Predicting Verified User Status on TikTok using Binomial Logistic Regression

Insights from Claims Classifications data

Project Overview

A Binomial Logistic Regression Analysis was conducted to examine how video characteristics relate to a user's verified status. The goal was to uncover interpretable factors that influence verification likelihood, supporting transparent and data-driven content moderation decisions.

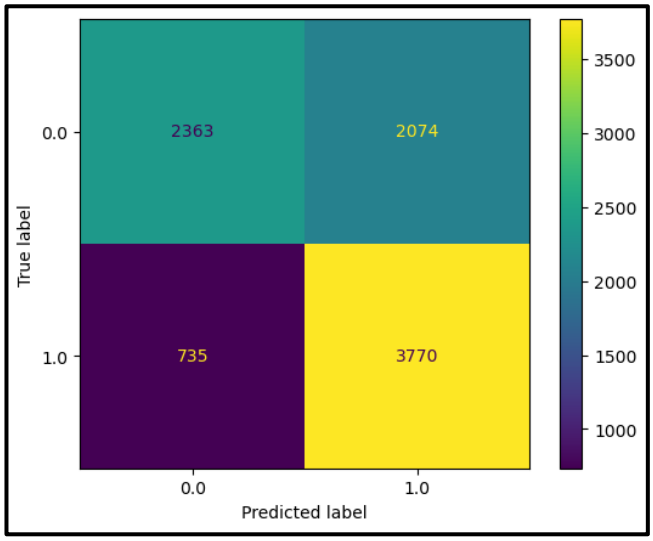
Key Insights

- The Model achieved **moderate predictive performance**: **Accuracy (69%), Precision (65%), Recall (84%), F1 Score (73%).**
- Better performance in identifying unverified users (**True Negatives = 2363**), compared to verified ones (**True Positives = 3370**), though some misclassification remain**.**
- Higher **Recall** for unverified users than for verified ones, suggesting sensitivity but room for improved **Precision**.

- **Positive Influences on Verification Likelihood:**
- claim_status_opinion (**1.683**) – Opinion content more likely linked to verified status.
- video_share_count (**5.09e-06**) – Shared Videos correlate positively with verification.
- video_comment_count **(0.0005)** – Higher comment counts are weakly associated with an increased likelihood of verification.

- **Negative Influences on Verification Likelihood:**
- video_download_count (**-0.0001**) – More frequent downloads are weakly associated with a lower likelihood of verification.
- video_duration_sec (**-0.0009**) – Longer video durations are weakly associated with a lower likelihood of verification.
- author_ban_status_banned (**-0.379**) – Enforcement history lowers verification probability.
- author_ban_status_under review (**-0.125**) – Enforcement history lowers verification probability.
- video_view_count (**-1.8074994377791O6e-07**) – High views correlate with lower verification likelihood.

Details



Analysis suggests that while sharing and opinion-based content increase the likelihood of verification, higher video views and a history of bans are associated with reduced verification chances. These results indicate that content type and engagement nature matter more than sheer popularity when predicting verified status.

Next Steps

1. **Address class imbalance** to improve precision
2. **Test advanced models** such as Random Forest or XGBoost to improve accuracy.
3. **Engineer new features** to capture engagement quality and context.
4. **Integrate model into moderation workflows** as a decision-support tool rather than a sole determinant of verification.