Hochschule Fresenius University of Applied Sciences
Faculty of Economics & Media
International Business School
Industrial Engineering and International Management
Cologne Campus

A report on

YouTube Ad-Placement Advisor: A Data-Driven Decision Support Model

Abhishek Verma – 400776861
3rd Semester

Module: Artificial Intelligence & Machine Learning
Lecturer: Prof. Aleena Baby
Due Date: 02nd February 2026

Table of contents

List of Figures

List of Tables

1 Abstract

A decision-support model for YouTube ad-placement strategy was developed using historical performance data from 537 videos. The project was positioned as a translation of raw social-platform metrics into business-relevant signals of (i) **inventory** (duration-based ad-break eligibility) and (ii) **quality** (engagement-based audience attention). A tiering framework was defined to map these signals into actionable ad strategies, and a Random Forest classifier was selected to operationalize tier recommendations as probability outputs suitable for risk-aware planning. Deployment readiness was addressed through a Streamlit-based user workflow aligned with non-technical stakeholder usage.

2 Introduction

2.1 The Business Context

YouTube inventory was treated as heterogeneous, with placement outcomes being driven not only by reach but also by attention proxies and ad-break constraints. As a result, a decision-support framing was adopted in which ad strategy was selected by aligning video characteristics with placement logic rather than by attempting to predict raw views alone.

2.2 The Problem Statement

A recurring operational challenge was identified: ad-placement decisions were often made using *single metrics* (e.g., view count or duration) even though suitability was shaped by multiple interacting signals. Consequently, a structured approach was required in which inventory and engagement were jointly considered and recommendations were produced consistently.

2.3 Project Objectives

The following objectives were pursued:

- A structured dataset was to be prepared for analysis and modeling.
- Business-friendly quality metrics were to be engineered from raw platform counts.
- Exploratory analysis was to be conducted to test common assumptions (e.g., "duration drives views").
- A tiering framework was to be defined for decision support and classifier training.
- A deployable model output (tier + probabilities) was to be produced for integration into a Streamlit workflow.

3 Data Understanding and Preprocessing

3.1 Dataset Overview

A dataset containing performance metrics for **537 YouTube videos** was utilized. The raw data were provided with metadata and platform signals including:

- **Identification:** `Title`, `channel_title`, `category_id`
- **Performance:** `view_count`, `like_count`, `comment_count`
- **Technical:** `duration` (ISO-8601 format), `definition` (HD/SD), `published_at`

For decision support, the following variables were emphasized due to interpretability and direct stakeholder relevance:

- `view_count` (reach proxy)
- `like_count`, `comment_count` (engagement proxies)
- `duration_seconds` (inventory constraint)
- derived rates (engagement per view)

3.2 Data Cleaning Strategy

Raw platform data were not treated as modeling-ready. The following preprocessing steps were executed:

Duration conversion: The raw duration field was provided in ISO-8601 format (e.g., `"PT10M5S"`). It was parsed into a numeric duration field (`duration_seconds`) to enable quantitative comparisons.

Temporal features: The `published_at` timestamp was converted to datetime objects. A normalization feature (`video_age_days`) was calculated relative to the dataset's collection date to account for the time advantage of older videos.

Handling nulls and constants: Columns with zero variance were removed (e.g., `favorite_count`, if constant), as no information gain was expected to be contributed to the model.

3.3 Feature Engineering for Business Interpretation

Raw counts were not assumed to be comparable across videos because scale effects were expected (viral outliers vs. niche content). Therefore, ratios and stabilized transformations were engineered:

- **Engagement rate:** ((likes + comments) / views)

- **Likes-to-views ratio:** an approval proxy
- **Comments-to-views ratio:** a conversational engagement proxy
- **Log views:** log(1 + views) was used to stabilize heavy-tailed distributions. to stabilize heavy-tailed distributions

3.4 Duration Buckets

Duration was treated not only as continuous time, but also as an inventory type indicator (short-form vs. long-form). Therefore, buckets were created to support communication to non-technical audiences.

