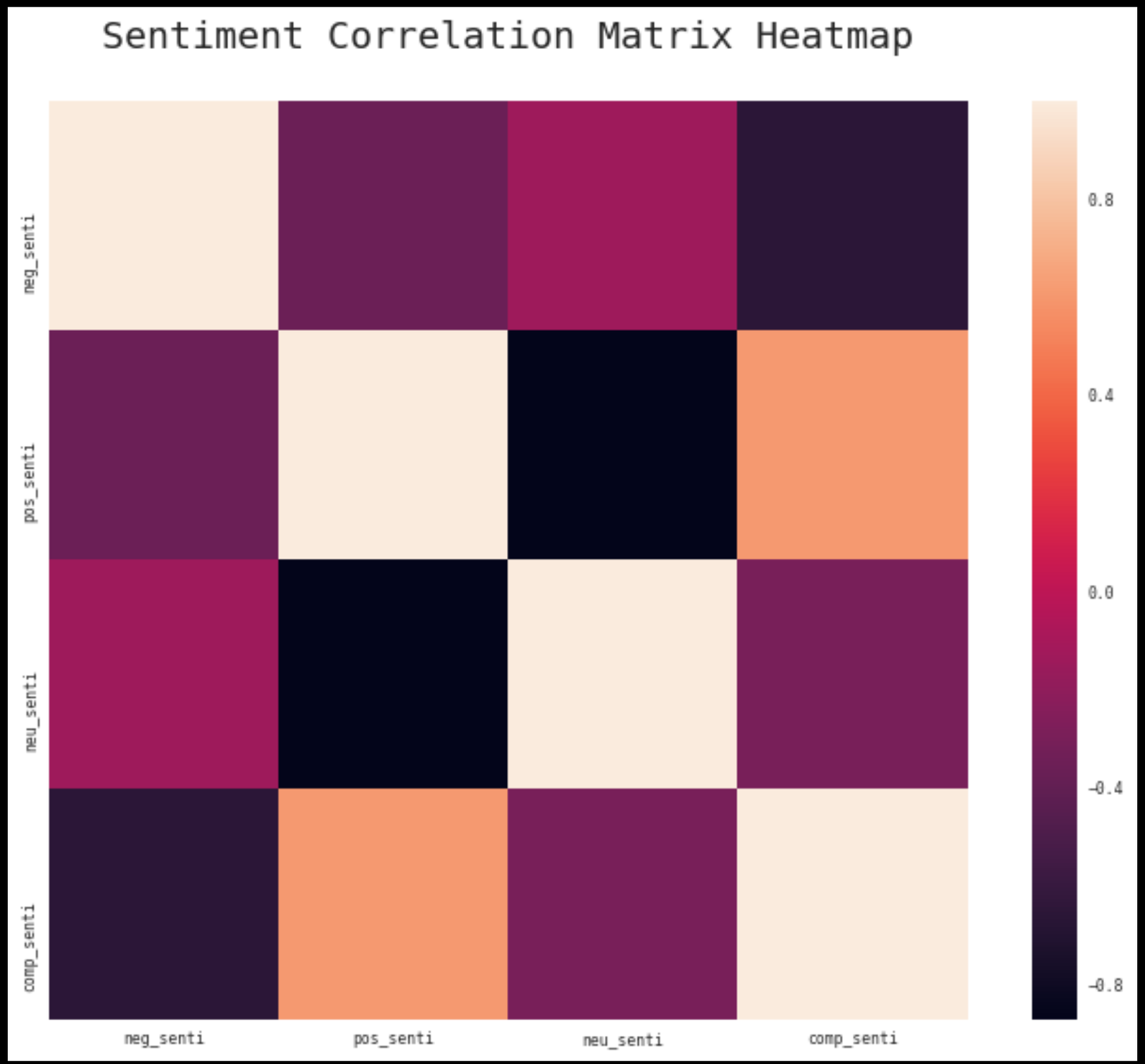
Word cloud

problem
kind
expect
deliciou
taste great
fresh
make
use
right
taste good
way
usually
lot
mix
tasty
gluten free
dog love
enjoy
come
stuff
happy
maybe
peanut butter
say
try
think
look
go
thing
actually
fact
bad
package
dog food
find
work
coffee
give
little
know
need
purchase
sure
year old
grocery store
want
highly
recommend
cup
nice
perfect
candy
green tea

