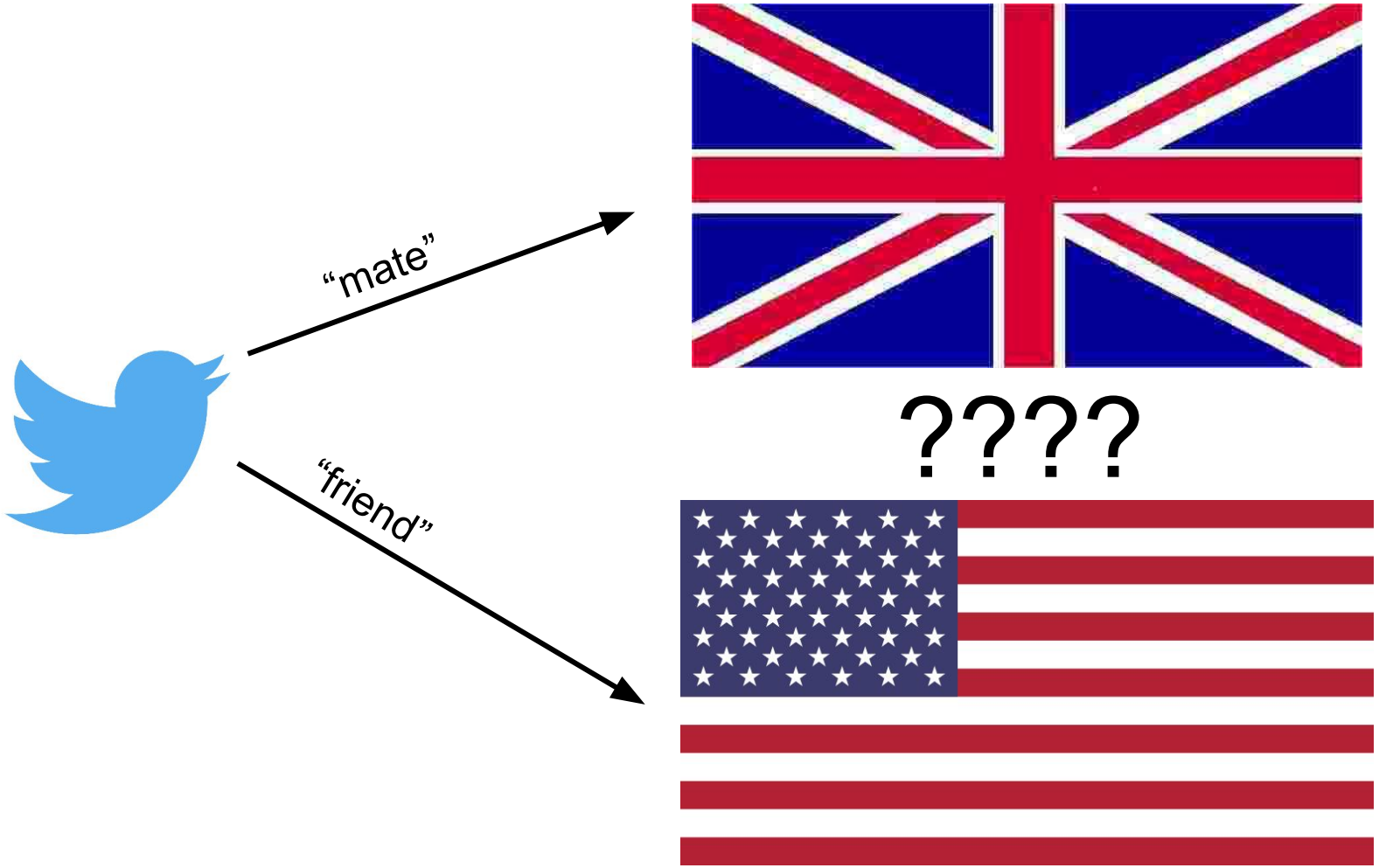
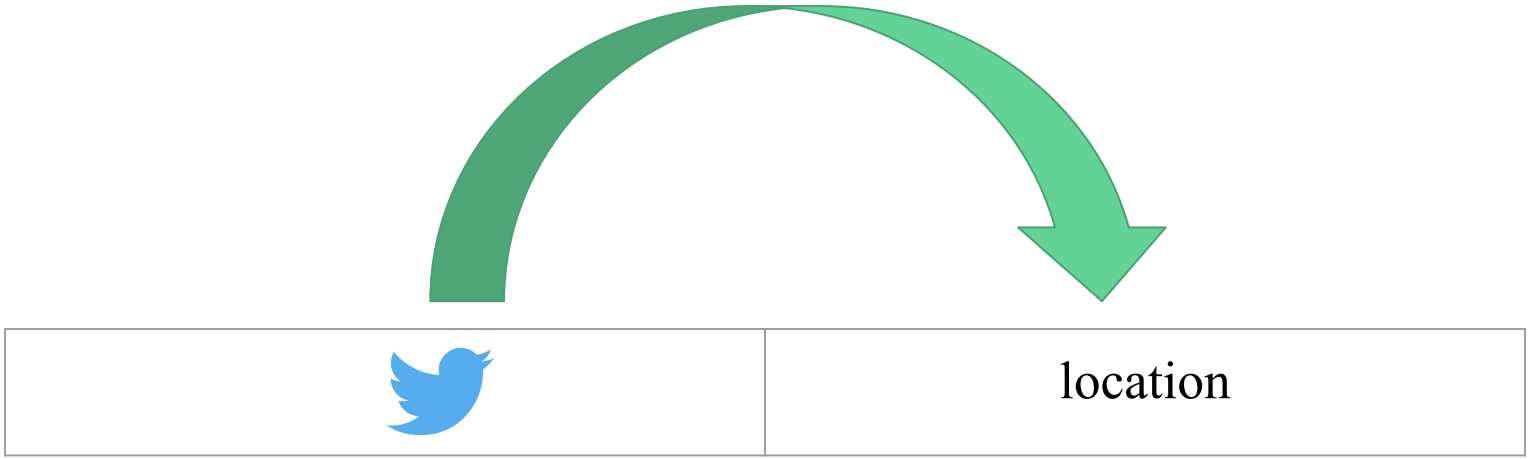
FriendOrMate?

