

Predicting Instagram Likes Based on Picture Features

Kalob Reinholz

- Introduction:

The aim for this project was to be able to predict the amount of likes a person would receive on Instagram based on the features included in the picture.

The goal of this project was to help Instagram users know what are the best features to have in a photo to increase the number of interactions they have on Instagram.

- Data:

The data set was created by randomly choosing public Instagram accounts that had a certain amount of followers. Once a person was selected, I chose 3 random pictures from their account and described the features.

There were 19 features in total that were used to describe each picture.

- Methods:

I used 3 different machine learning algorithms to try and predict Instagram likes. Random Forest, K-Nearest Neighbor, and Decision Trees.

- Results:

- Random Forest:

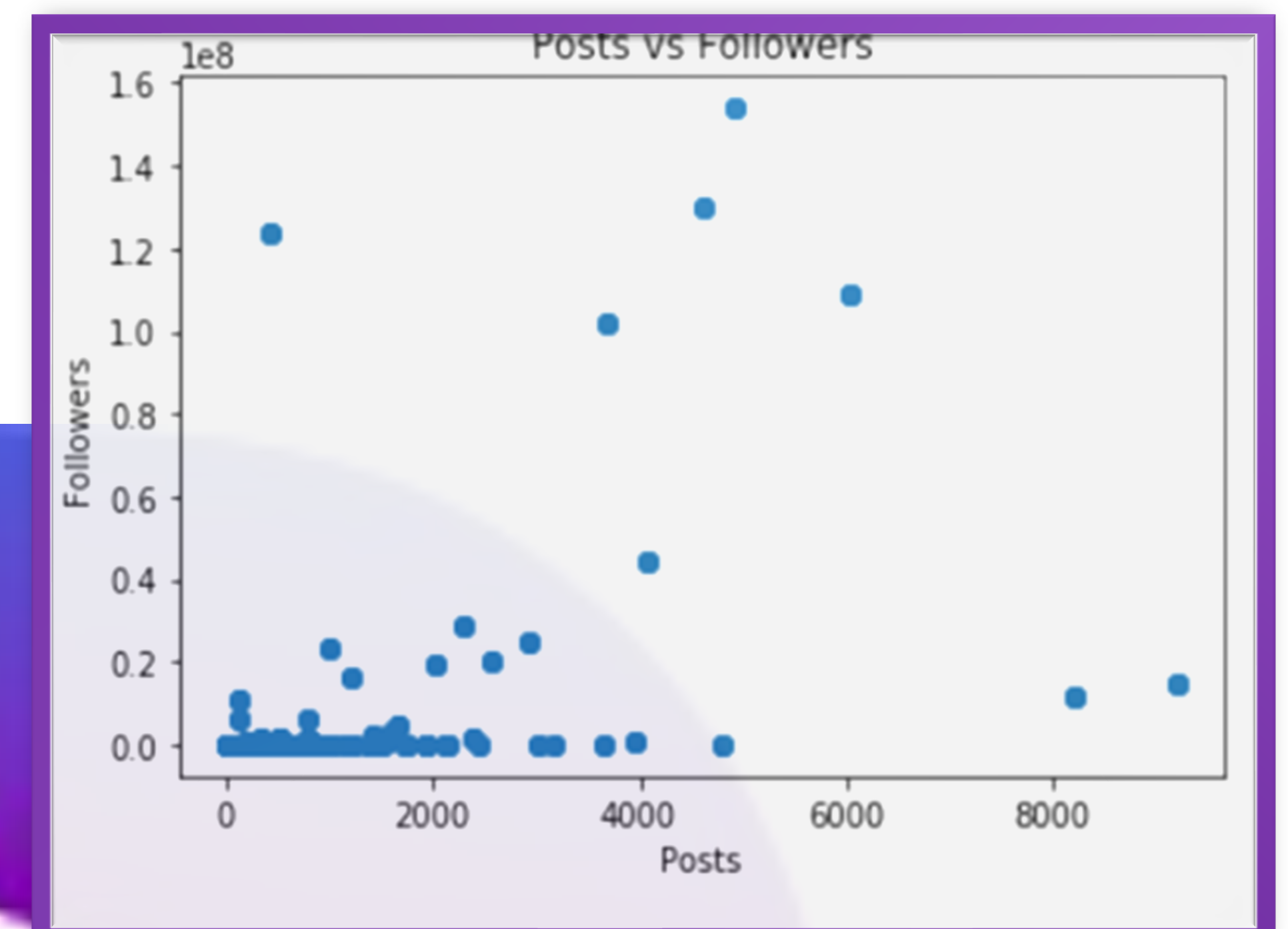
The Random Forest algorithm displays very sporadic results. It sometimes predicts you will get more likes than you have followers. Other times it gives fairly accurate results.

