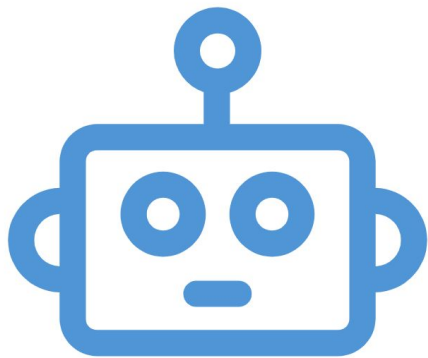


# Twitter Bot or Not?

---



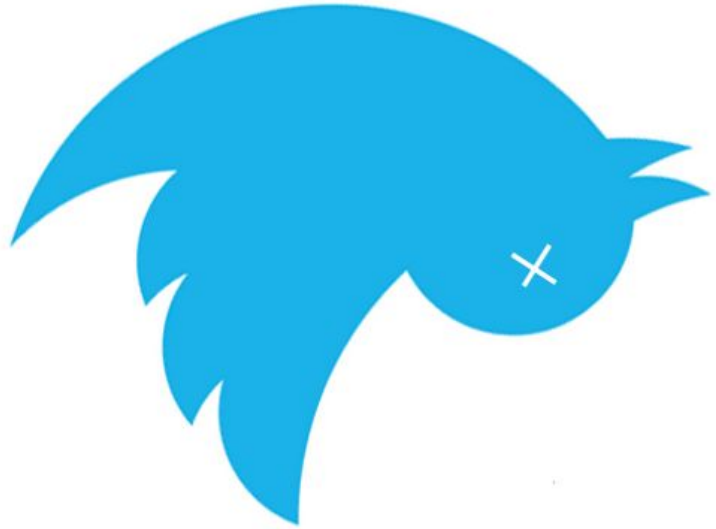
# Why do we care about bots?

---

A **Twitter bot** may autonomously **tweet, retweet, like, follow, unfollow,** or **direct message** other accounts

Which can mean...

- Spreading **misinformation**
- **Scams** and **exploitation**
- **Degrading** social **trust**
- Social media **platform deterioration**
- **Increased polarization**



# Twitter account-level features

Profile images

Network size

Date & time created

User activity

The image shows a Twitter profile page for a user named kris doenlen (@scrapfishies). The profile picture is a circular image of a man with a mustache and sunglasses. The header image is a photo of a man's face inside a squirrel-shaped cutout. The profile name is kris doenlen, and the handle is @scrapfishies. The location is Oakland, CA, and the website is github.com/scrapfishies. The user joined in July 2020. The profile shows 65 Following and 2 Followers. The user activity section is highlighted with a dashed orange box and contains links to Tweets, Tweets & replies, Media, and Likes. Annotations with orange dashed arrows point to various features: 'Profile images' points to the profile picture and header image; 'Network size' points to the 'Following' and 'Followers' counts; 'Date & time created' points to the 'Joined July 2020' text; and 'User activity' points to the activity tabs.