Where does your customer live?

Julia Tsitron
Insight Data Science

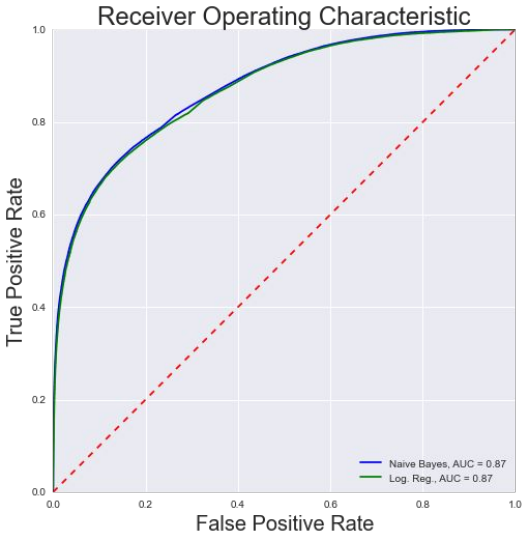

“mate”
“friend”
????



Advertisers don’t want to waste money targeting people who don’t have access to their product

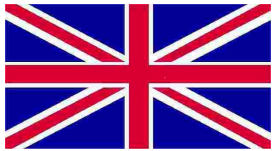
~ 1 million

Training::Cross Validation::Testing
Naive Bayes:
Prediction Accuracy: 79%
Logistic Regression:
Prediction Accuracy: 79%



80% / 20%
Training / Testing

or



What did the model learn?

