

PREDICTING POLITICAL AFFILIATION

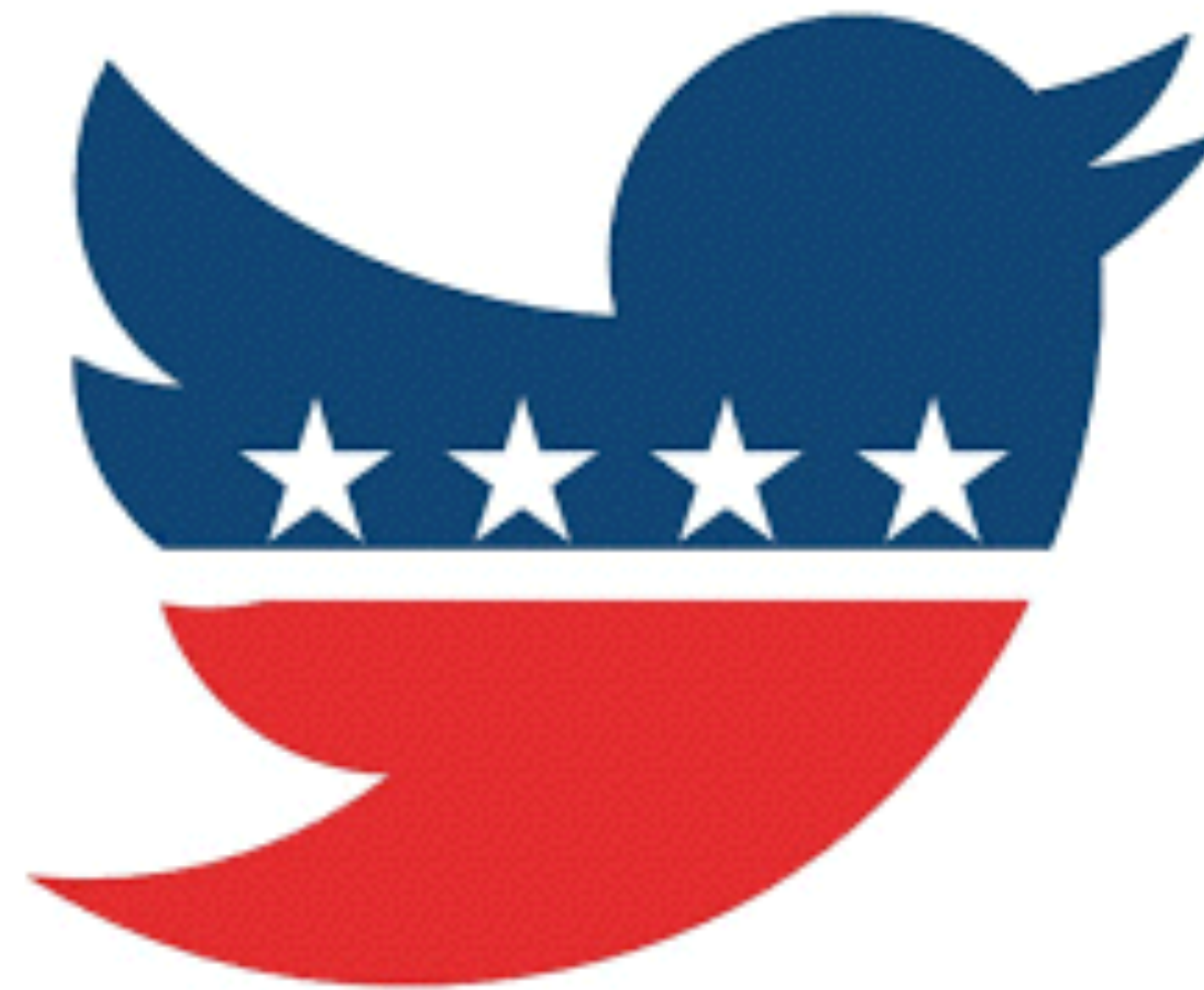
Using Twitter Data



Tiffany Moeller

February 20, 2018

2020ELECTION



HELP
PLEASE!