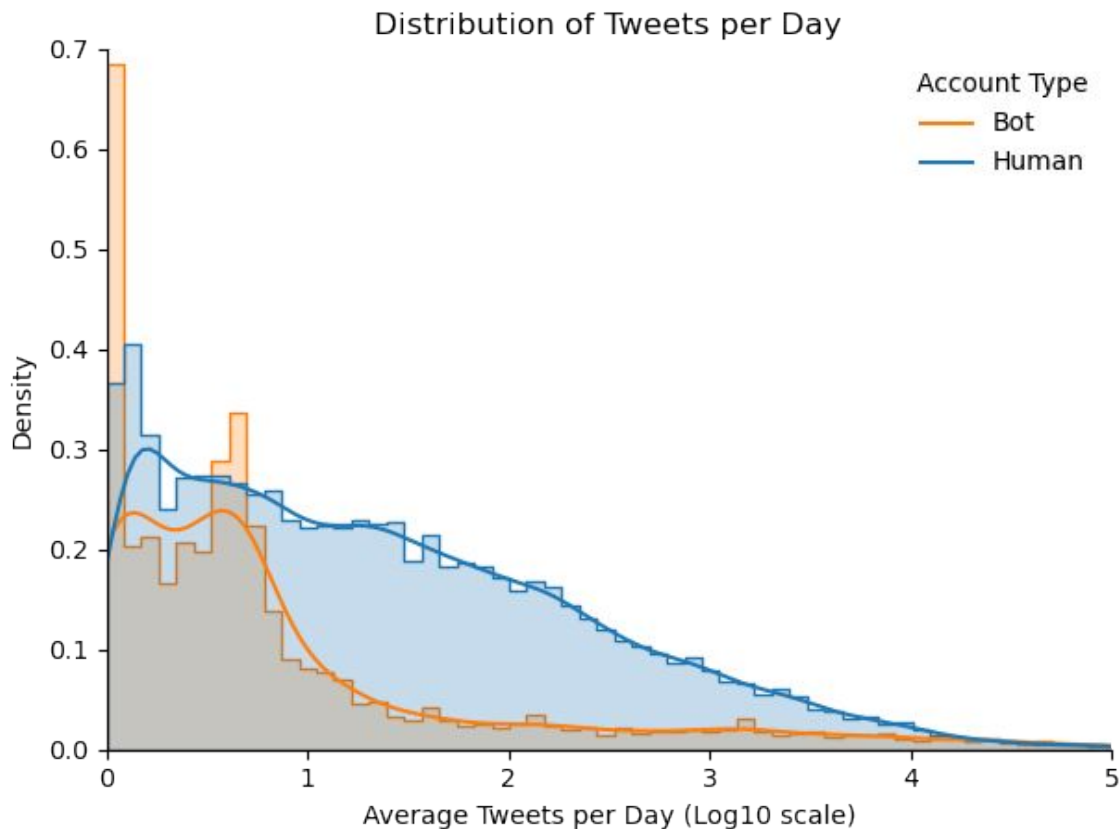
**kris doenlen**  
@scrapfishies

📍 Oakland, CA 🔗 [github.com/scrapfishies](https://github.com/scrapfishies) 📅 Joined July 2020

65 Following 2 Followers

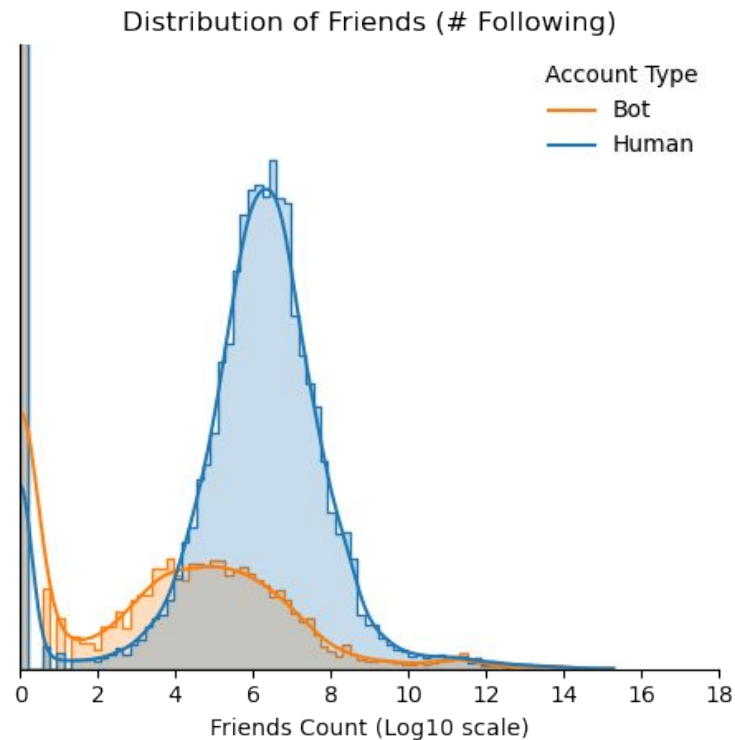
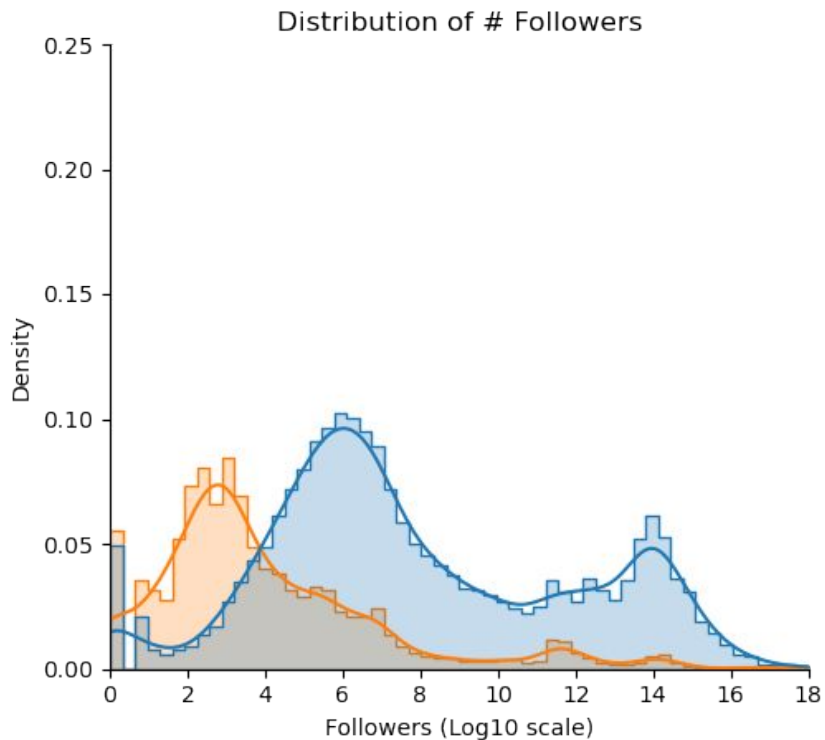
[Tweets](#) [Tweets & replies](#) [Media](#) [Likes](#)

