Machine Learning Model Results
Executive summary report for TikTok project

Overview

The TikTok data team aims to create a machine learning model for categorizing videos into claims or opinions. Initial analysis of the data indicated that video engagement strongly correlated with claim status. The team is confident the model will fulfill performance expectations.

Problem

TikTok encounters numerous user reports on its videos for various reasons. Due to limitations, not all reported videos can be manually reviewed by moderators. Videos with claims (rather than opinions) are more likely to potentially violate the platform's terms. TikTok is exploring ways to detect claim-making videos to prioritize them for review.

Solution

Two tree-based classification models were developed and evaluated on a validation dataset. The model with the highest recall score was selected and then applied to a test dataset to estimate its future performance.

Results

Both model types, namely Random Forest (RF) and XGBoost, exhibited impressive performance. The RF model stood out with an outstanding recall score of 0.995, making it the preferred choice. On the test holdout dataset, the performance was almost flawless, with just five misclassified instances out of a total of 3,817.

Further examination confirmed the anticipated outcome: the main predictors were closely linked to video engagement metrics. Among these predictors, parameters such as video view count, likes, shares, and downloads encompassed the majority of predictive patterns in the dataset. Based on these findings, we can infer that videos with elevated engagement were significantly inclined towards being claims. Notably, no opinion-oriented video exceeded 10,000 views

Random forest - test set

Feature Importances
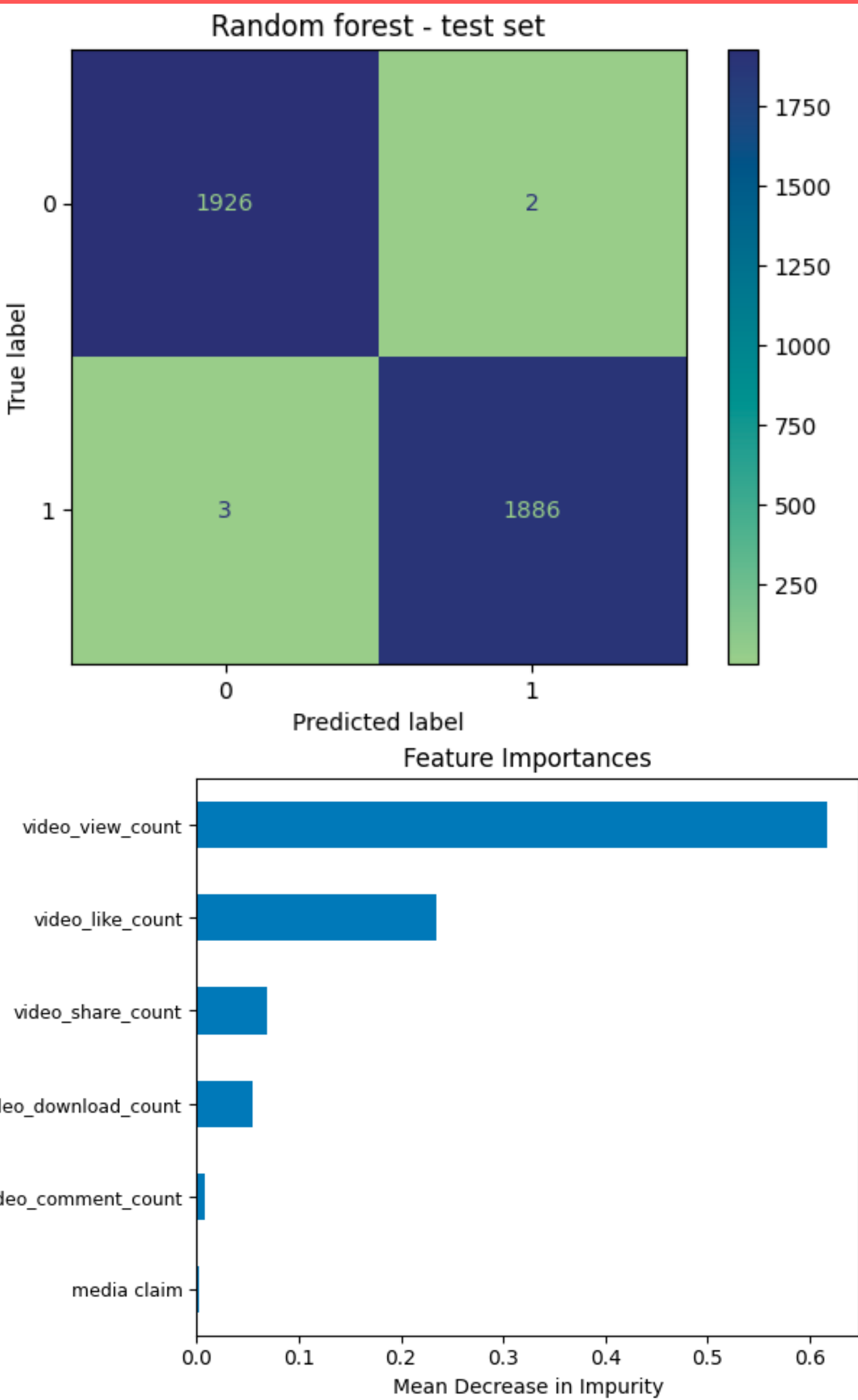
Next Steps

1. The model's strong performance on test holdout data supports its reliability. However, before proceeding with deployment, the data team strongly advises conducting further evaluations using additional subsets of user data to ensure the model's consistency and generalization across various scenarios.
2. The data team also emphasizes the importance of continuous monitoring of the distributions of video engagement levels. This proactive approach will help to safeguard the model's robustness against potential feature fluctuations, ensuring that it continues to deliver accurate results over time.