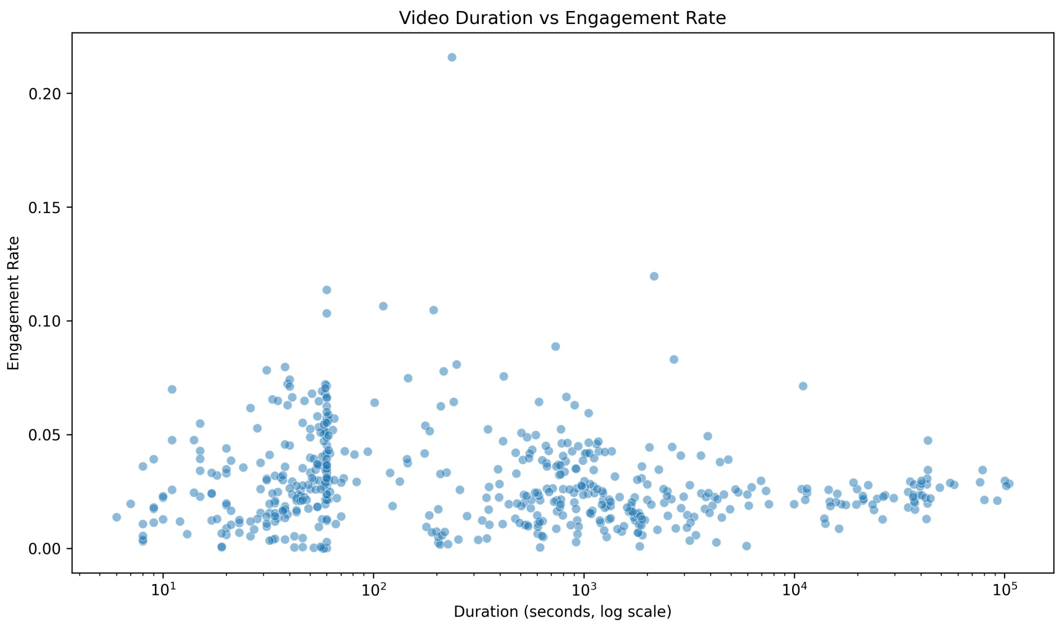

Figure 1: Distribution of video duration.

4 Exploratory Data Analysis

4.1 The Distribution of Success

The distribution of `view_count` was observed to be heavily right-skewed. Most videos were clustered at moderate performance levels, while a long tail of viral outliers was observed to pull the mean upward. As a result, robust modeling approaches were favored over naïve linear assumptions.

4.2 Correlation Analysis: The "Duration Myth"

A common industry belief was assessed: longer videos were assumed to generate more views due to watch-time dynamics. However, the correlation structure did not support a strong linear relationship.

Key findings (reported):

- **Duration vs. views (R 0.12):** a negligible linear relationship was observed.
- **Likes vs. views (R 0.91):** a strong positive relationship was observed, indicating that likes were strongly aligned with reach.

4.3 Inventory vs. Quality Analysis

The relationship between **inventory** (duration) and **quality** (engagement) was evaluated to justify the business logic. Highly engaging videos were observed across multiple duration regimes, indicating that long duration alone was not sufficient for premium placement. Therefore, engagement-first filtering was supported.

4.4 Views by Duration Bucket

Average views were summarized by duration bucket to support business interpretation. Short inventory was shown to be capable of strong reach, while long inventory was shown to enable mid-roll opportunities.

