Twitter Bot Detection

Anantha Natarajan Selvaganapathy
N16989511
ans599@nyu.edu

The problem

What

The objective of this project is to use machine

learning techniques to detect whether a given Twitter account is a bot or not.

Why

16% of the US adults use Twitter for news

Bots can be used to aggressively retweet and share, favorite or reply to certain articles thus creating a ‘fake’ trend

How

Aim to come up with efficient ways to use machine learning algorithms to detect Twitter bots.

Challenges deep-dive

Challenge 1

Data Collection

How do you collect the labeled dataset?

Challenge 2

Data Cleaning

Is the data clean?

How do we clean the data?

Challenge 3

Dataset Size

Is the dataset large enough to train a good classifier?

Solution

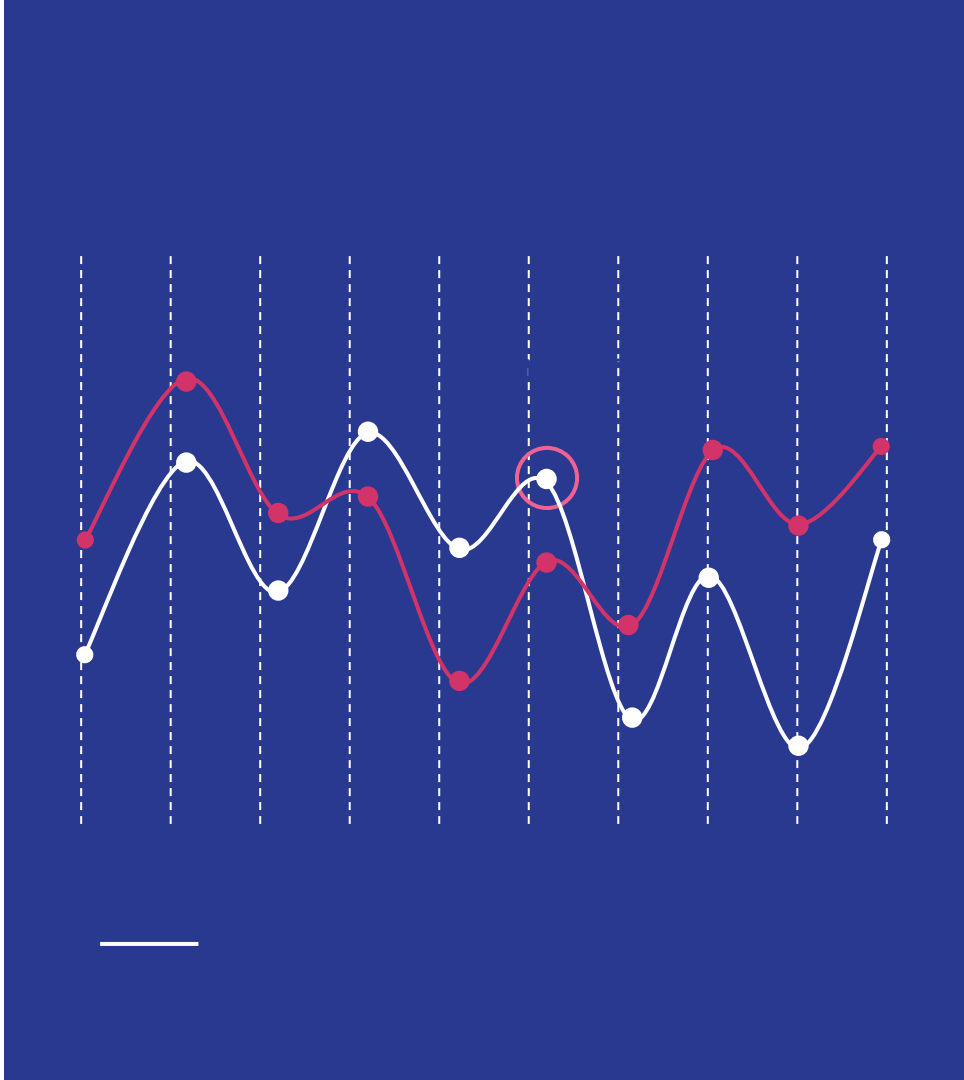
Use Random Forest Classifier to predict bot or not.