friend

mom

realize

sports

color

math

roommate

soccer

mate

mum

realise

sport

colour

maths

rubbish

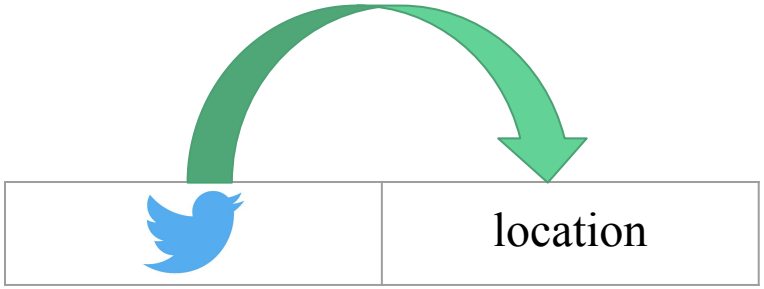
football

1

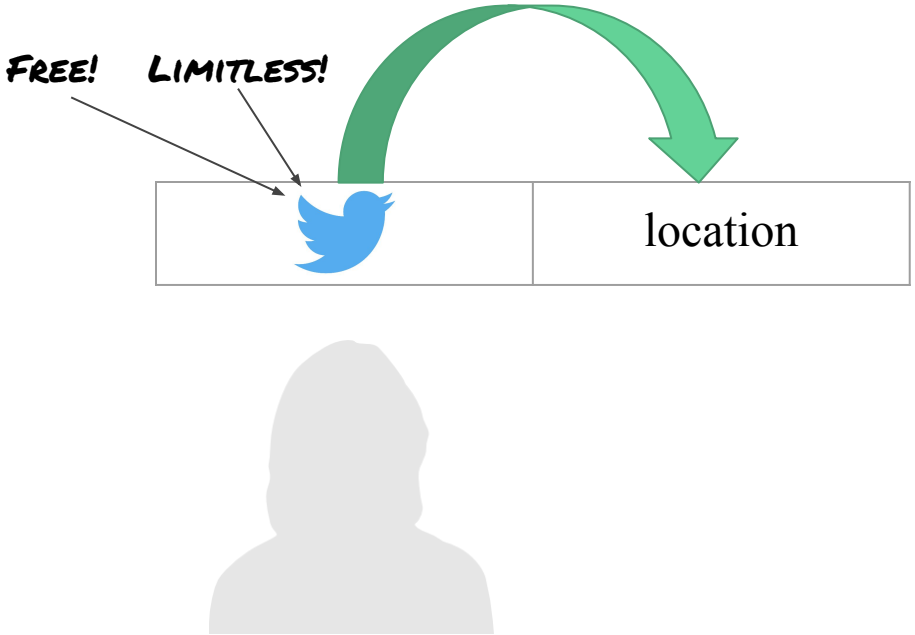
0

1

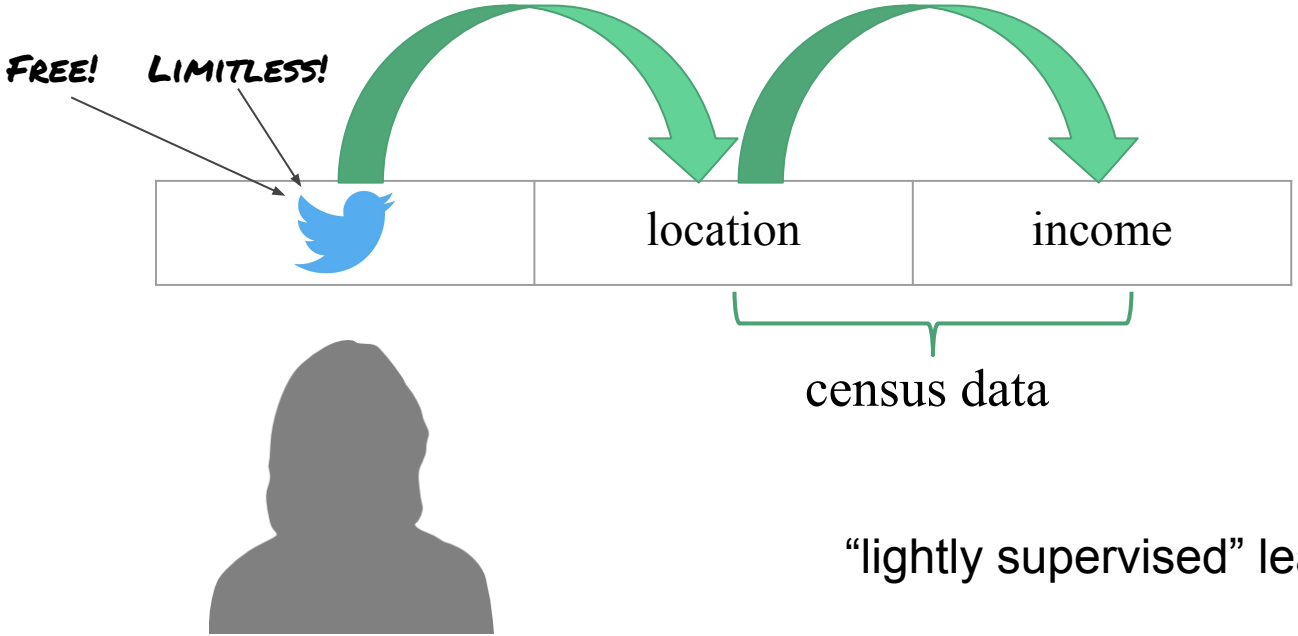
The Takeaway: We can learn user attributes from tweets!



The Takeaway: We can learn user attributes from tweets!