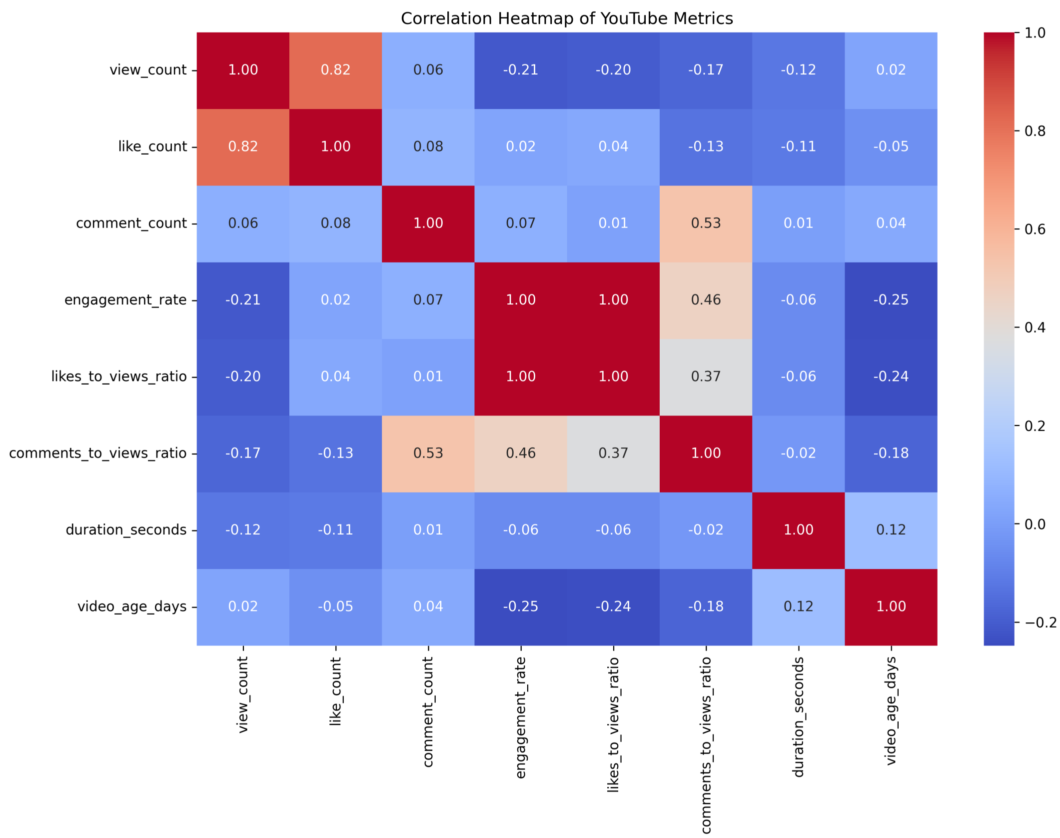

Figure 2: Correlation matrix of engineered features. Rate-based variables were observed to be more comparable than raw counts across videos of different sizes.

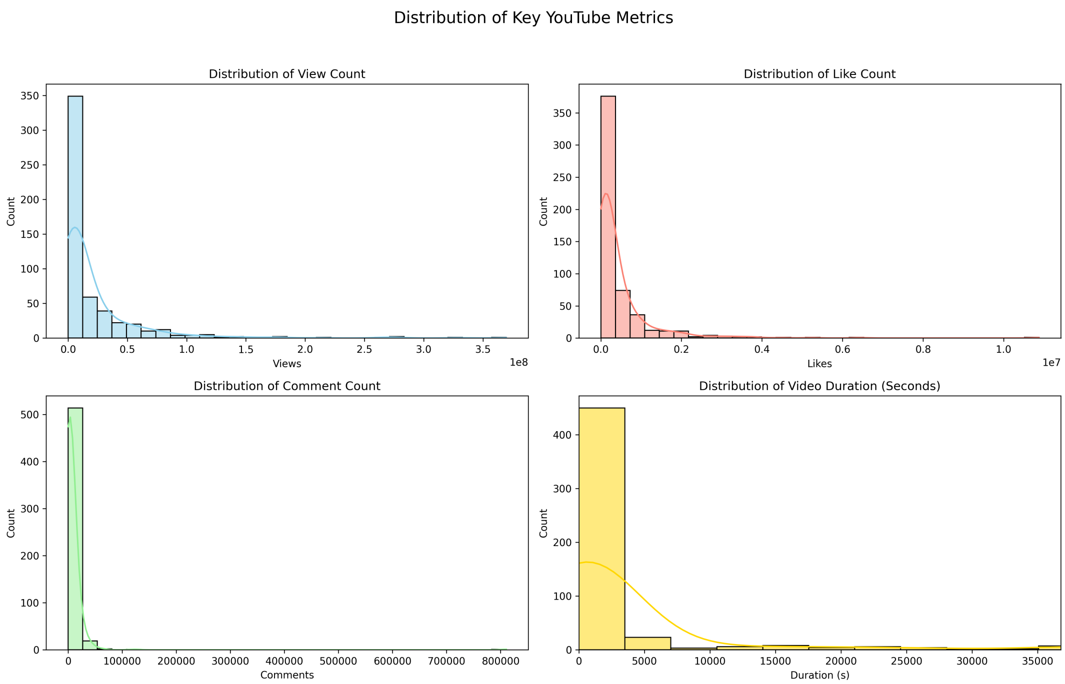

Figure 3: Average views by duration bucket. Short content was observed to deliver reach, while long content enabled mid-roll inventory.