DELIVERABLES

#1

PREDICTIVE
MODEL



#2

INTERACTIVE
VISUALIZATION



DATA SOURCES

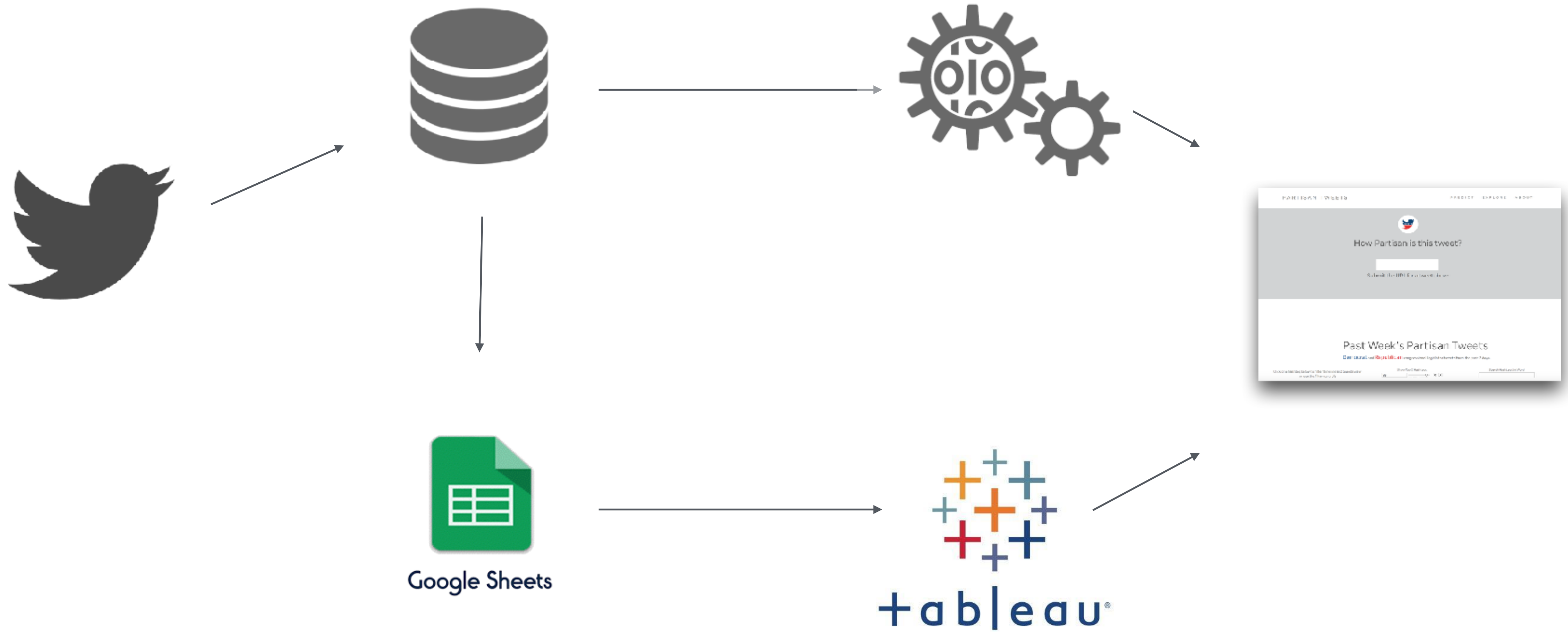
CURRENT
LEGISLATORS

TWITTER
API

POSTGRES DB
ON AWS



DATA PIPELINE



ANATOMY OF A TWEET

METADATA

- Day of Week Created
- Hour of Day Created
- Length of Tweet
- General Sentiment
- Retweets
- Followers



