# Executive Summary (Task 5)

## Overview

The TikTok data team seeks to develop a machine learning model to assist in the classification of claims for user submissions. Earlier, the data team observed that if a user is verified, they are much more likely to post opinions. Since the end goal is to classify claims and opinions, it’s important to build a model that shows how to predict the behavior of the account type (verified) that tend to post more opinions. So, in this part of the project, the data team built a logistic regression model that predicts verified\_status (verified or not verified.)

|  |  |
| --- | --- |
| **Key Insights**   * **Accuracy (65%):** the model correctly classified 65% of all predictions, whether "verified" or "not verified". * **Precision (61%):** the model is able to correctly predict 61% of the cases labeled as ‘not verified’. * **Recall (84%):** the model successfully identified 84% of the actual ‘not verified’ cases. * **F1-Score (71%):** indicating better performance in detecting not verified accounts.   **What this means:**  -The model is better at predicting "not verified" outcomes than it is at predicting "verified" ones.  -We might need to refine the model further if we want to improve the accuracy for "verified" predictions, depending on the importance of these classifications. | **Details**  **Evaluation Metrics report** **ROC curve**  \*The model is better than random guessing but still has room for improvement in terms of distinguishing between "verified" and "not verified" cases. The AUC score of 0.65 shows it’s somewhat effective but not perfect.  **Coefficients**  \*Based on the logistic regression model, video duration has the highest coefficient. Each additional second of the video is associated with 0.009 increase in the log-odds of the user having a verified status. In other words, if we exponentiate e^0.009 = 1.009, then for each additional second of video duration, the odds of being verified increase by about 0.9%. |
| **The next steps**  The next step is to construct a classification model that will predict the status of claims made by users. That is the final project and original expectation from the TikTok team. Now, there is enough information to analyze the results of that model with helpful context around user behavior. | |