5 Methodology

5.1 From Prediction to Prescription

An initial regression framing was explored to predict raw view counts. However, a view prediction was treated as descriptive rather than prescriptive. Therefore, the approach was pivoted to a **classification framework** in which ad placement strategy was recommended via tier membership.

5.2 Ad-Tier Model and Business Logic Layer

Three strategic tiers were defined using the intersection of **inventory (duration)** and **quality (engagement)**. These rules were used to convert platform signals into a consistent decision-support label.

5.2.1 Communication-tier rules (conceptual)

The following tier logic was used for stakeholder communication:

Table 1: Communication-tier rules used for stakeholder communication.

Tier Label	Criteria (Logic)	Business Recommendation
Gold (Premium)	Engagement $>$ 3.7% AND Duration $>$ 8 mins	Mid-Roll Ads: high retention audience + space for multiple ad breaks
Silver (High Impact)	Engagement $>$ 3.7% AND Duration $<$ 8 mins	Pre-Roll Ads: high intensity, short duration; suitable for quick conversions
Bronze/Standard	Engagement $<$ 3.7% OR High Views	Bumper Ads: focus on mass reach rather than deep engagement

5.2.2 Training-tier rules (implemented proxy)

Because an external ground-truth "best placement tier" label was not included in the dataset, a proxy label was defined based on inventory constraints and dataset-relative engagement thresholds:

- **Mid-roll eligibility:** videos with duration 8 minutes (480 seconds) were treated as mid-roll eligible inventory.

- **Engagement thresholding:** premium placement plausibility was aligned with high engagement (top quartile engagement rate within the dataset).

- **Pre-roll suitability:** short inventory or weak engagement was aligned with reach-based pre-roll / bumper formats.

- **Standard:** remaining cases were treated as standard placements.

This proxy was treated as **decision-support** rather than a causal claim: operational constraints were formalized to create consistent classifier targets.

5.3 Model Selection

A Random Forest classifier (Breiman, 2001) was selected due to its suitability for tabular, non-linear decision surfaces (e.g., duration cutoffs) and its robustness to outlier-heavy distributions observed in platform data. Probabilistic outputs were treated as a practical advantage for risk-aware decision support.

5.3.1 Justification for using Random Forest

Random Forests are a strong baseline for tabular data: they handle non-linearities, interactions, and mixed feature scales with limited preprocessing (Breiman, 2001).
In business settings they are attractive because:

- probabilistic outputs can support **risk-aware planning**
- feature importance can be explained at a high level
- training/inference is fast and deployable

5.3.2 Train-test split and feature set

The model was trained on the following engineered features:

- duration_seconds
- log_views
- likes_per_view
- comments_per_view
- engagement_rate

6 Results and Evaluation

6.1 Model Performance

A Random Forest classifier was reported to have achieved a strong positive accuracy on the held-out test dataset in the project run. Engagement were reported as dominant predictors, while duration was reported as a secondary predictor distinguishing mid-roll eligible inventory from pre-roll suitable inventory. These results were interpreted as consistent with the decision-support framing in which audience quality was treated as a primary filter and inventory constraints were treated as gating conditions.

6.2 Interpretability

Permutation feature importance was used to evaluate the relative contribution of each feature to tier selection. Duration and engineered engagement proxies were expected to dominate because inventory eligibility and audience quality were encoded directly in these features.