TEXT

- Top 1750 Most Common Words
- Includes Hashtags

METADATA
MODEL

Gradient Boost
Classifier

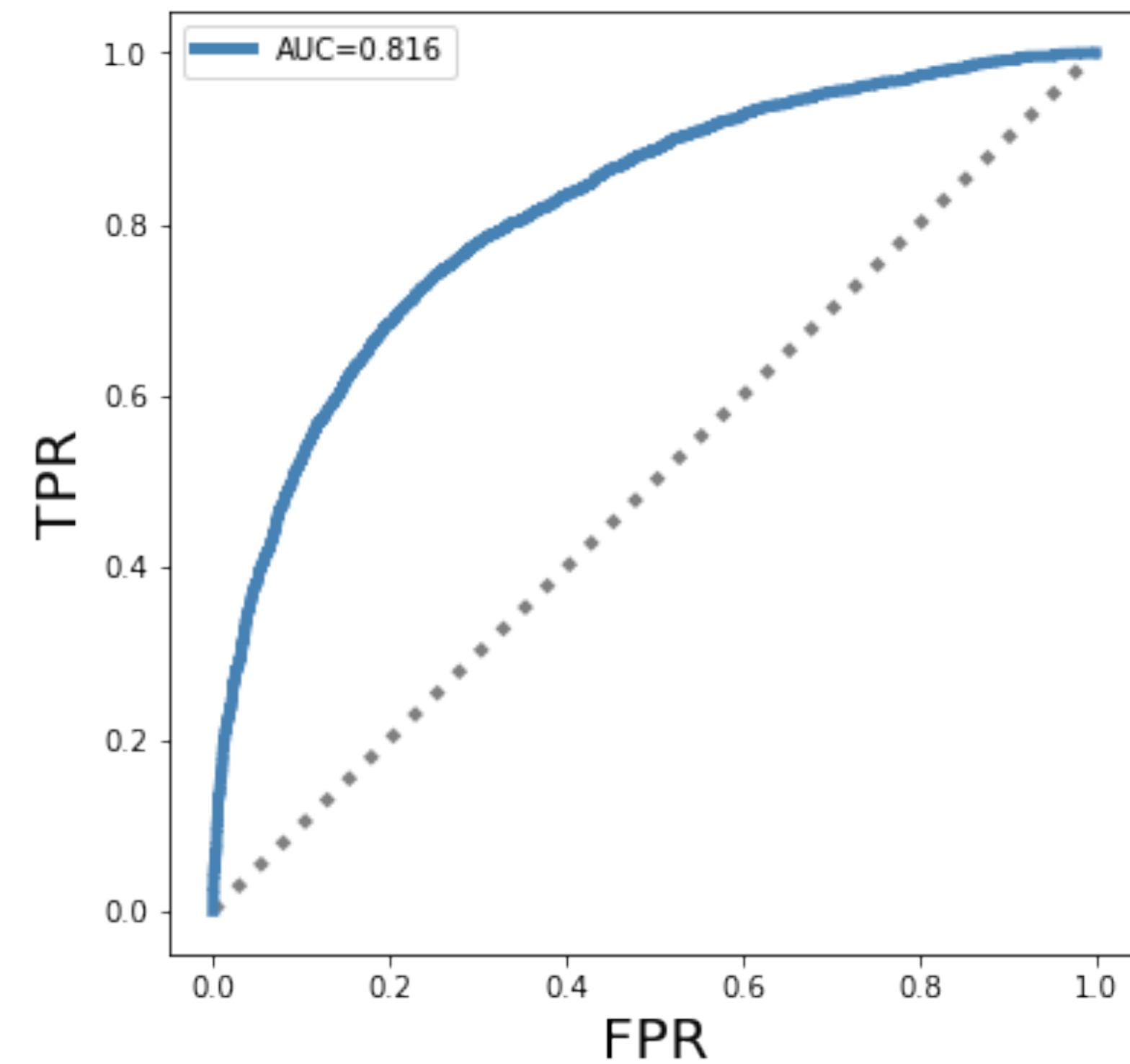
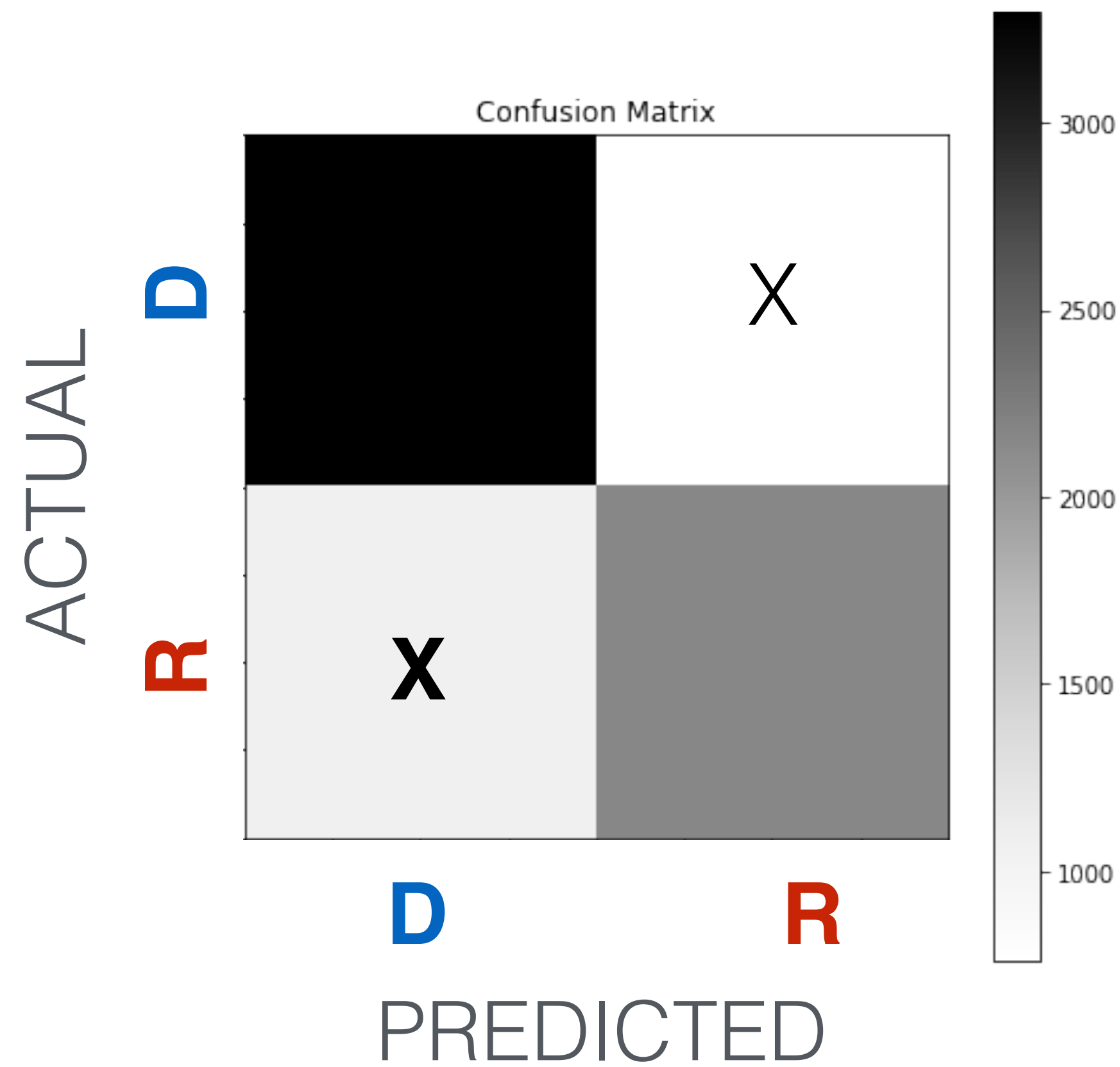
TEXT
MODEL