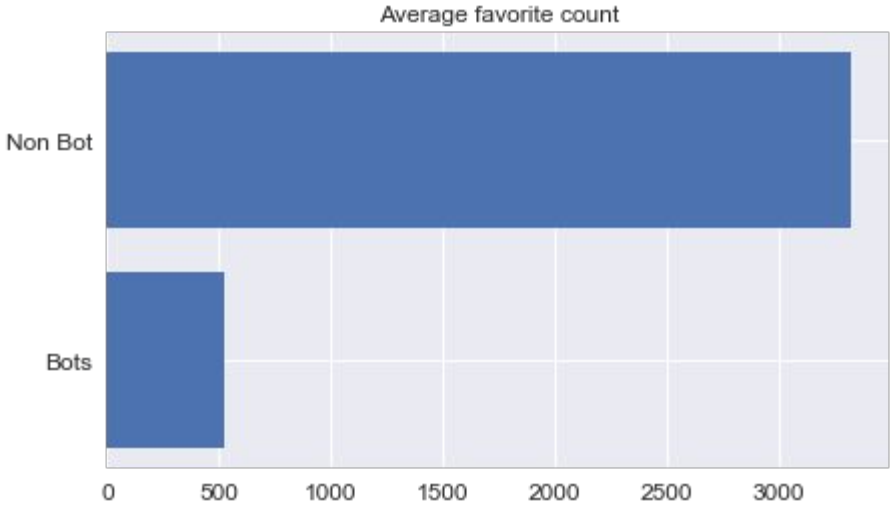
But is that enough?

Implementation

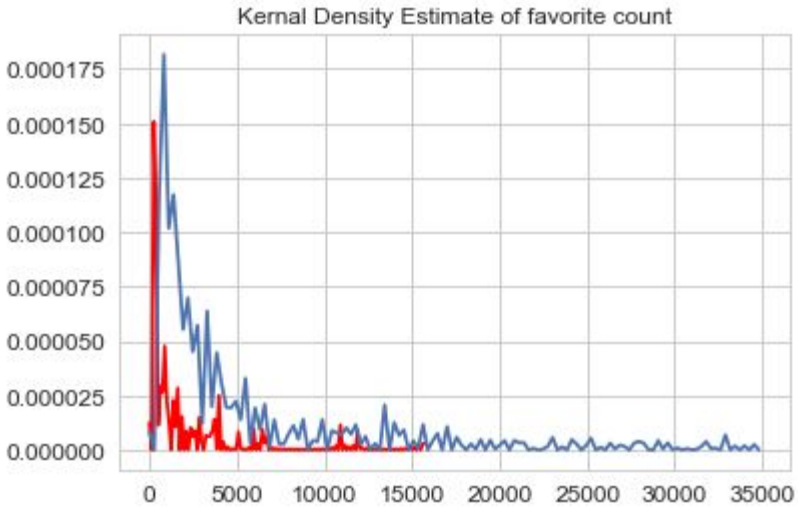
Feature Visualization

What is important?