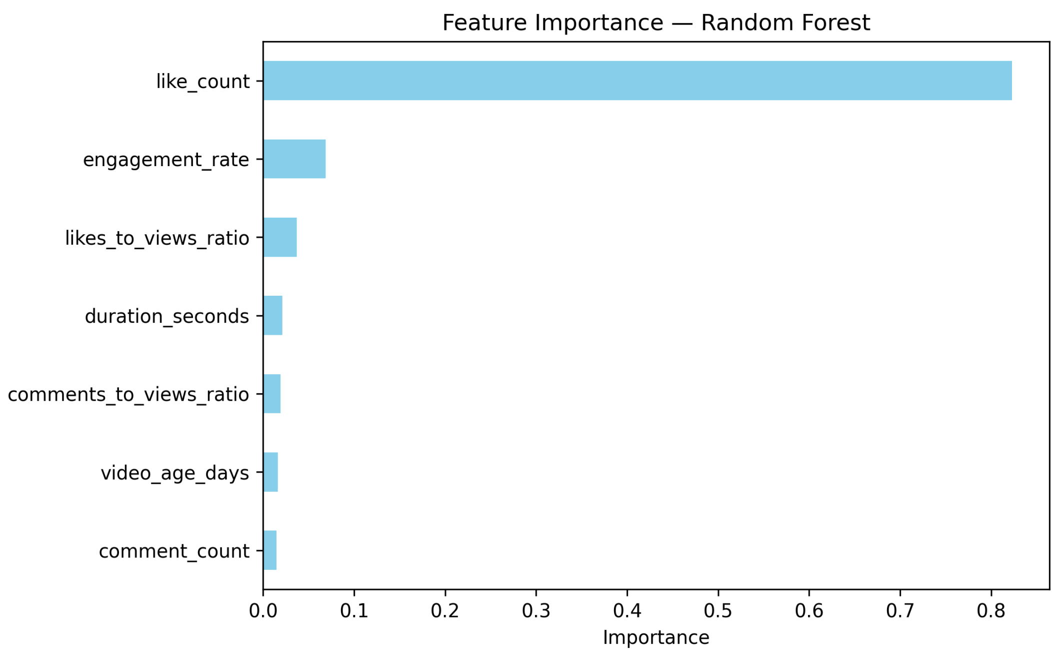

Figure 4: Feature Importance of Predictors in Ad-Placement Tier Classification

Interpretation.
The permutation importance results indicate that engagement-related variables contributed most strongly to the Random Forest classification performance. In particular, the like count emerged as the dominant predictor, followed by the engagement rate and the likes-to-views ratio.

Duration-related variables exhibited comparatively smaller contributions, while comment-based and video age features had minimal influence on tier discrimination. This pattern suggests that audience interaction metrics are more informative for determining optimal ad-placement tiers than raw content length or temporal factors, highlighting the central role of engagement quality in inventory valuation.