Bernoulli Naive Bayes
Classifier

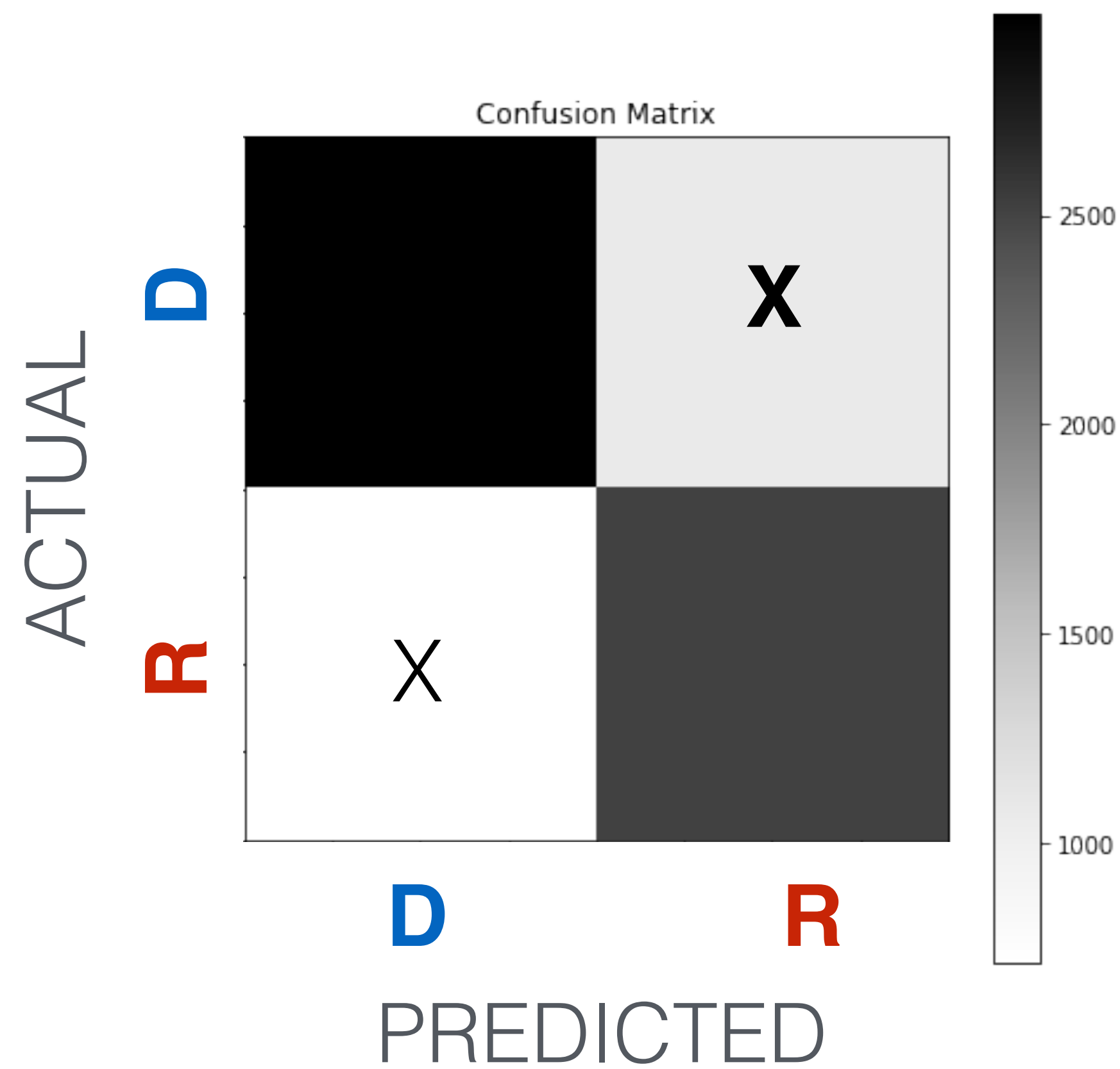
METADATA

