PROJECT MILESTONE 6

PACE: Plan Stage

- What am I trying to solve or accomplish?

I'm trying to build a machine learning model that can be used to determine whether a video contains a claim or whether it offers an opinion.

- What resources do I find myself using as I complete this stage?

Jupyter notebook.

- Do I have any ethical considerations at this stage?

It's very important to identify videos that break the terms of service, even if that means some opinion videos are misclassified as claims. The worst case for an opinion misclassified as a claim is that the video goes to human review. The worst case for a claim that's misclassified as an opinion is that the video does not get reviewed and it violates the terms of service. A video that violates the terms of service would be considered posted from a "banned" author.

- What metric should I use to evaluate success of my business/organizational objective? Why?

It's better for the model to predict false positives when it makes a mistake, and worse for it to predict false negatives. Predicting opinion when actually it is claim is worse scenario. Therefore, the goal is to minimize false negatives which leads the use of the 'recall score' metric.

PACE: Analyze Stage

- Why did I select the X variables I did?

Previous investigation into the available data revealed that video engagement levels were highly indicative of claim status. Therefore, I selected all 5 count variables plus 'video_duration_sec', 'text_length', and two encoded variables 'verified_status' and 'author_ban_status' because I believe that those variables have the power to predict the outcome variable 'claim_status'.

- What has the EDA told me?

EDA reveals that there are 298 missing values in 7 features. There are no duplicates.
All 5 count variables have a lot of outliers. Because I plan to build tree-based model which is resistant to outliers, there is no need to impute or drop the outliers.

PACE: Construct Stage

- How well do the models fit the data? What are their validation scores?

The Random Forest model made only 25 mis predicts out of 3.817 predictions on the validation set.
The XQBoost model made only 26 mis predicts out of 3.817 predictions on the validation set.
Therefore, I choose the Random Forest as a champion model to predict over the test set

PACE: Execute Stage

- What key insights emerged from the models?

The champion model (RF) performs even better on the test set - it makes only 19 mispredictions that the actual video will be opinion, when it is actually claim.

- What are the criteria for model selection?

The only criterion for model selection was the recall score on the validation set.

- Does my model make sense? Are my final results acceptable?

I would strongly recommend using the RF model because it performed well on both the validation and test holdout data. Furthermore, both precision and F1 scores were consistently high. The model very successfully classified claims and opinions.

- Do I think my model could be improved? Why or why not? How?

Because the model currently performs nearly perfectly, I think it is not necessary to seek for future improvements.

- What business/organizational recommendations do I propose based on the models built?

Because the model's most predictive features were all related to the user engagement levels associated with each video, it was classifying videos based on how many views, likes, shares, and downloads they received.

- Given what I know about the data and the models I was using, what other questions could I address for the team?

It would be helpful to have the number of times the video was reported, and the total number of user reports for all videos posted by each author.

- Is the model ethical?

The model is ethical because it is focused on minimizing the worst cases for a claim that's misclassified as an opinion when it is actually a claim, thus violated the terms of service. A video that violates the terms of service would be considered posted from a "banned" author.

- When the model makes a mistake, what is happening? How does that translate to my use case?

The very rare occasions when the model makes mistakes are the cases when it predicts opinion when actually it is claim (false negatives), and vice versa (false positives).