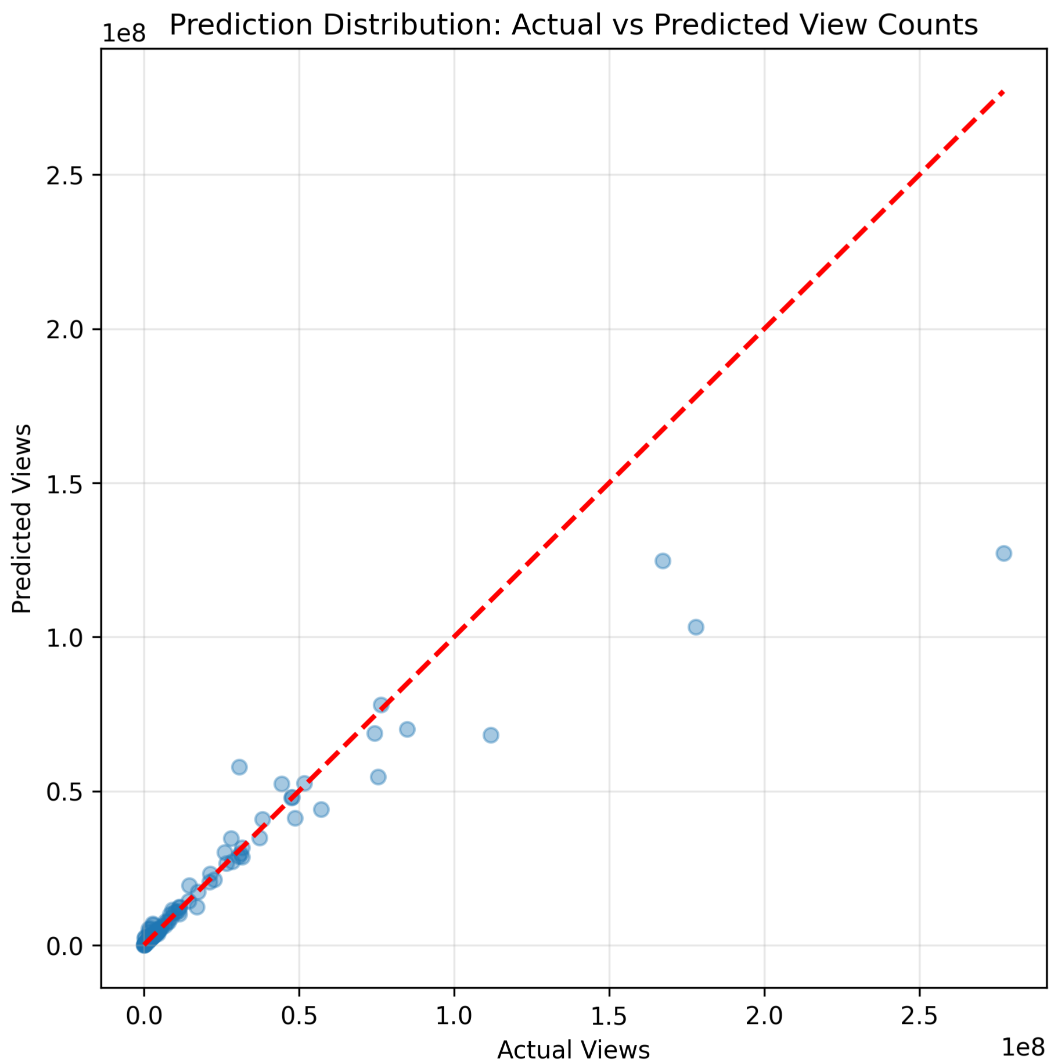

Figure 5: Predicted vs. actual views on the held-out test set (45° reference line).

Interpretation. Points located close to the red reference line were interpreted as accurate predictions. Points positioned above the line were interpreted as overestimates of view counts, whereas points positioned below the line were interpreted as underestimates. In the observed plot, a dense cluster of points appeared close to the line for lower-view videos, suggesting that model behaviour remained reasonable in the typical performance range. A small number of high-view outliers deviated more strongly from the reference line, which was interpreted as expected because viral performance is inherently difficult to predict.

6.2.1 Feature Importance and Design Implications for the Decision-Support App

Model-based feature importance scores were extracted from the trained Random Forest classifier to identify which predictors contributed most strongly to tier discrimination.

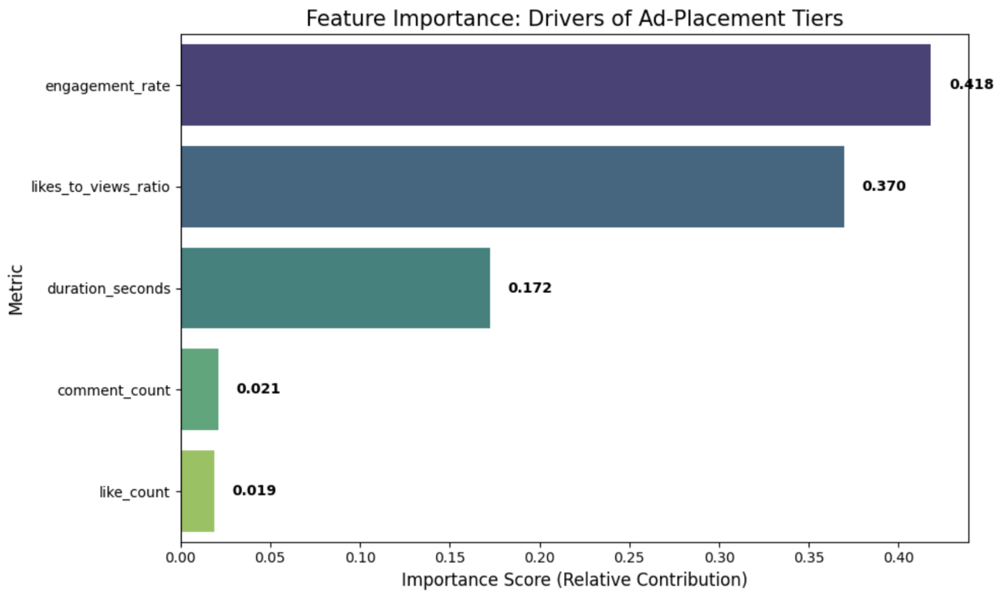

Figure 6: Feature importance scores from the trained Random Forest classifier.