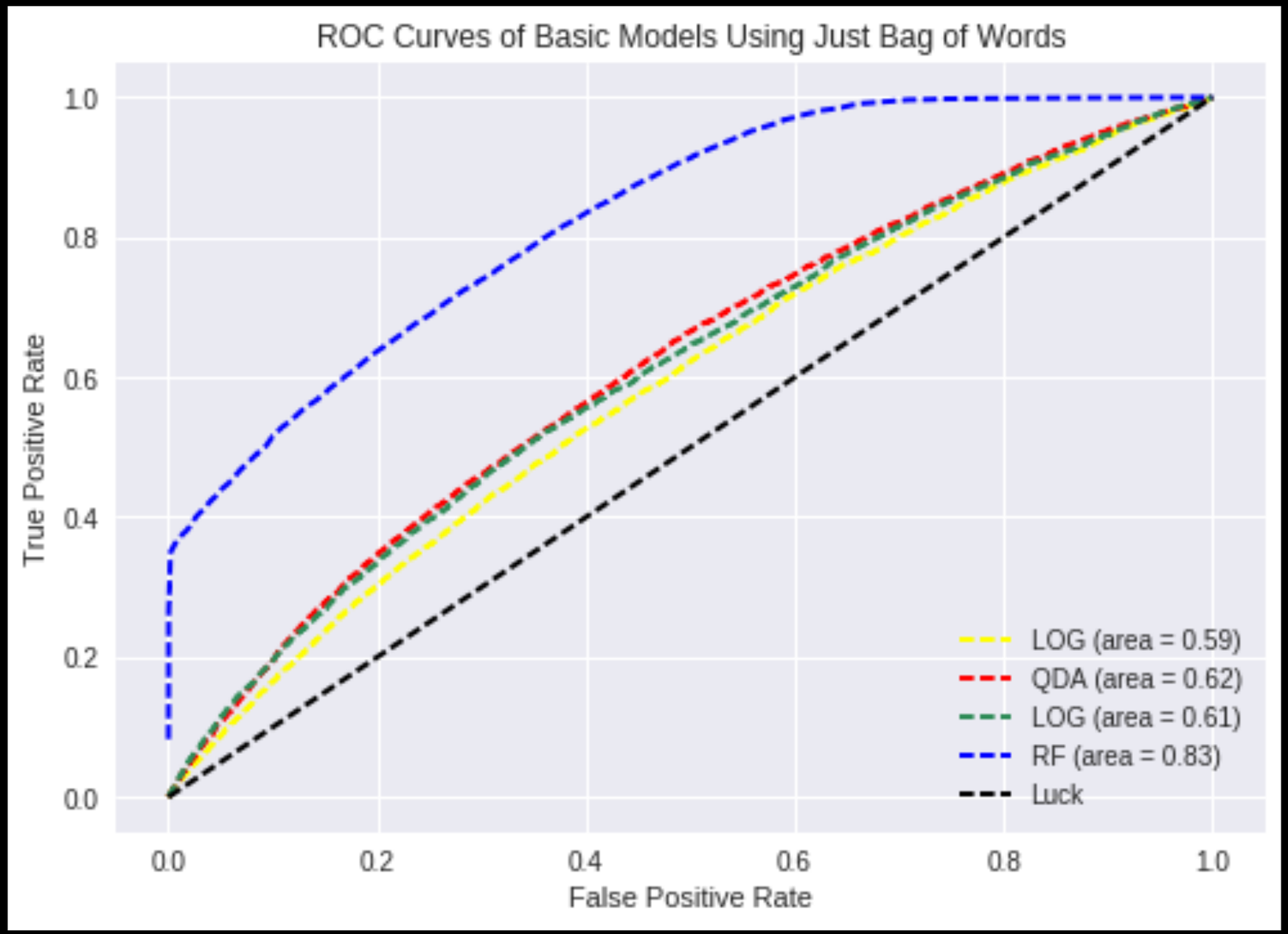

brand
know
sure
make
find
product
come
purchase
work
delicious
go
try
think
order
coffee
eat
treat
flavor
taste
buyday
good
great
use
cup
little
want
way
thing
love
add
mix
time
look

deliciou
way
give
order
problem
try
taste great
look
mix
go
find
enjoy
cup coffee
sure
come
make
purchase
grocery store
know
taste
good
want
need
peanut butter
dog food
time
right
dog love
work
say
stuff
think
bad
use
thing
highly
recommend
actually

