Executive Summary: TikTok Video Claim Detection Project

Project Overview

Our team has completed an initial data exploration and cleaning of the TikTok video dataset, assessed key variables, and performed a preliminary analysis of relationships between these variables. This work lays the foundation for building a machine learning model to identify claims versus opinions in TikTok videos.

Key Insights

1. Dataset Overview:
   - We have information on 19,382 videos with 12 different features for each.
   - The dataset is well-balanced: about half the videos contain claims, half contain opinions.
   - Very few data points are missing, giving us a complete picture to work with.
2. Patterns in Claim vs. Opinion Videos:
   - Videos with claims are viewed about 100 times more often than opinion videos.
   - Claim videos get more user interaction: 30% of views lead to likes, compared to 20% for opinions.
   - TikTok is more likely to flag or review authors of claim videos than opinion videos.
3. User Engagement Trends:
   - Users interact more (likes, comments, shares) with claim videos than opinion videos.
   - Videos shared more often are more likely to be reviewed by TikTok for inappropriate content.
   - How much users engage doesn't depend on whether the video author has been banned or is under review.
4. Data Quality Notes:
   - Some numbers aren't stored in the right format, but this is easy to fix.
   - A few videos are unusually short, which we'll look into.
   - Whether an account is verified doesn't seem very important for our purpose.
   - We can remove some unnecessary information to simplify our dataset.

Details

	video_duration_sec	video_view_count	video_like_count
count	19382	19084	19084
mean	32	254708	84304
std	16	322893	133420
min	5	20	0
25%	18	4942	810
50%	32	9954	3403
75%	47	504327	125020
max	60	999817	657830

	video_share_count	video_download_count	video_comment_count
count	19084	19084	19084
mean	16735	1049	349
std	32036	2004	799
min	0	0	0
25%	115	7	1
50%	717	46	9
75%	18222	1156	292
max	256130	14994	9599

Statistical summary

Next Steps

1. Clean up the data (fix number formats, check unusual videos, remove unnecessary information).

2. Analyse the text of video transcripts to find patterns in how claims and opinions are phrased.

3. Create new ways to measure user engagement that might help spot claims vs. opinions.

Executive Summary: TikTok Video Claim Detection Project

Exploratory Data Analysis phase

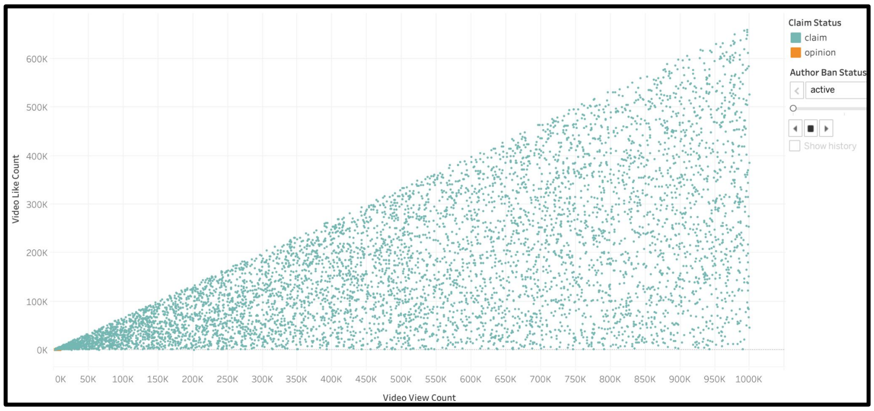
Project Overview

Our project aims to analyze a TikTok video dataset to distinguish between claim and opinion videos. Through exploratory data analysis, we have examined key engagement metrics and identified patterns in user interactions, preparing the data for a machine learning model to enhance video classification.

Key Insights

- Our analysis of 19,382 TikTok videos revealed that claim videos consistently receive higher engagement than opinion videos across all interaction metrics, such as views, likes, shares, and comments.
- The data is skewed towards videos with high engagement, resulting in a number of outliers. However, those values reflect typical behavior on social media platforms and provide valuable insights into user interaction patterns.
- Only 1.5% of the dataset contained missing information, which was removed without affecting the overall analysis.
- We found a strong link between an author's ban status and their video's popularity, with banned or under-review accounts receiving significantly higher views.
- Verified users are more likely to post opinion videos than claims, a trend that may influence content strategies.

Details



Likes vs Views classified by video claim status and active authors

Next Steps

1. We recommend retaining the outliers as they offer critical insights into user behavior.
2. To handle the skewed data, we will use techniques like the median and interquartile range for further analysis.
3. The next phase will focus on developing machine learning models to classify videos as claims or opinions based on these insights.