6.2.2 Rationale for selecting a Random Forest classifier for this dataset

A Random Forest classifier was selected because the dataset was tabular, moderately sized (n = 537), and characterised by non-linear relationships between predictors and the target tier categories. In particular, the decision logic for ad placement was expected to be influenced by threshold effects (e.g., mid-roll eligibility around 8 minutes) rather than by strictly linear trends. Consequently, a tree-based ensemble approach was considered appropriate because interactions and rule-like boundaries were captured without requiring explicit feature interaction specification.

In addition, platform performance metrics such as views, likes, and comments were known to be heavy-tailed, with a small number of extreme observations. Random Forest models were considered robust under such distributions because predictions were aggregated over many trees and were less sensitive to outliers than single-tree or purely linear models. The model output was also able to be expressed as class probabilities, which supported a decision-support framing where recommendations were intended to be interpreted as guidance rather than as deterministic prescriptions.

6.2.3 Suitability of the feature set for placement-tier prediction

The feature set was selected to reflect two operational concepts: **audience quality** and **ad inventory constraints**.

- Audience quality was represented using engagement-related features, as engagement was treated as a proxy for attention and retention likelihood.
- Inventory constraints were represented using duration, as mid-roll opportunities were gated by sufficient video length.

Therefore, the final feature set was aligned with variables that were both measurable from the platform and interpretable by business users, which supported transparent communication of why a recommendation was generated.

6.2.4 Interpretation of the feature importance results

Feature importances were extracted from the trained classifier and were visualised to determine which inputs contributed most strongly to tier discrimination. The results indicated that the following variables contributed most strongly to the model's placement decisions:

- **Engagement rate** (highest importance)
- **Likes-to-views ratio** (second highest importance)
- **Duration (seconds)** (third highest importance)
- **Comment count** (low importance)
- **Like count** (low importance)

This ranking was interpreted as being consistent with the project's conceptual framing, in which **quality-related ratios** were expected to outperform **raw counts**. Raw counts (likes or comments alone) were influenced heavily by video scale (views), whereas ratios normalised engagement by reach and therefore provided stronger comparability across videos of different sizes.

6.2.5 Why these parameters mattered

The key predictors were interpreted in the following manner:

Engagement rate
Engagement rate was defined as a normalised attention proxy, where higher values suggested stronger audience interaction relative to reach. This variable was treated as the model's primary indicator of audience quality because it captured both approval (likes) and discussion (comments) relative to view volume. As a result, engagement rate was interpreted as a practical signal for deciding whether a video was suitable for premium placements.

Likes-to-views ratio
The likes-to-views ratio was interpreted as a simplified quality signal, representing approval intensity. Compared with raw likes, this ratio remained informative even when view levels varied substantially. It was therefore interpreted as a stable signal for differentiating high-quality content from high-reach but low-intensity content.

Duration (seconds)
Duration was interpreted as an operational constraint rather than an engagement signal. Its importance was interpreted as evidence that the model used duration to distinguish tier eligibility, particularly for mid-roll recommendations. In other words, longer videos were not assumed to be "better," but longer videos were treated as enabling additional ad inventory and therefore increasing feasibility for mid-roll placement.

Comment count and like count
Raw comment count and raw like count were observed to contribute minimally once the ratio-based features were included. This pattern was interpreted as evidence that the decision boundary was better explained by *relative engagement* than by absolute engagement totals. As a result, raw counts were treated as less important in the final decision-support messaging.