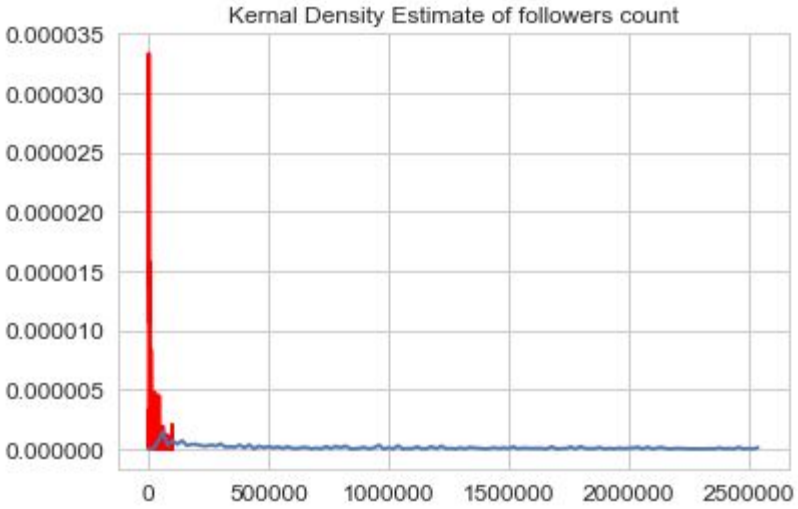


Avg. Favorite Count



Kernel Density Estimate analysis

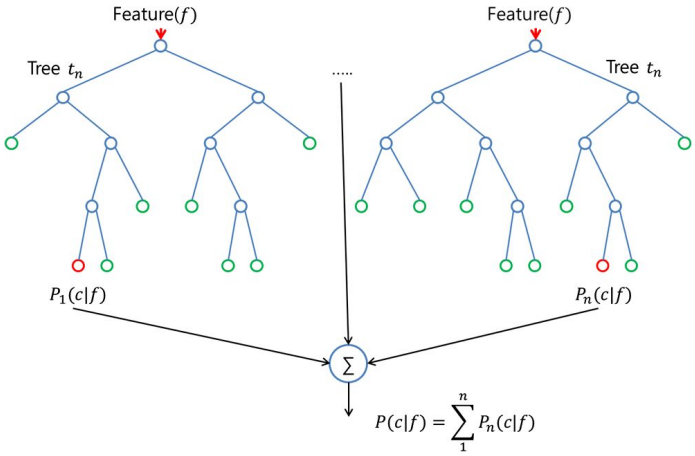
Of Favorite count



Kernel Density Estimate analysis

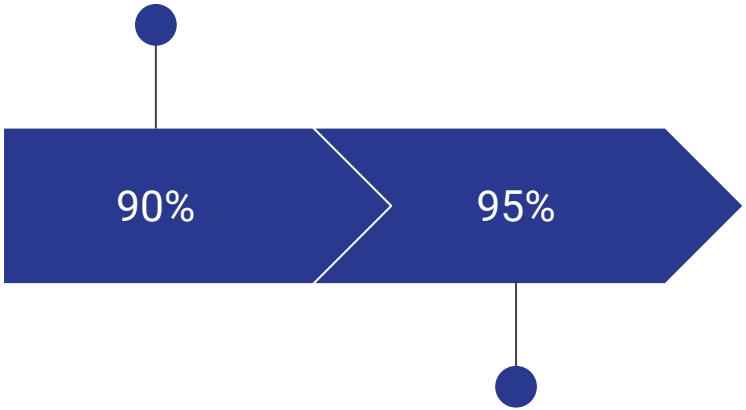
Of Followers count

Let’s train!

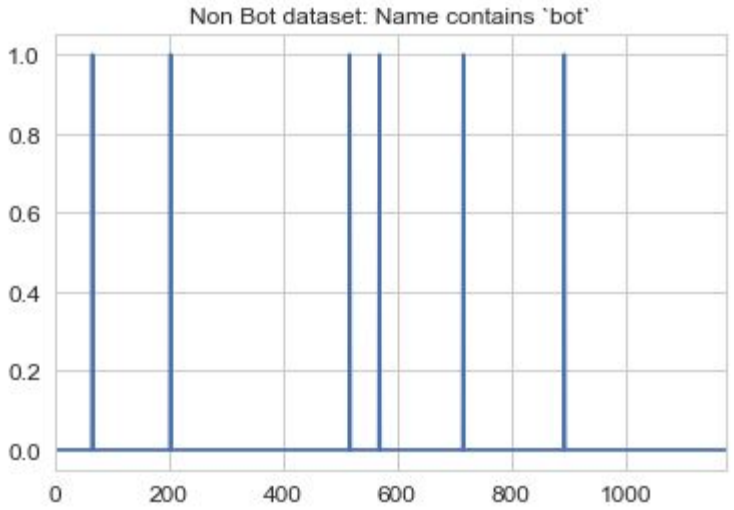


['screen_name', 'description', 'followers_count', 'friends_count', 'listedcount', 'favourites_count', 'verified', 'statuses_count', 'status', 'default_profile', 'url', 'location', 'name', 'id_len', 'default_profile_image']

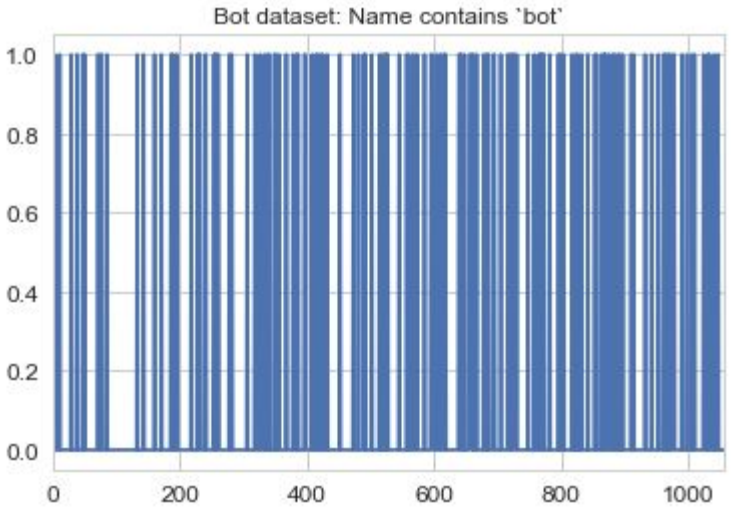
Random Forest with most of the numerical features