- KNN:

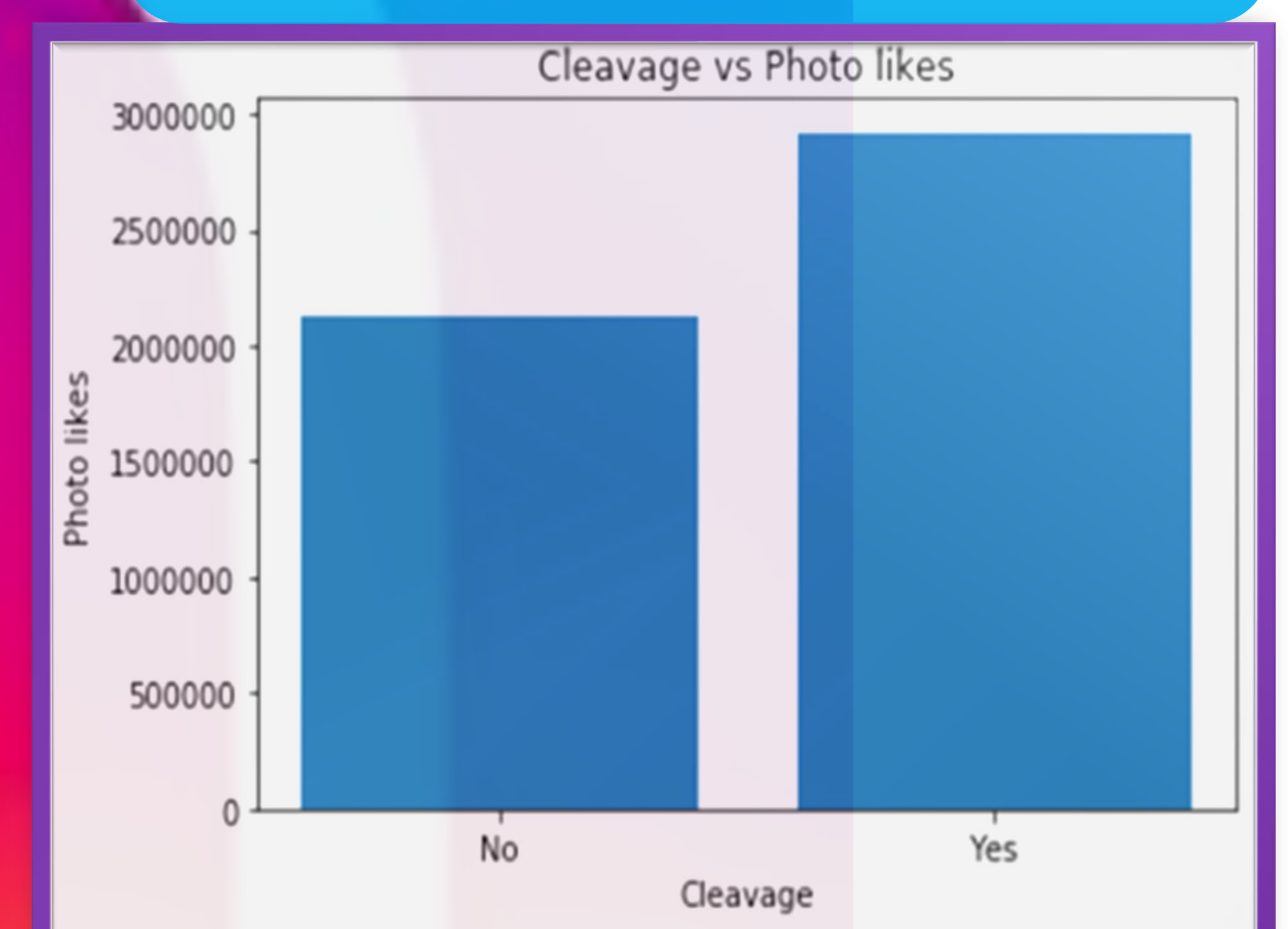
The KNN algorithm always gives accurate results, but if you want to predict more than one picture with different features, the algorithm will predict the same number of likes for every picture.