# Bots behave differently...



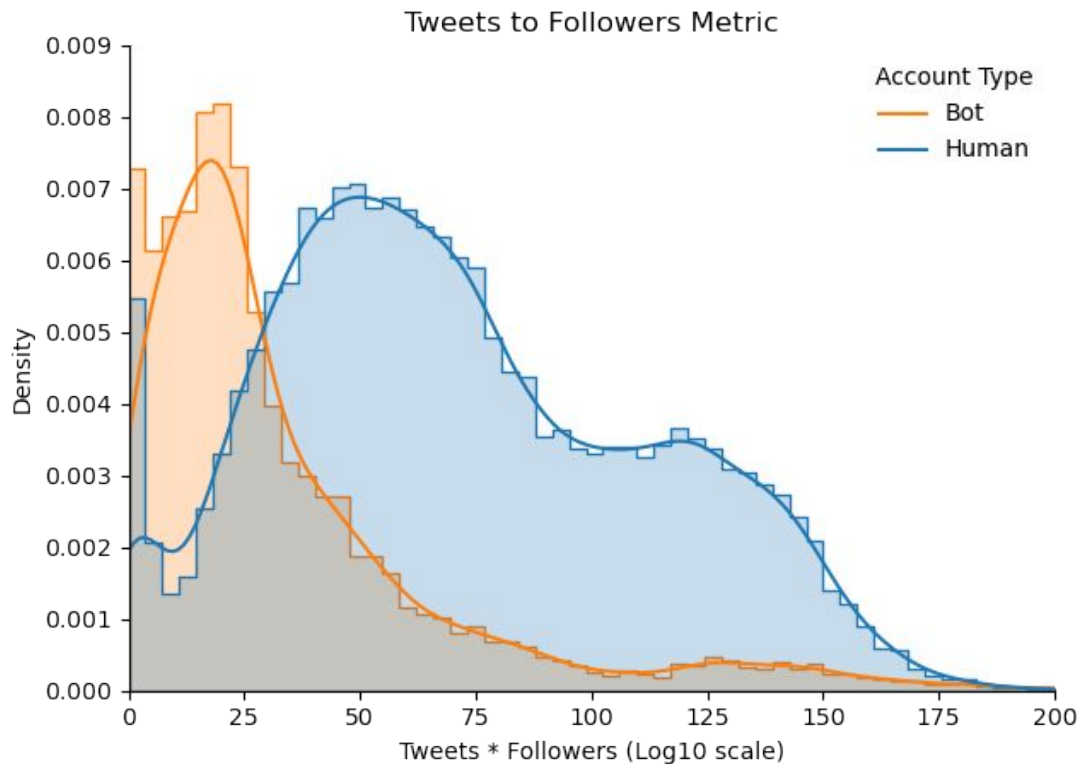
# ...and their networks look different, too