Create better features: eg. is ‘bot’ present in the name?



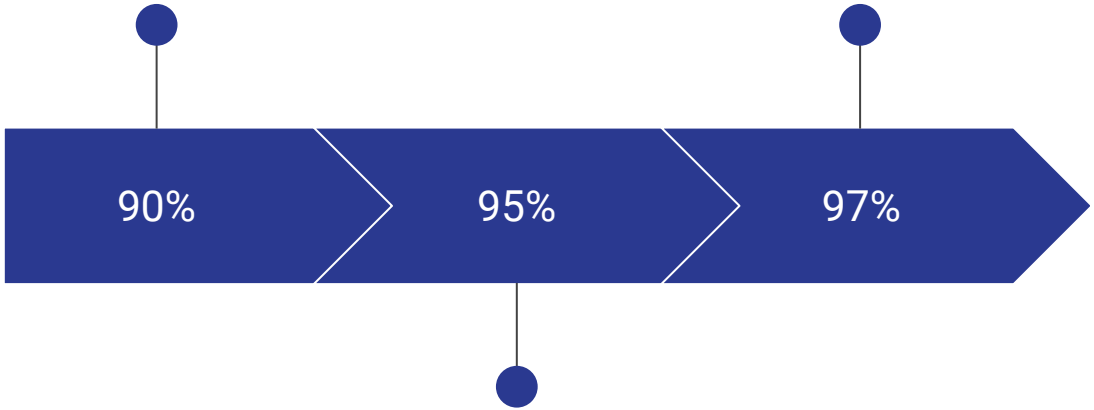
Is ‘bot’ present in name?



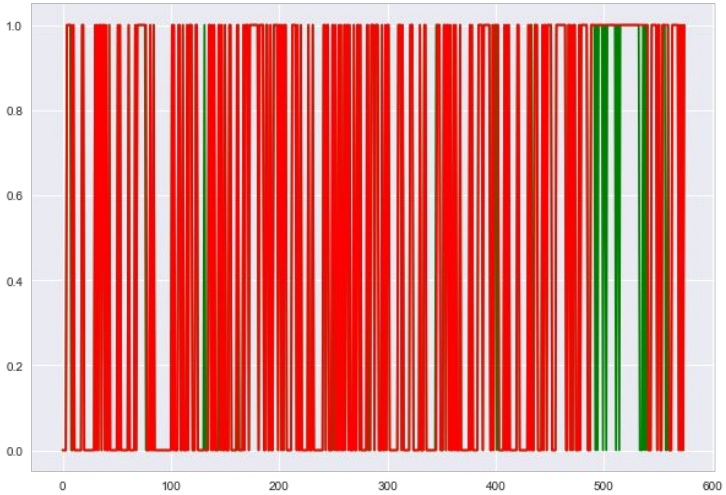
Is ‘bot’ present in name?

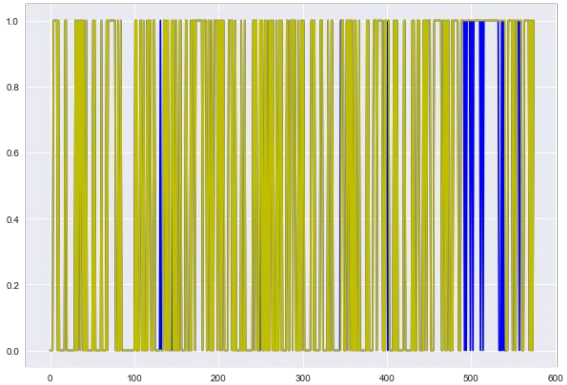
Random Forest with most of the numerical features

Further feature engineering



Create better features: eg. is ‘bot’ present in the name?