Gradient Boost Classifier

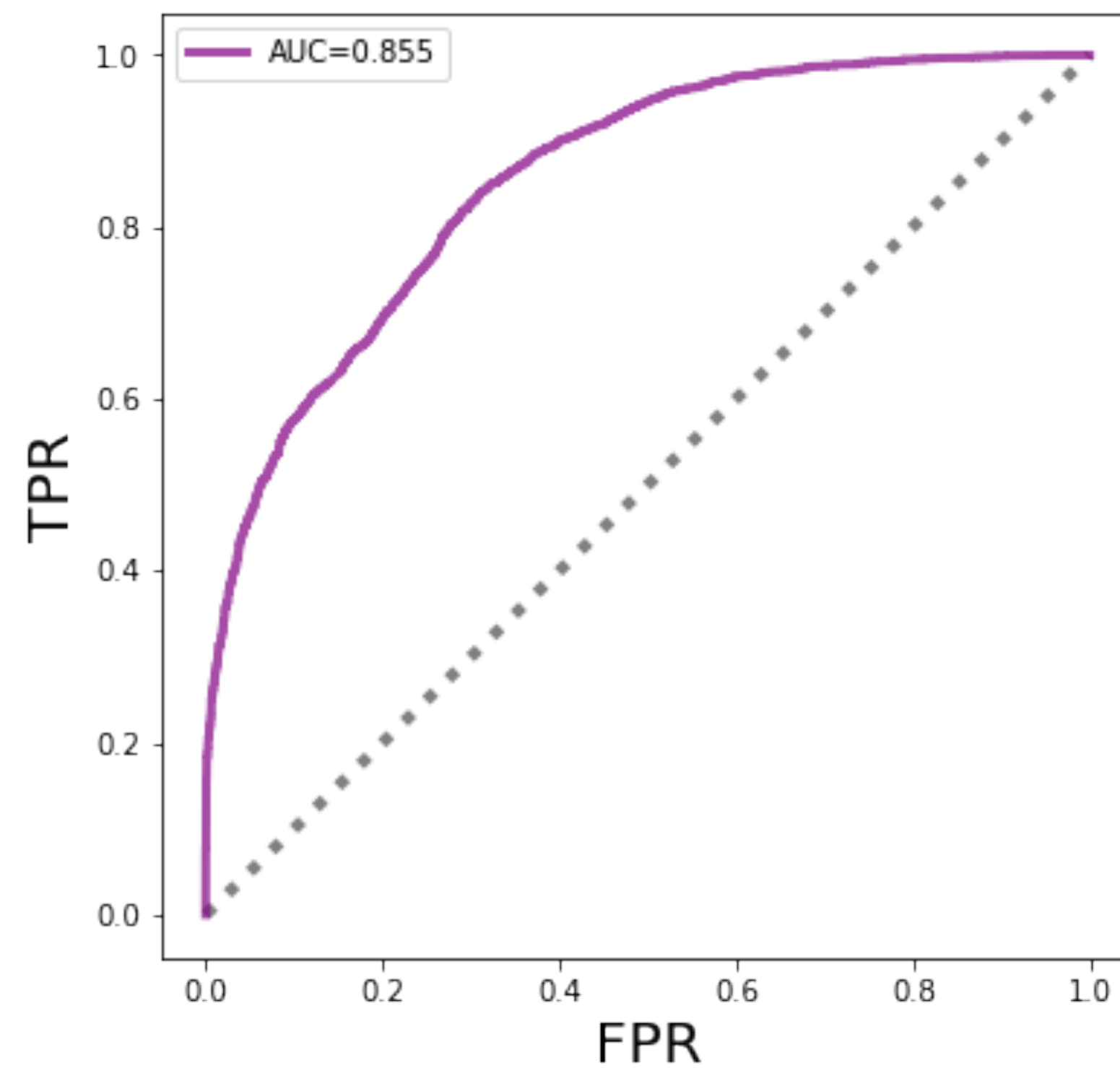
ROC AUC: **.816**



TEXT Naive Bayes Classifier