- Decision Tree:

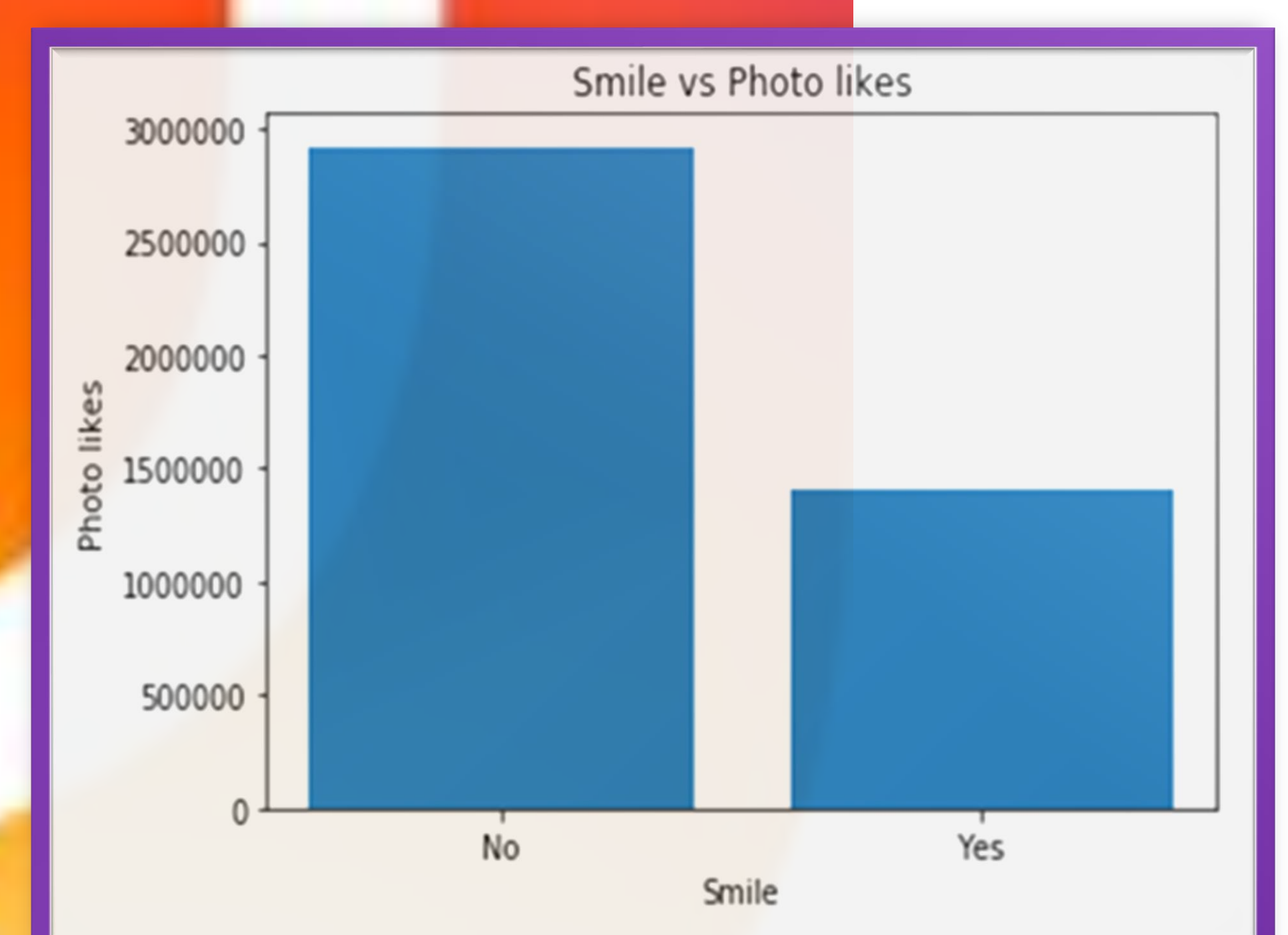
The Decision tree Algorithm was the worst algorithm at predicting the number of likes out of the three. It almost never gave an accurate prediction. It was either significantly greater or less than an accurate prediction.



This graph shows that there is no correlation between how many post a person has vs the amount of followers they have.



This graph shows that if a women shows cleavage, she is more likely to receive a higher number of likes.

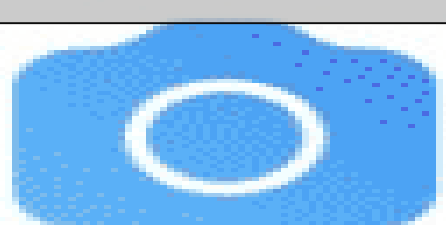


From my dataset more people that had higher likes did not smile, but more people smile in their photos than not

- Conclusion:

This project predicted likes based off of 19 personal features. However there are many, many different variables that can affect the data outside of these features and therefore these predictions are not always trustworthy.

Instagram Scraper used:
<https://github.com/realsirjoe/instagram-scraper>
Website used to help find most followed people on Instagram:
<https://www.trackalytics.com/the-most-followed-instagram-profiles/page/1/>



Write a message...