Distribution of Followers & Friends



# Model construction and selection

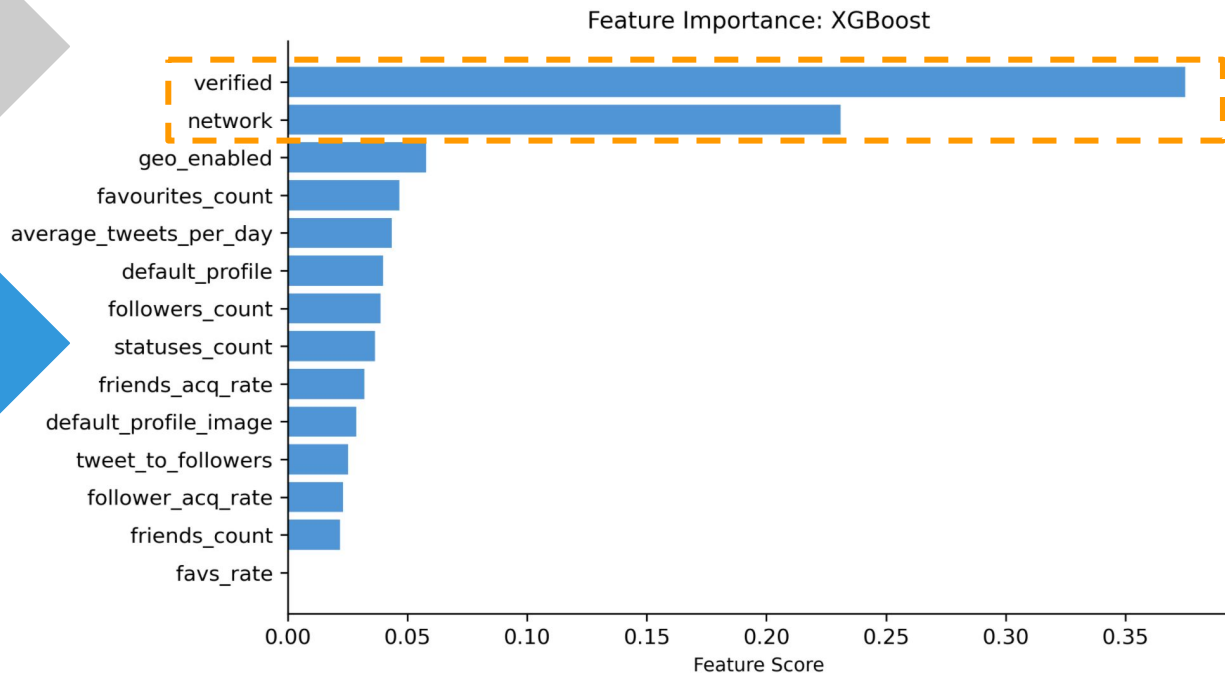
## Feature Engineering



# Model construction and selection

Feature Engineering

Feature Selection

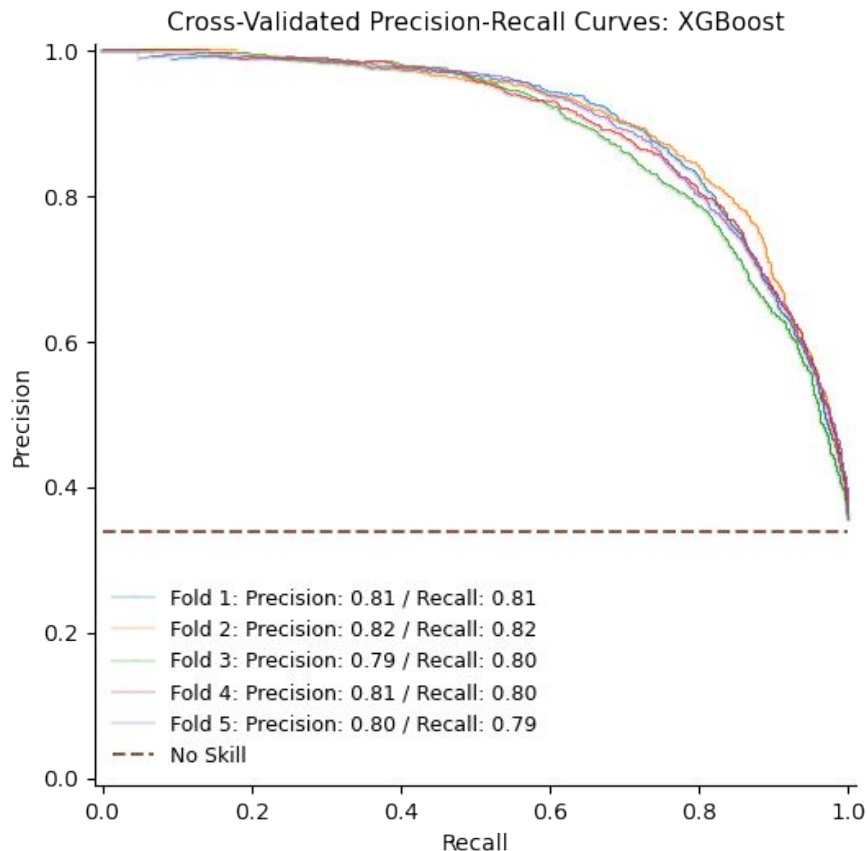


# Model construction and selection

Feature Engineering

Feature Selection

Parameter Tuning





# XGBoost model results

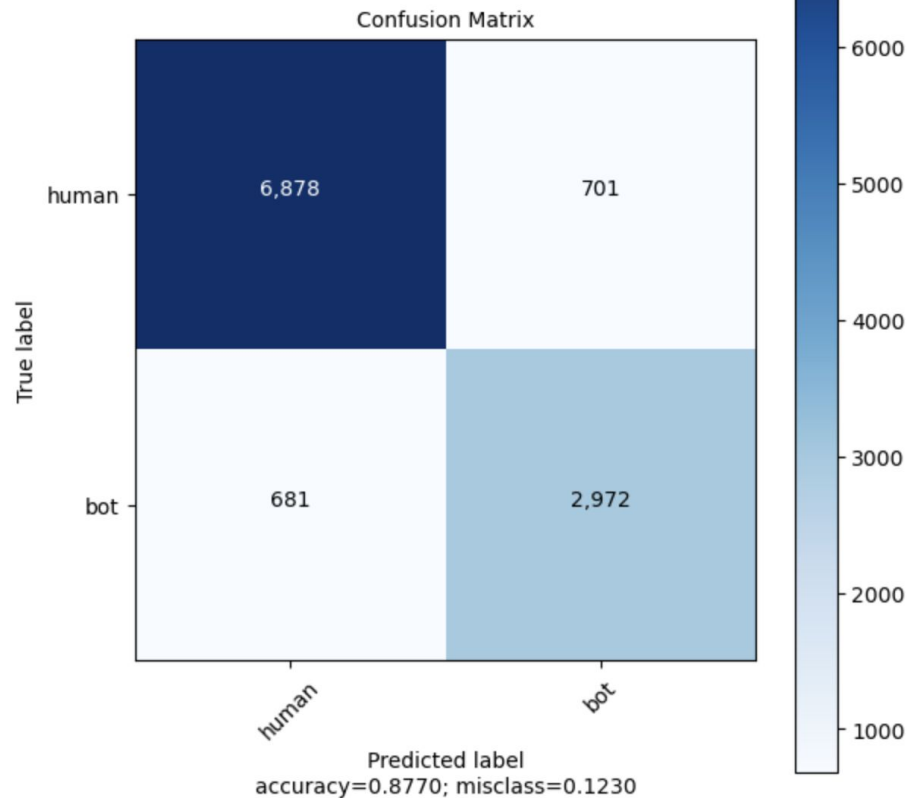
| Metric | Test Score (%) |
|--------|----------------|
|--------|----------------|

|          |      |
|----------|------|
| Accuracy | 87.7 |
|----------|------|

|           |      |
|-----------|------|
| Precision | 81.3 |
|-----------|------|

|        |      |
|--------|------|
| Recall | 80.9 |
|--------|------|

|         |      |
|---------|------|
| ROC AUC | 93.4 |
|---------|------|





## Twitter Bot or Not

Enter a Twitter handle



scrapfishies

nytimes

willdoenlen

atotalpirate

aoc

sdf

Twitter Bot or Not by Chris Doenlen |    

# Final thoughts

---

## Not all bots are bad!

- Sometimes they're **creative, helpful**, or **funny**
- It all depends on **how they were *designed***
- **Bot awareness is key**

## Suggestions for future work

- Tweet-level features
- **Build a bot to detect bots**
- Classify *types* of bots (political, fake followers, trolls, financial, spammers....)