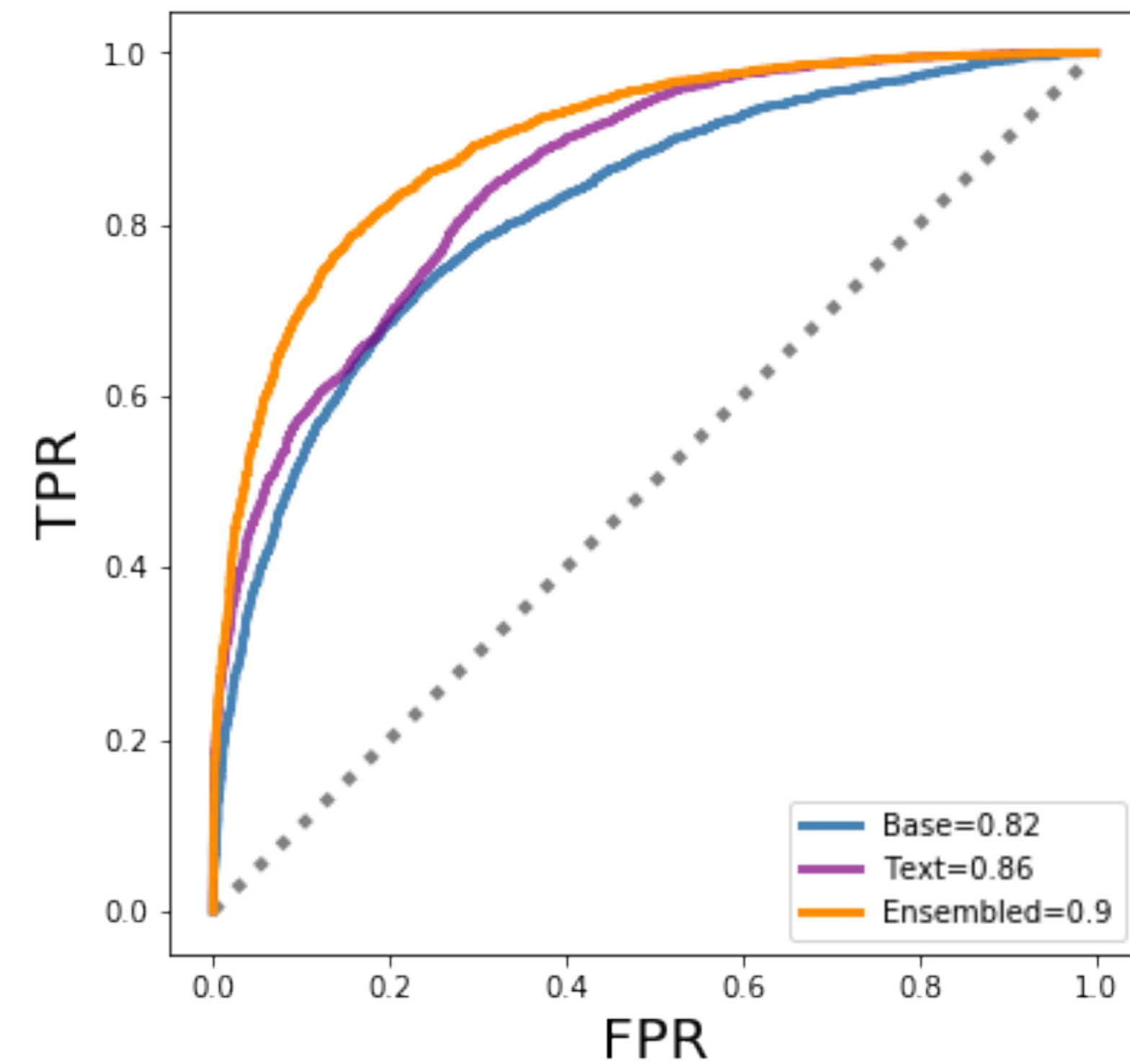
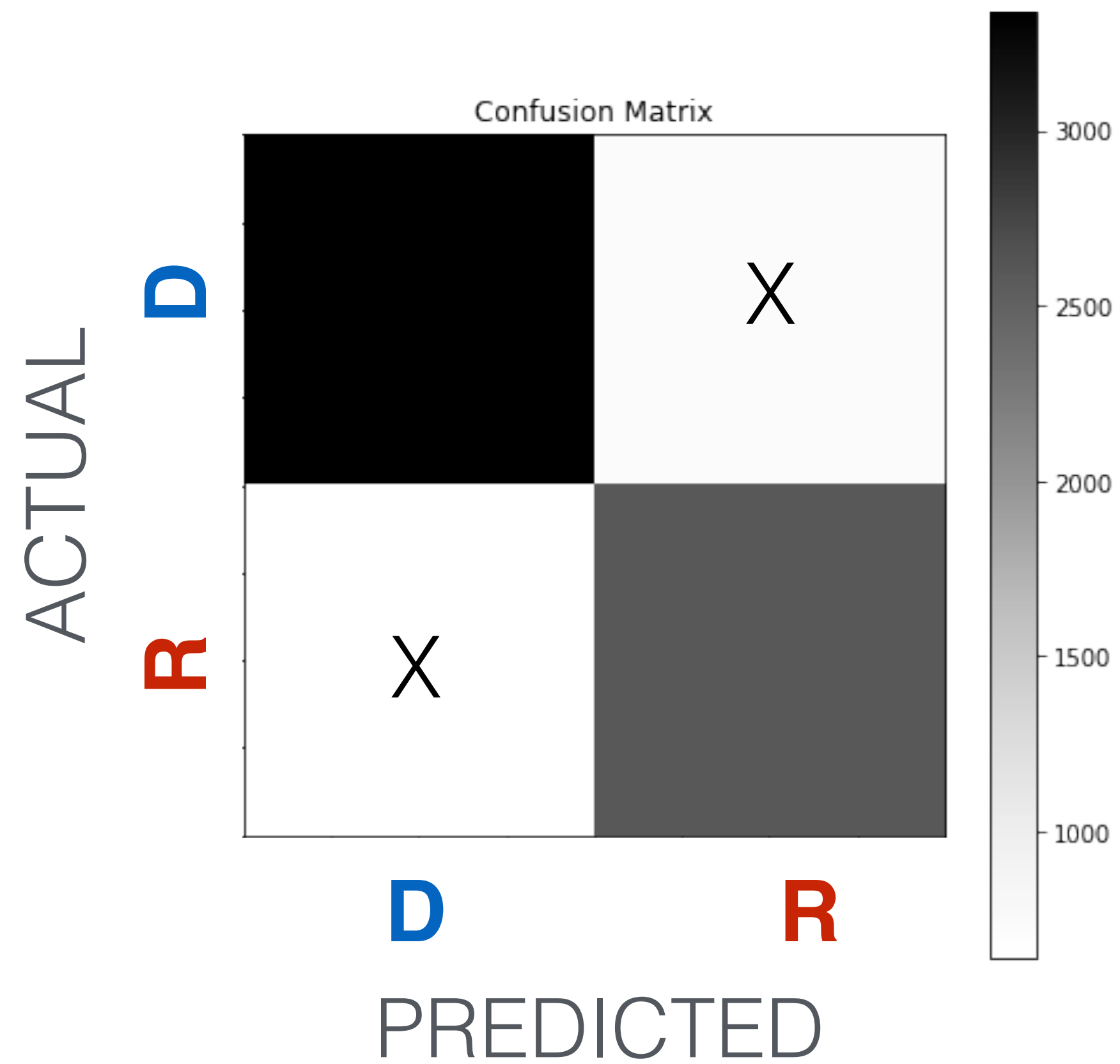