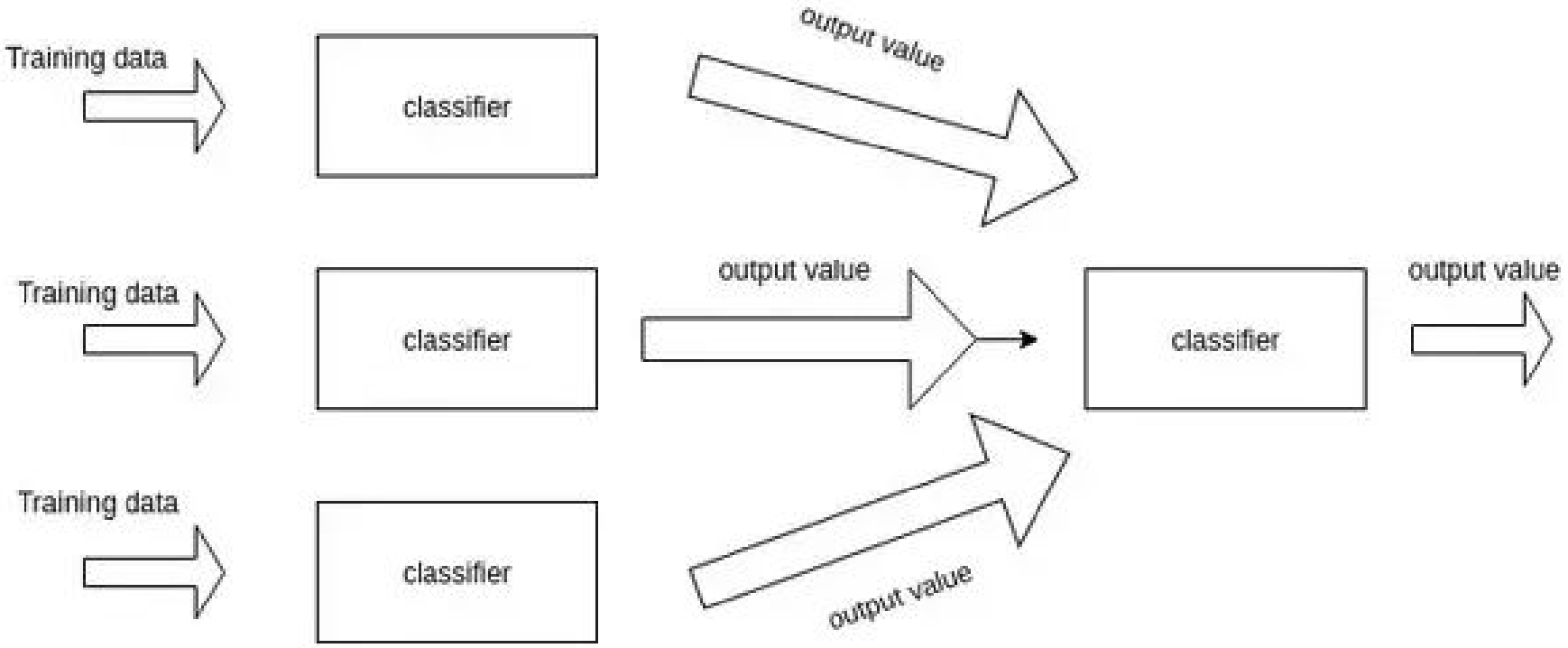


Predictions are *almost* the same across models - except a few outliers.