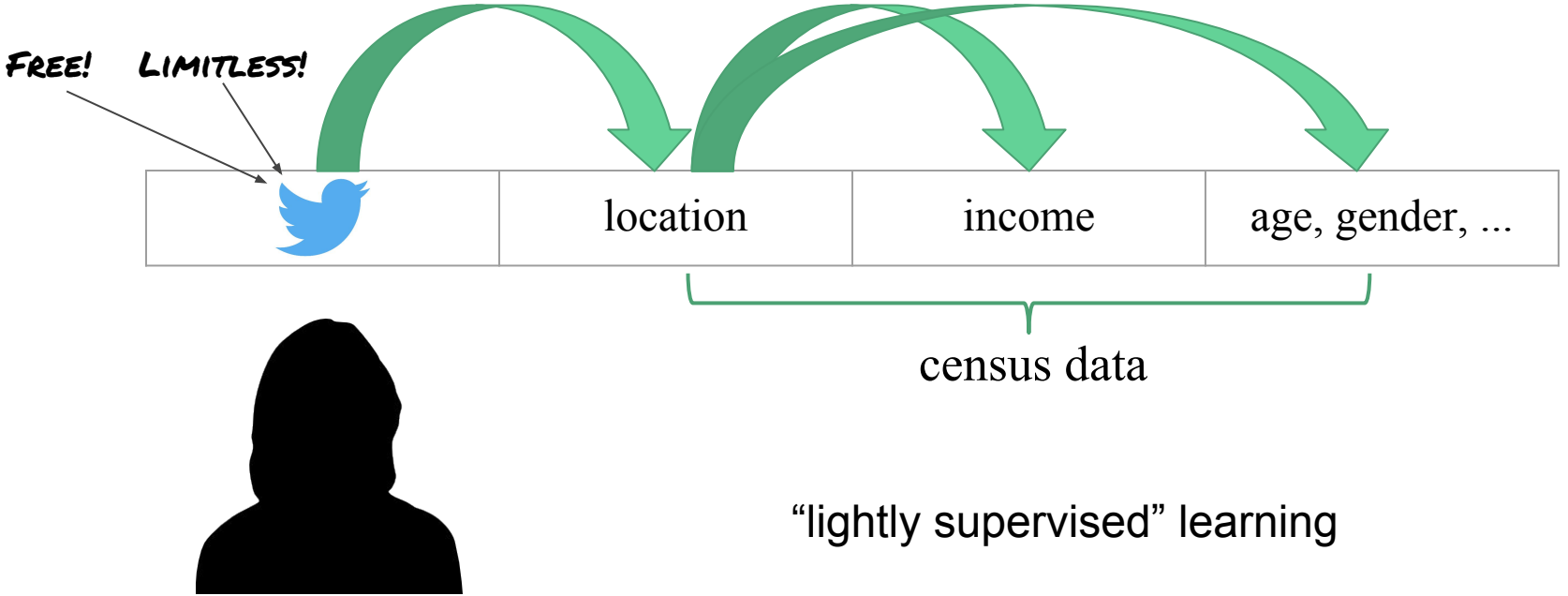


The Takeaway: We can learn user attributes from tweets!



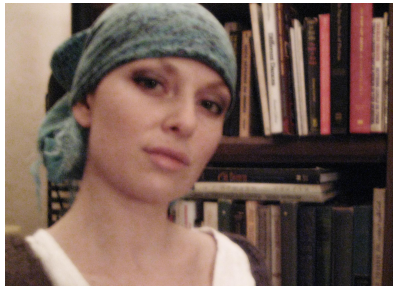
“lightly supervised” learning

The Takeaway: We can learn user attributes from tweets!