ROC AUC: **.855**



ENSEMBLED MODEL

ROC AUC: **.895**



FLASK APP



How Partisan is this tweet?

Submit the URL for a tweet above

Past Week's Partisan Tweets

Democrat and Republican congressional legislator tweets from the past 7 days

Click on a hashtag below to filter timeline and tweets view
or use the filter controls

Show Top X Hashtags

85



Search Hashtags by Word

APPENDIX

DATA GATHERING

SETUP

REQUIREMENTS.TXT
CONFIG.INI

INSTALL

- Package dependencies w/ versions
- Credentials, API keys and DB config

DB_FUNCTIONS.PY

DEFINE

- Twitter API Class
- Helper functions to:
 - Connect to DB
 - Transform JSON to pandas

EXECUTION

DB_APP.PY

CLI COMMANDS

- Initial Data Gathering
- Initial Data Load DB

EXECUTION

DB_APP.PY

CLI COMMANDS

- Initial Data Gathering →
- Initial Data Load DB

INPUT

YAML Files
Setup Files



Load In
Legislator
YAML File

Load In
Social Media
YAML File



Pickle as Dataframes



Isolate Valid Twitter Screen Names



Use Twitter's API to
fetch the user timeline
for legislators over
past 30 days



Create
User Profile
Log Dataframe

Create Tweet
Dataframe

Pickle Dataframe



OUTPUT

Legislators Pkl
Social Media Pkl
User Profiles Pkl
Tweet Pkl

EXECUTION

DB_APP.PY

CLI COMMANDS

- Initial Data Gathering
- Initial Data Load DB →

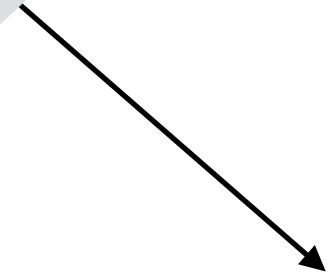
INPUT
Legislators Pkl
Social Media Pkl
User Profiles Pkl
Tweet Pkl



Load In 4
Pickled
Dataframes



Transform
Dataframes for
Schema



Define Postgres
Database Schema



Create DB Schema