Analysis of Video Views for Verified vs. Non-Verified TikTok Users

Hypothesis testing to evaluate video views based on user verification status

Overview

Continuing with our claim classification project for TikTok, and after concluding the Exploratory Data Analysis phase, we assess whether verified users tend to receive different levels of video engagement compared to non-verified users. Descriptive statistics and hypothesis testing were employed to understand it.

Problem

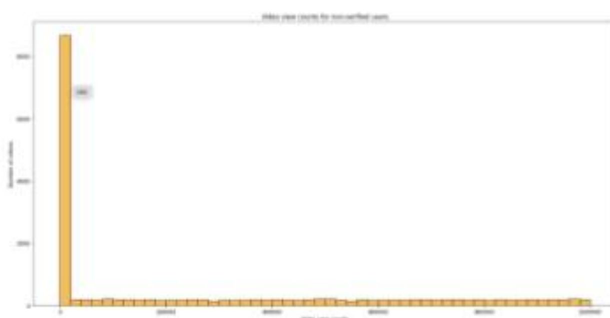
TikTok needs to determine whether there is a statistically significant difference in video views between verified and non-verified users on TikTok. Despite the suggestion of using a two-sample t-test for it, due to the non-normal distribution of video views, we consider that a more appropriate statistical method must be used.

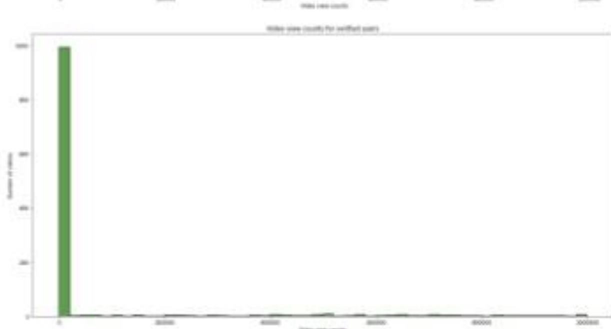
Solution

Results with a two-sample t-test indicated that at 5 % significance level, **verified and non-verified users receive different numbers of video views**. However, due to the non-normal distribution of video views, a more appropriate statistical method is needed to confirm these findings. We propose using **Mann-Whitney U test**, which does not assume a particular distribution.

Details

1. General statistics:
    - Verified Users: Mean video views = 91,439, Std. Dev. = 221,139
    - Non-Verified Users: Mean video views = 265,664, Std. Dev. = 325,682
    - Both distributions do not follow a normal distribution and are highly skewed to the right, with a small number of videos garnering a disproportionately high number of views, as expected in social networks.
2. Hypothesis Testing:
    - Null Hypothesis (H0): There is no difference in video views between verified and non-verified users.
    - Alternative Hypothesis (Ha): There is a significant difference in video views between the two groups.
    - Significance level: 5 %
    - Test: 2-sample t-test.





Next Steps

- Conduct additional non-parametric tests on other engagement metrics like video status, likes, comments, shares, to determine if the difference in user status extends beyond video views.
- Include verified status as a key variable in the machine learning model for claim classification. Verified status could affect user behavior and content engagement, influencing classification outcomes.

Verification Status Prediction on TikTok

Logistic Regression Analysis to Predict User Verification Status

OVERVIEW

This part of the classification study explores how various features of TikTok videos are associated with verification status and seeks to develop a model to predict whether a video belongs to a verified or unverified user. A logistic regression model was applied to identify significant predictors and assess its performance.

PROJECT STATUS

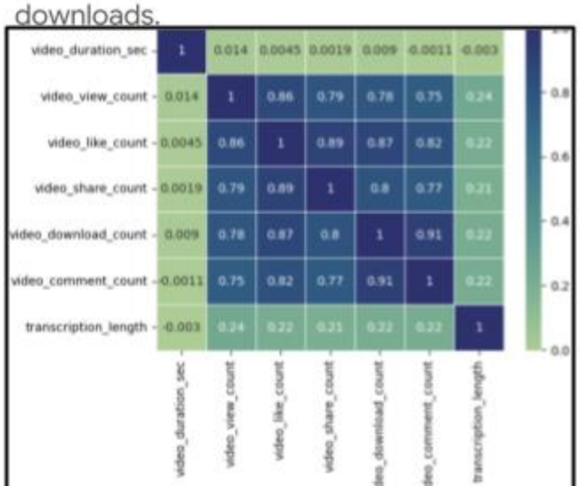
- The data set includes 19,382 videos with 12 features; 298 entries with missing data were removed. Since only 6,29 % of the videos were from verified users, these ones were resampled by allowing repetitions to obtain an equal number (17884) for each.
- Variables were processed for modeling:
  - Categorical variables were one-hot encoded.
  - Highly correlated variables were excluded to address multicollinearity.
- Logistic regression was trained and tested, achieving Recall of 84% and Precision of 61%.
- The model struggled to correctly classify non-verified users, with a false positive rate of 56%.