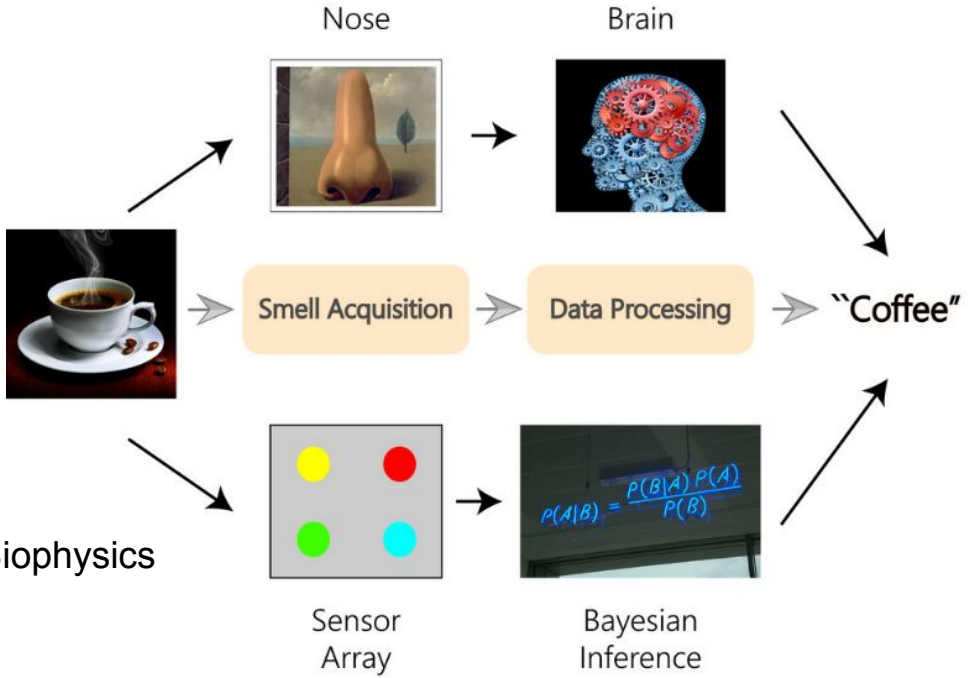


"lightly supervised" learning

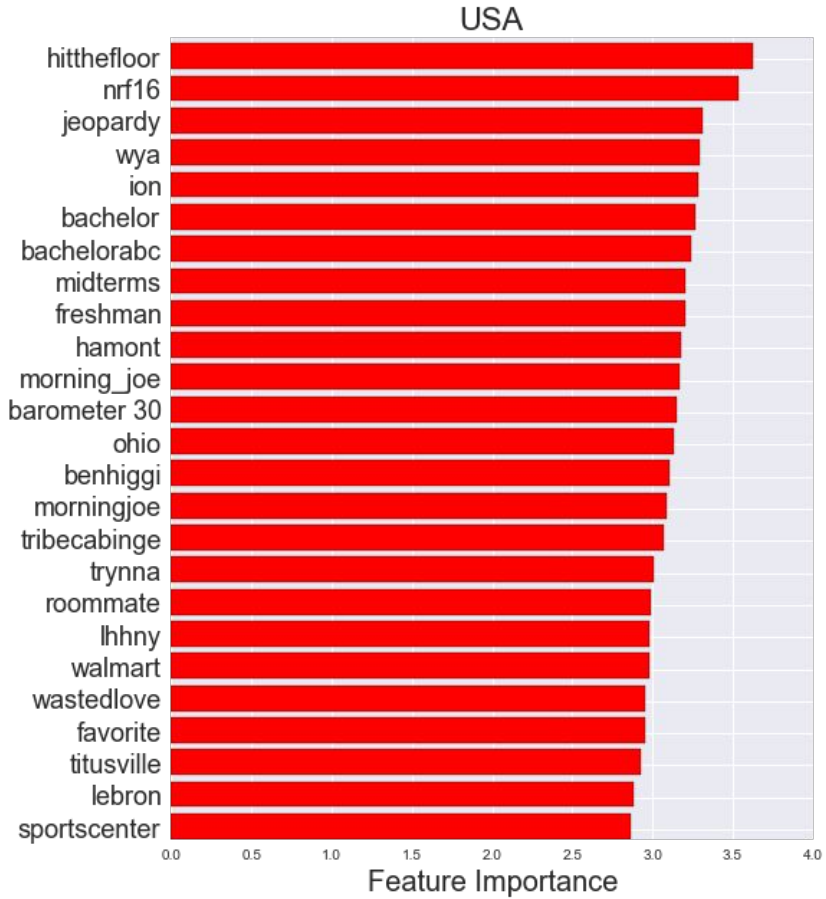
Julia Tsitron

