TikTok project: detect claims and opinions

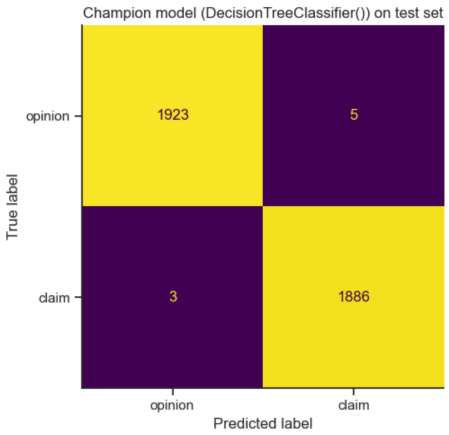
Milestone #5: Classification of claims and opinions.

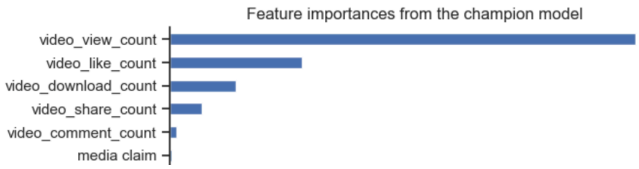
Overview

The objetive of this project is to implement a machine learning model to classify the videos uploaded to TikTok into claims (unsourced statements) and opinions (personal thoughts) in order to reduce moderating times and reduce economical costs.

Objective

For this milestone, we evaluated two ensemble classifiers for the classification of claims and opinions. First, we applied basic EDA to remove missing values and winsorize outliers. Then, we create new features from the video transcription text feature using natural language processing. Finally, we applied grid-search cross-validation to find the two optimal Random forest and XGBoost classifiers. The selected the champion model using a validation set and we estimated its generalization capability on the test set.





- Both the Random forest and the XGBoost classifier obtained close to perfect performance on the validation set. However, the performance of the Random forest was slightly superior and therefore, this model was considered the champion.
- The performance of the champion model on the test set was also close to perfect. Consequently, we can conclude the model generalizes soundly to unseen data.
- Since the performance of the model was outstanding, no need more additional tuning is required.
- Regarding feature importance, the champion model gave the most feature importance to the engagement variables: likes (the most important), shares, comments and downloads. This agrees with the previous analyses presented in former milestones.
-

Next Steps

In case the stakeholders were interested in continuing with the analysis, we recommend the investigation of the different user profiles as indicated by the ban status. Further analyses could also be carried out in order to understand the differences in the transcription texts between claims and opinions. The classification objective has already been accomplished.