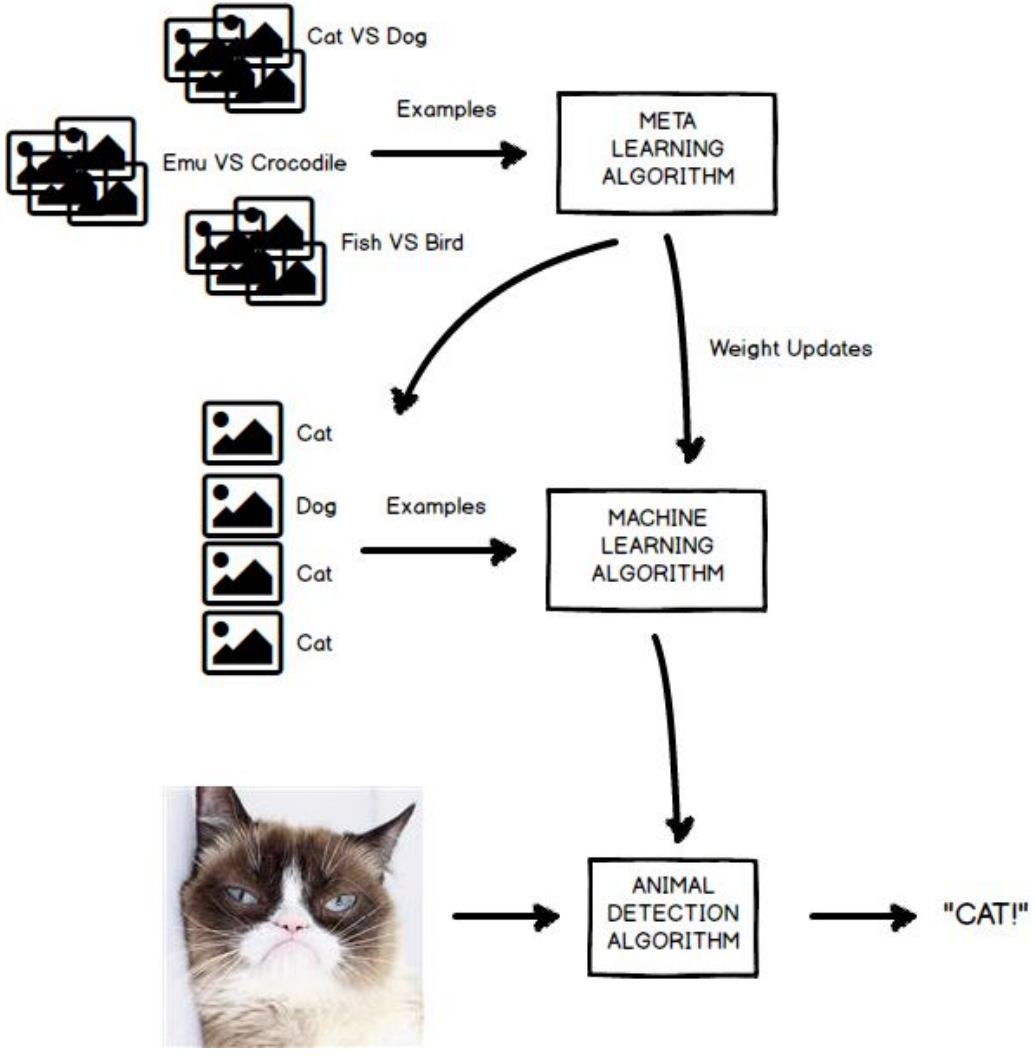
Some models have high True Positives for bots some have for non-bots.

Can we weight the predictions? How?