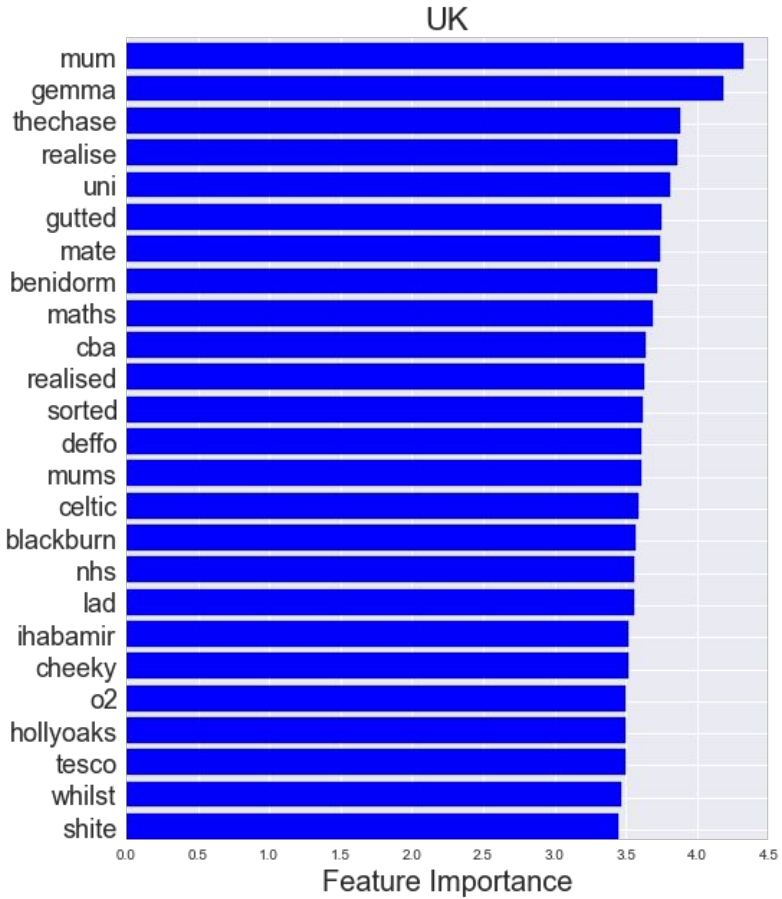


Ph.D., Computational Biology and Molecular Biophysics
M.A., Physics
B.A., Physics