# Thank you!

---

Chris Doenlen

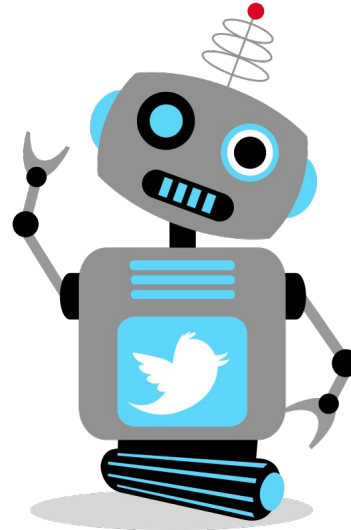
[Twitter Bot or Not](#)

[GitHub Repository](#)

[scrapfishies@gmail.com](mailto:scrapfishies@gmail.com)

[@scrapfishies](#)

(Probability of being a bot: 23.9%)



# Appendix

---

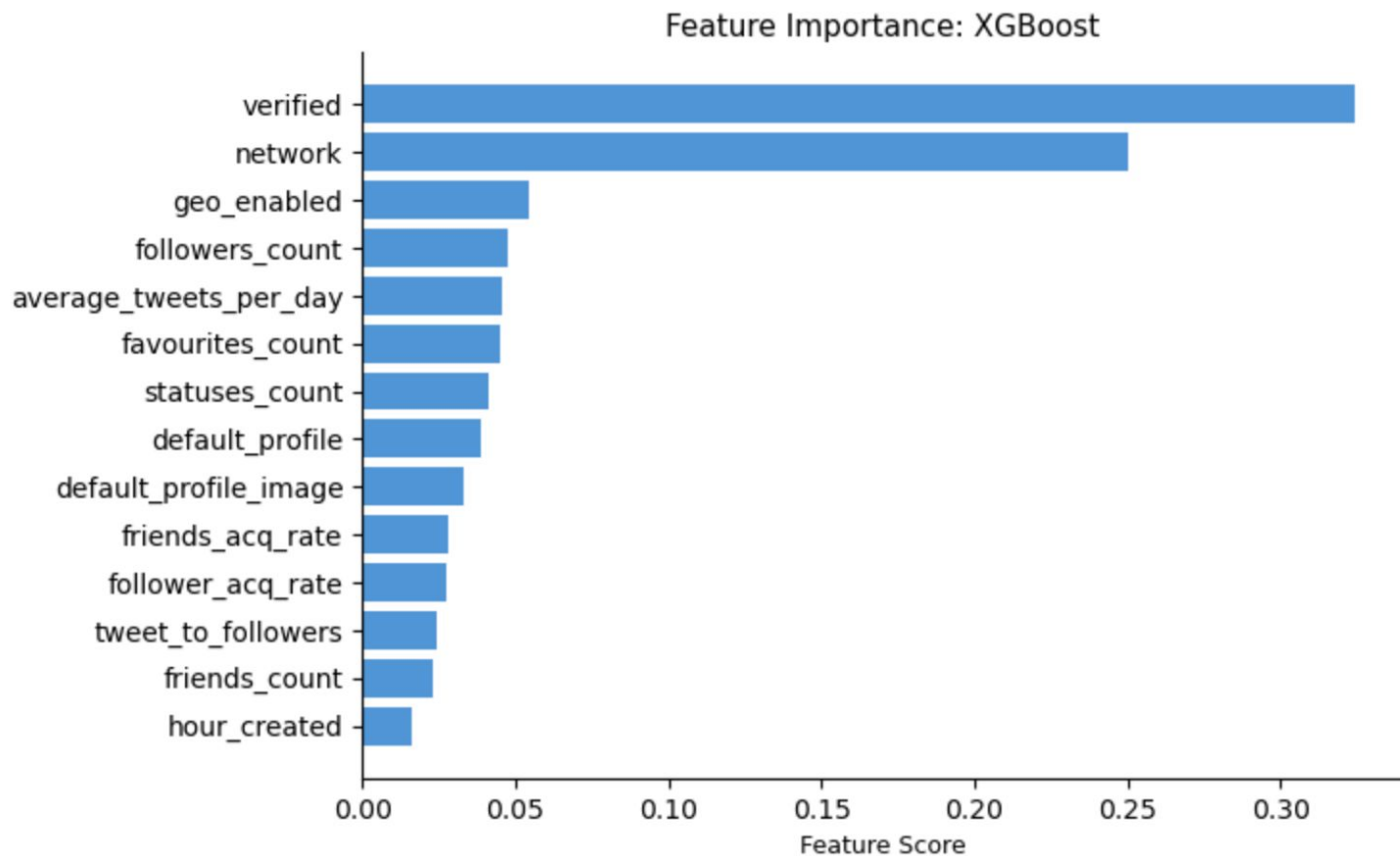
Try out the live app: [Twitter Bot or Not](#)

