TikTok Claims Classification Project

Results of Logistic Regression Model

ISSUE / PROBLEM

TikTok aims to efficiently prioritize claim reports

Objective: develop a reliable machine learning model which effectively classifies claim reports in order to streamline their processing

Steps: prepare data for modelling, build regression model

RESPONSE

The data team:

- Confirmed that there is a statistically meaningful difference between `'video_view_count'` and `'verified_status'`
- Corrected for imbalance in the `'verified_status'` variable
- Built a robust logistic regression model to reveal how video characteristics relate to the variable `'verified_status'`
- Evaluated model

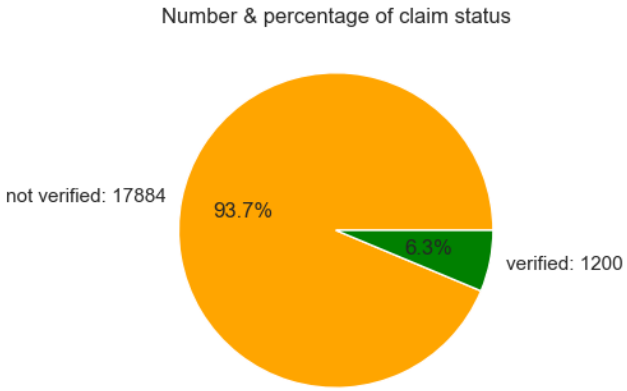
IMPACT

Results of the model exceeded expectations by achieving a recall score of 82%.

Recall measures the model's efficacy in predicting whether a video is coming from a verified account or not.

Recall	Precision	Accuracy	F1
0.82	0.64	0.68	0.72

UNDERSTANDING VERIFICATION



- Almost 94% of authors are not verified
- EDA revealed that Opinion videos were more likely to be posted by verified authors

MODEL EVALUATION

With 3681 True positives, the model did a very good job at predicting videos posted by verified authors

Upper left: # of videos posted by not verified authors accurately predicted. Lower right: # of videos posted by verified authors accurately predicted. Upper right: # of videos posted by not verified accounts inaccurately predicted. Lower left: # of videos posted by verified accounts inaccurately predicted.

KEY INSIGHTS

In a logistic regression, the coefficients represent the odds of an event occurring per 1-unit of change in a predictor variable. Here the coefficient for:

- `'claim_status_opinion'` indicates that moving from Claim to Opinion is associated with an increase in the probability that a video is posted by a verified account
- `'author_ban_status_banned'` is negatively correlated to the outcome that the video is posted by a verified account

Perhaps most interesting is that `'likes_per_view'` is negatively correlated with verified accounts.