Feature Importance

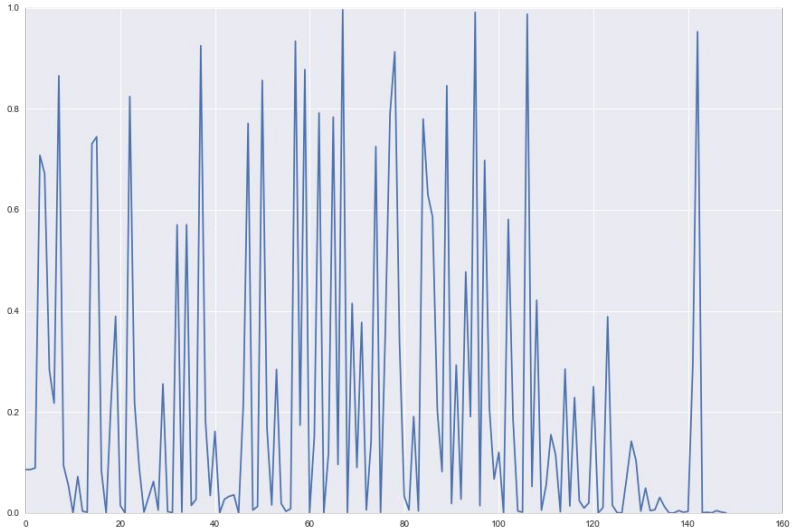




A Few Case Studies

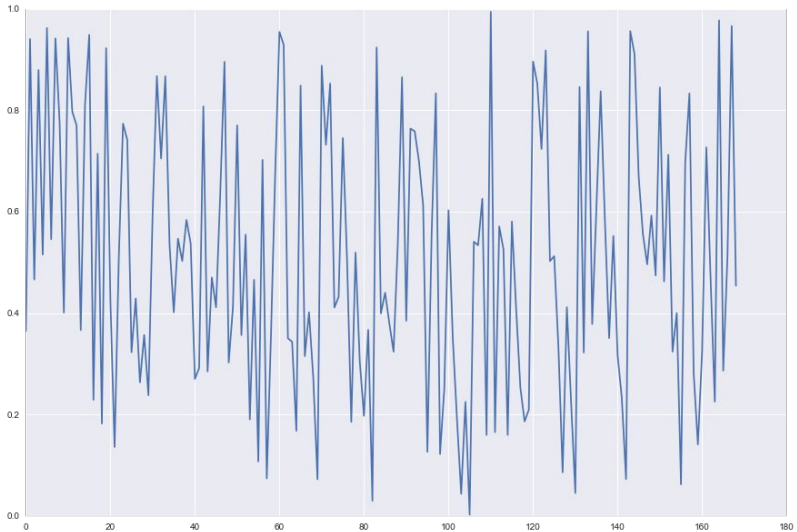
