

TikTok Claims Classification

Milestone 3

In Milestone 3 of the TikTok Claims Classification project, we focused on **Exploratory Data Analysis (EDA)** and visualizations to understand various engagement metrics associated with TikTok videos. The goal of this milestone was to gain insights into the data and explore how engagement metrics such as **video views**, **likes**, **comments**, and **shares** correlate with **claim** and **opinion** videos. These insights will inform the next steps for model development and content classification.

Key Insights

Video Duration Trends

Most videos are between 20 and 45 seconds, with common peaks at 10, 20, 40, and 60 seconds. Short, standardized durations suggest a structured posting behavior among users.

Highly Skewed Engagement Metrics

Metrics like views, likes, shares, comments, and downloads show strong right-skewness. A small portion of videos accumulates very high engagement while the majority receive low interaction. This is consistent with viral social media patterns.

Dominance of Claim Content

Claim videos account for 99% of total video views. Opinion videos receive minimal engagement, emphasizing claims' higher potential for reaching large audiences.

Author Ban Status and Views

Banned and under-review authors have significantly higher median view counts compared to active authors. This suggests that controversial or policy-violating content may achieve more engagement before moderation actions occur.

Verification Status and Posting Behavior

Verified users are more likely to post opinions, while unverified users are more likely to post claims. Verification appears to correlate with more cautious or opinion-based content creation.

Details