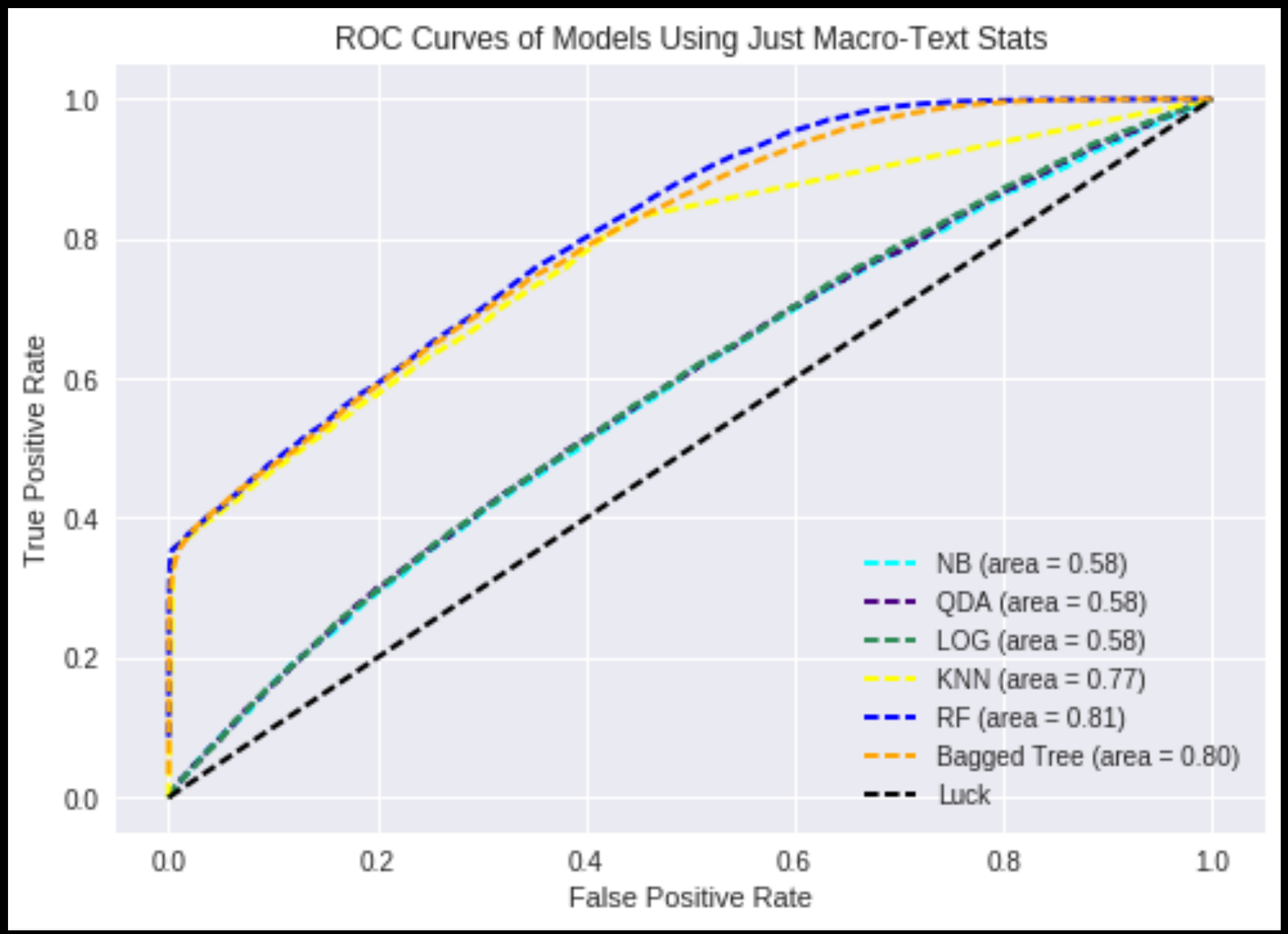
Correlation



Model - Bag of Words tf-idf



Model - Macro level tf-idf



Time to run models

Winner: KNN
Loser: RF