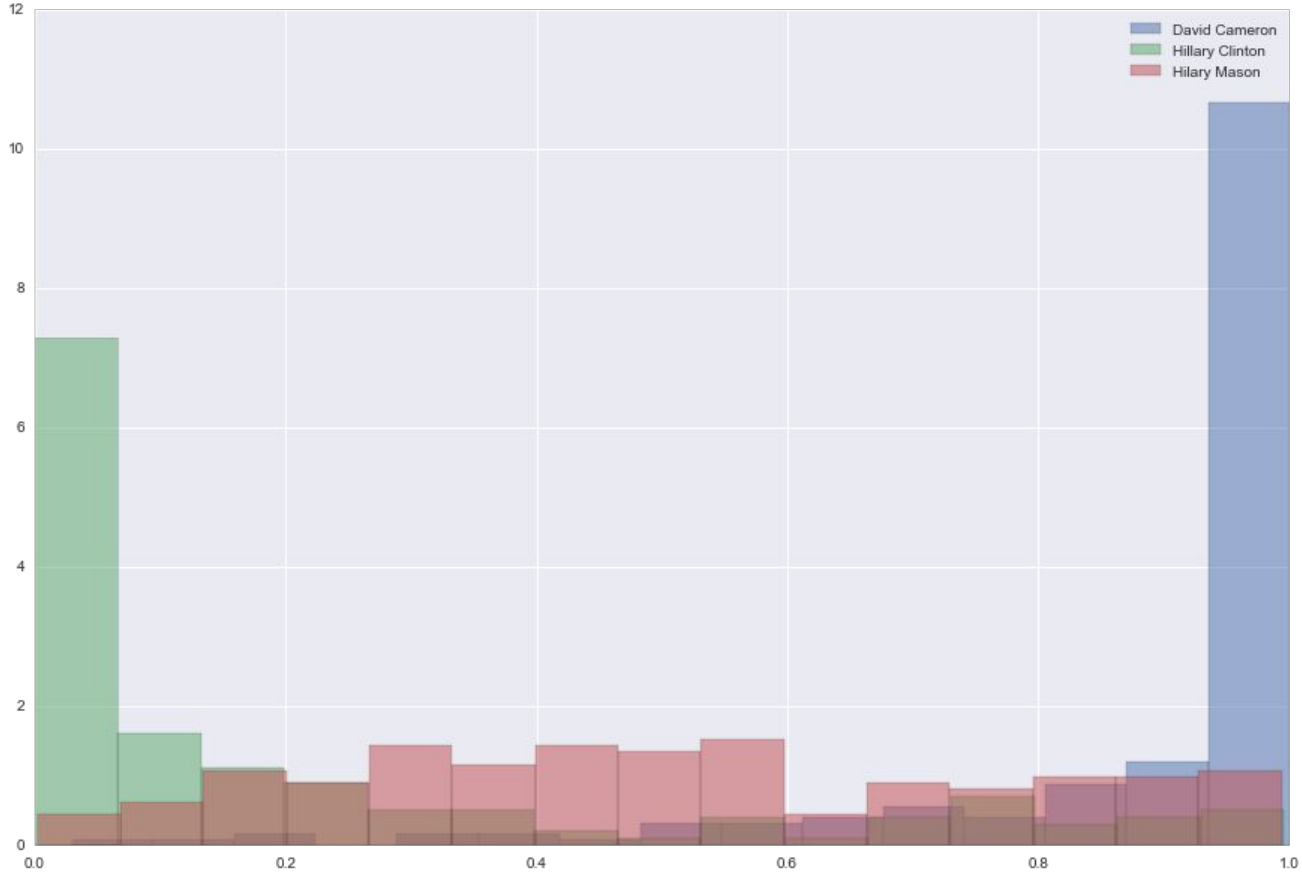
David Cameron

Hillary Clinton



A Few Case Studies

Hilary Mason



A Few Case Studies, continued