## Sources & References

- [Kaggle: Twitter Bot Accounts](#)
- [Bot Repository](#)
- [Tweepy](#)
- [Twitter](#)

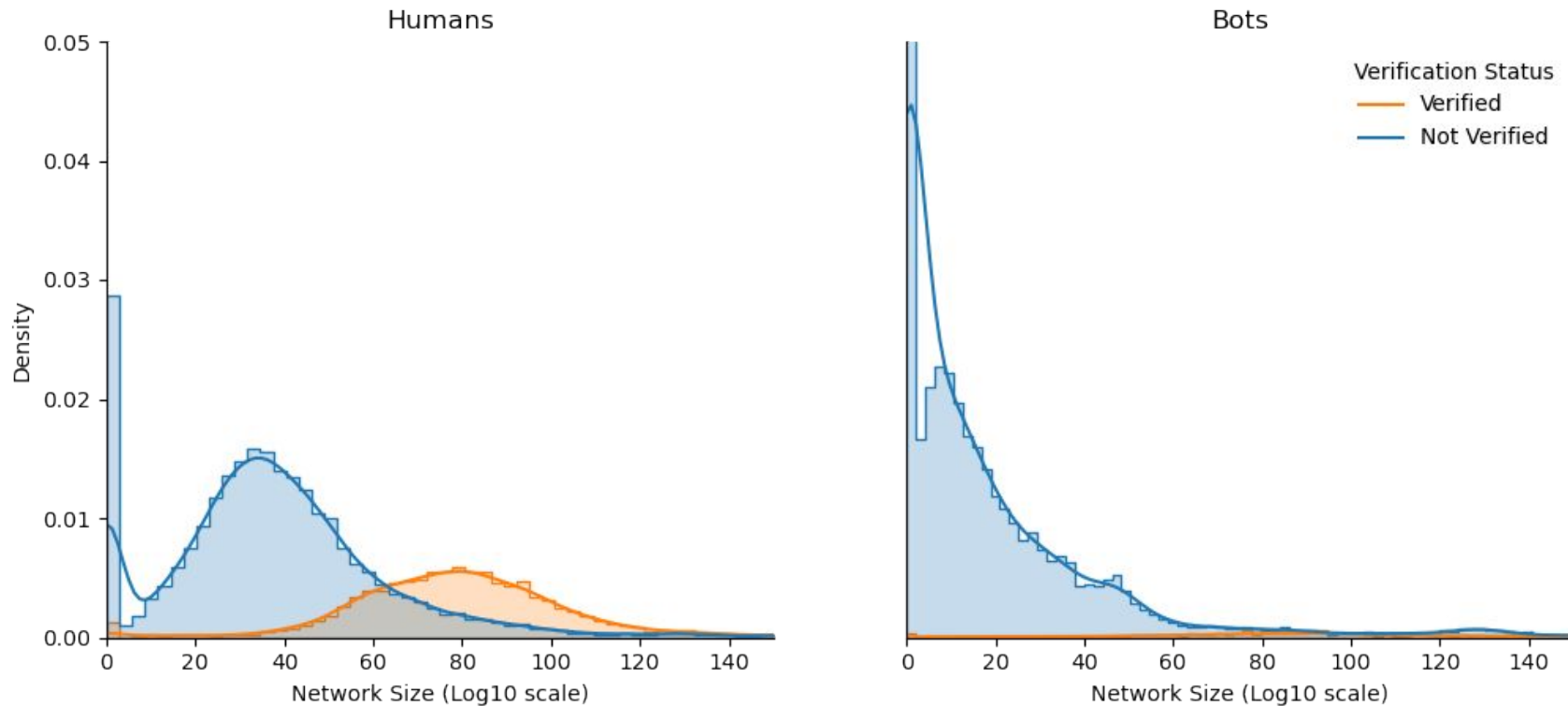
## Appendix: Final XGBoost model features

---



# Appendix: Top model features

Network Size by Verification Status



# Appendix: Final XGBoost model performance curves