NEXT STEPS

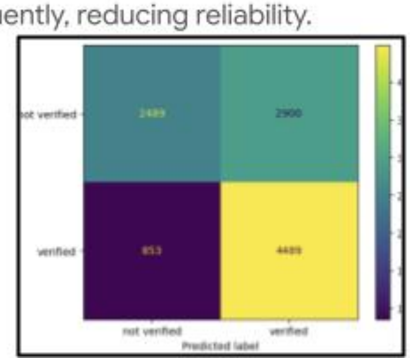
1. Model Refinement:
   - Exclude non-significant variables like the number of comments.
   - Experiment with models less sensitive to multicollinearity.
2. Data Improvements: Address class imbalance by resampling or collecting more data for verified users.
3. Additional Research:
   - Explore keyword analysis in video transcriptions to enhance prediction capabilities.
   - Investigate predictors beyond engagement metrics, such as user activity history.
4. Ethical Considerations: Ensure the model does not inadvertently promote bias or lead to impersonation risks.

KEY INSIGHTS

- Verified users account for 6.29% of the dataset, highlighting class imbalance.
- Strong correlations exist among engagement metrics, e.g., likes, views, shares, and downloads.



- Key findings from logistic regression:
- The odds of a video being from a verified user increase by 0.85% per additional second of video duration.
- The number of comments was not a statistically significant predictor (p = 58.1%).
- The model misclassifies non-verified users frequently, reducing reliability.



Machine learning model results

Final executive summary for TikTok project

Overview

The TikTok data team built a machine learning model to classify videos as either claims or opinions. Identifying content that presents a claim is critical, as it is more likely to require fact-checking and moderation. The model aims to help TikTok prioritize moderation resources and ensure platform integrity.

Problem

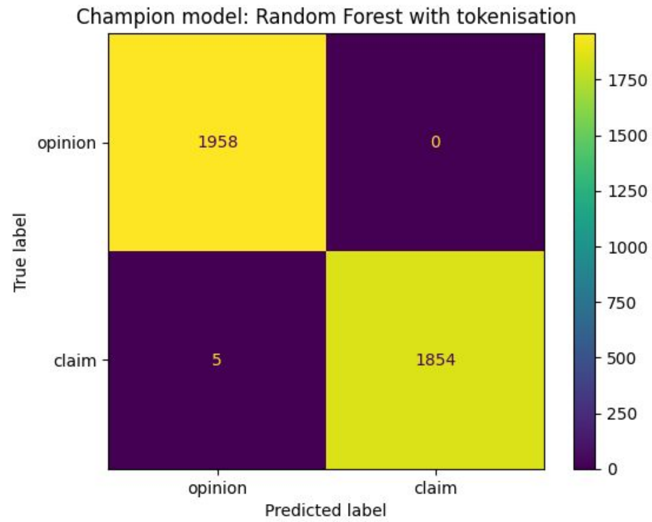
Due to the volume of user reports and the risk of misinformation, TikTok needs an efficient method to automatically identify videos containing claims or opinions. Moderators cannot manually review all videos. The ones presenting claims are more likely to violate community guidelines and cause harm.

Solution

The team developed two models—Random Forest and Gradient Boosting—to classify TikTok videos. Engagement metrics (views, likes, shares, downloads, comments) author and verification status and text features (length and key word pairs) were used as predictors. Model selection was based on recall to minimize the risk of missing true claims.

Details

Both models performed well, but the Random Forest model outperformed Gradient Boosting achieving over 99 % in all metrics. In particular, recall was of 99.73 % and 100 % precision on test data. This means it accurately detected almost all claim videos—only 5 wrong out of 1,859—whilst no opinion videos were incorrectly flagged as claims. Engagement metrics were the strongest predictors; for example, no opinion video had more than 10,000 views. Author ban status, transcription length and text patterns like “media claim” and “friend read” also contributed to prediction accuracy.



Next Steps

The team recommends testing the model on larger, unseen data to ensure consistent performance. Monitoring engagement feature distributions over time will help maintain model reliability.

Ethical risks, such as unjustified censorship, should be mitigated through human oversight and transparency in classification criteria.