Conclusion

▸ The gradient boosting classifier was able to perform with 90% accuracy on the testing data set, providing evidence that bot comments are distinguishable from non-bots.

▸ My non-bot user group is rather small.

- While I attempted to use a random selection process, I may have introduced bias based on my own opinion of what 'non-bot' user comments look like.

- Because several of my non-bot users were well-known public figures, I believe the scores for non-bots are skewed higher than they would be for the entire population of Reddit users.

Conclusion

‣ The gradient boosting classifier was able to perform with 90% accuracy on the testing data set, providing evidence that bot comments are distinguishable from non-bots.

CAVEATS

‣ My non-bot user group is rather small.

‣ While I attempted to use a random selection process, I may have introduced bias based on my own opinion of what 'non-bot' user comments look like.

‣ Because several of my non-bot users were well-known public figures, I believe the scores for non-bots are skewed higher than they would be for the entire population of Reddit users.

Conclusion