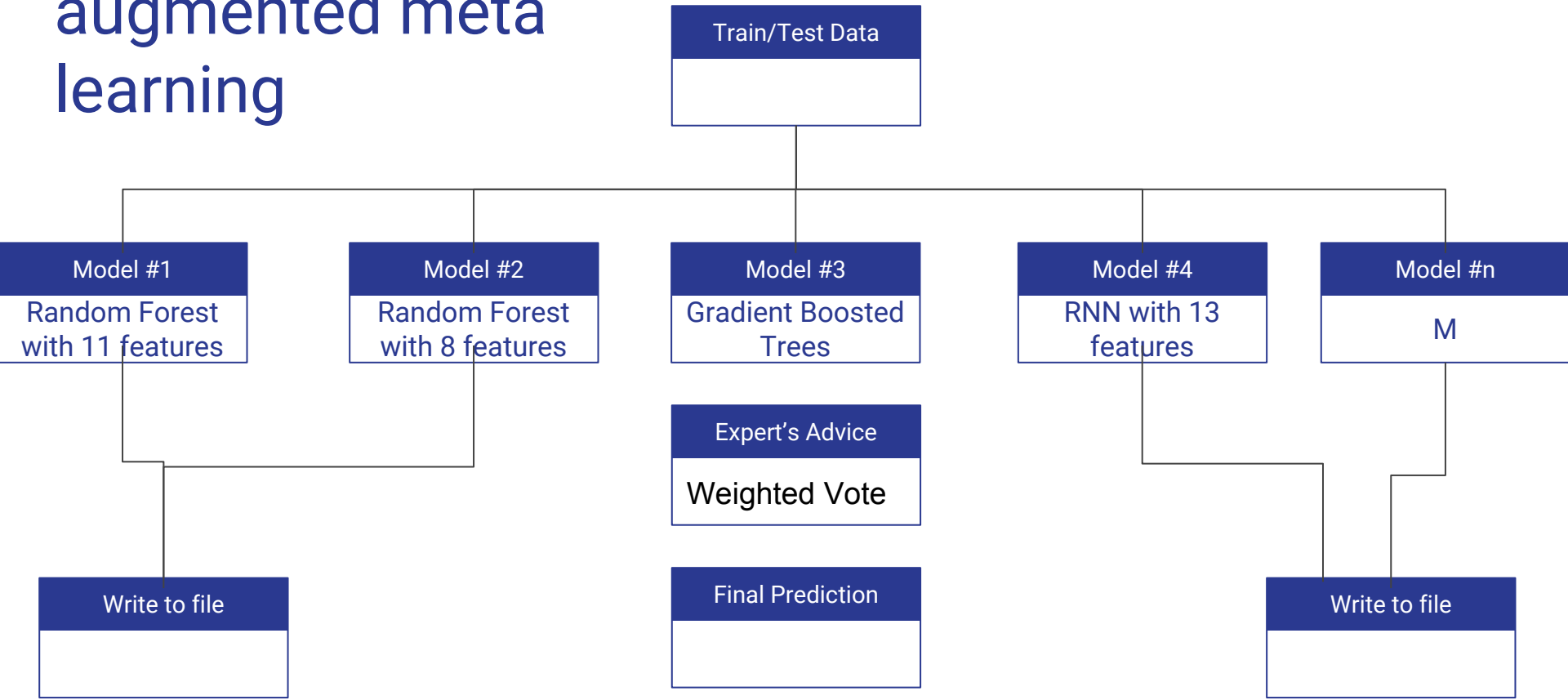
Stacking


Training data
classifier
output value
Training data
classifier
output value
classifier
output value
Training data
classifier
output value

Randomized weighted majority algorithm



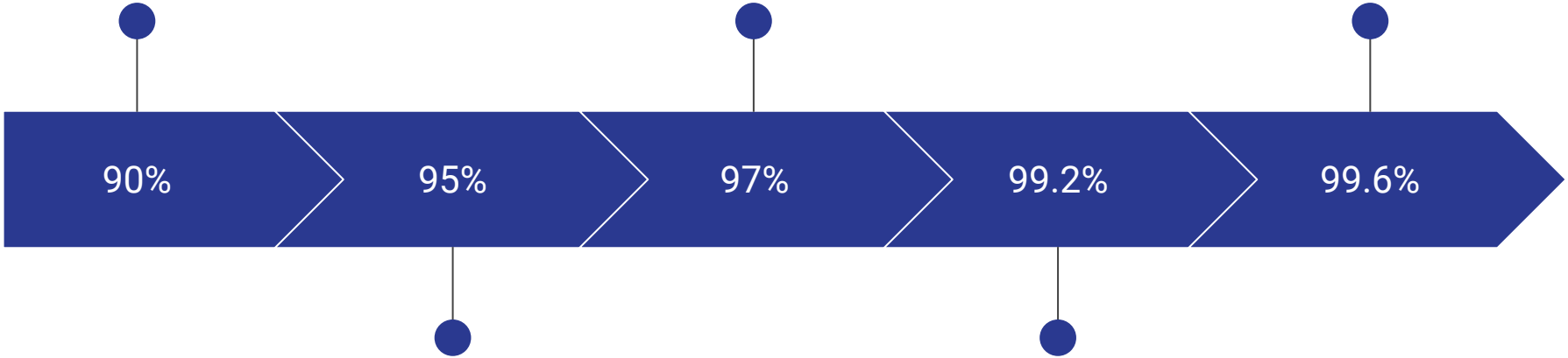
Memory augmented meta learning



Random Forest with most of the numerical features

Further feature engineering

Ensemble output pipelined into an ‘Expert’s Advice algorithm



Create better features: eg. is ‘bot’ present in the name?

Ensemble methods with aggressive parameter tuning

References:

http://sebastiaanboer.com/2016/05/28/Meta-Learning.html

https://en.wikipedia.org/wiki/Randomized_weighted_majority_algorithm

https://stats.stackexchange.com/questions/244842/how-to-combine-different-predictions-together

https://arxiv.org/pdf/1605.06065.pdf