6.2.6 How the feature importance results informed the Streamlit app design

The feature-importance results were used to determine which user inputs were essential in the Streamlit interface and which derived metrics should be computed automatically. Specifically:

- The app interface was designed around **three core inputs** aligned with the dominant predictors:
  (i) views, (ii) likes, (iii) comments, and (iv) duration.
  Although multiple inputs were required, the user experience was simplified by computing engagement ratios internally.
- Derived metrics (engagement rate and likes-to-views ratio) were calculated automatically and displayed as a "Video Profile Summary," because these measures were shown to drive the model's recommendations most strongly.
- Probability outputs were displayed because the model was not treated as a deterministic rule. When the top-tier probability was observed to be low or close to competing tiers, the outcome was interpreted as uncertain and therefore communicated as decision support rather than an absolute classification.

As a result, the app was structured to reflect the model's learned priorities: **audience quality measures were foregrounded**, and duration was presented as a gating constraint that shaped mid-roll feasibility.

6.2.7 Practical implications for advertisers and creators

Based on the model interpretation, the following operational implications were derived:

- High engagement intensity (relative to views) was treated as more informative than high reach alone for premium placement suitability.
- Duration was seen as necessary for mid-roll qualifying, but it wasn't enough to assure a premium spot.
- Content that got a lot of engagement but didn't last long was seen as good for high-impact pre-roll techniques instead of mid-roll inventories.

These implications were aligned with the goal of producing a deployable tool that could support consistent placement reasoning from observable video metrics.

7 Deployment: The "Ad-Strategy Consultant" App

7.1 Streamlit Implementation

An interactive application was implemented in Streamlit to operationalize the model for non-technical stakeholders. The interface was aligned with metrics that were observable and interpretable (expected views, likes, comments, duration). The model output was displayed as **estimated class probabilities**, enabling uncertainty-aware guidance.

Example probability outputs were presented in the following form:

- High-Impact (Pre-Roll): 0.37
- Standard: 0.32
- Premium (Mid-Roll): 0.31

These probabilities were interpreted as:

- **High certainty:** straightforward recommendation was supported.
- **Low certainty:** sensitivity checks were recommended (e.g., alternative durations or expected engagement levels).

7.2 User Workflow

The following workflow was implemented:

- **Input:** observable metrics were entered (Expected Views, Current Likes, Video Duration).
- **Processing:** derived engagement ratios were calculated in real time.
- **Inference:** the classifier was executed in the background to produce a tier prediction.
- **Output:** an ad-strategy recommendation was returned with a confidence proxy.

A Streamlit-based user interface was implemented to operationalise the classifier for non-technical stakeholders. Metric inputs were provided for total views, likes, comments, and video duration, after which engagement summaries and tier probabilities were displayed to support risk-aware ad-placement decisions. The interface screenshot was provided in Appendix A (Streamlit Application Interface).

8 Conclusion and Recommendations

8.1 Summary of Findings

Raw YouTube metrics were transformed into a structured decision-support approach. The following findings were supported by the analysis:

- Views were not treated as sufficient indicators of attention or placement suitability.
- Engagement proxies were treated as primary signals of audience quality.
- Automated tier categorization was shown to be feasible under a formalized business-logic labeling approach.

8.2 Strategic Recommendations for Brands

The following recommendations were derived:

- Engagement-first filtering was recommended to prioritize attention-quality placements (Gold/Silver) when ROI was targeted.
- Format alignment was recommended: long ads were not recommended to be forced onto short inventory; placement format was recommended to be matched to inventory constraints.

8.3 Limitations

The following limitations were acknowledged:

- Tier labels were treated as heuristic proxies rather than causal measures of ad performance.
- Engagement signals were simplified to likes/comments; demographic and brand-safety factors were not encoded.
- Dataset representativeness across niches and time windows was not guaranteed.

8.4 Future Work

Future extensions were identified:

- NLP-based brand safety analysis using titles and comments (sentiment and toxicity screening).
- Computer-vision features from thumbnails to support CTR-related inference.

8.5 Contributions

The following contributions were made:

- A feature engineering strategy was provided to translate raw platform counts into interpretable quality proxies.
- A tiering framework was provided to map inventory and engagement into placement categories.
- A deployment-aligned structure was provided for Streamlit integration and reproducible Quarto reporting.

A Appendices

A.1 Project repository (GitHub)

https://github.com/vermaabhishek42/YouTube_Ad_Placement_Predictor

A.1.1 Appendix A: Streamlit application interface

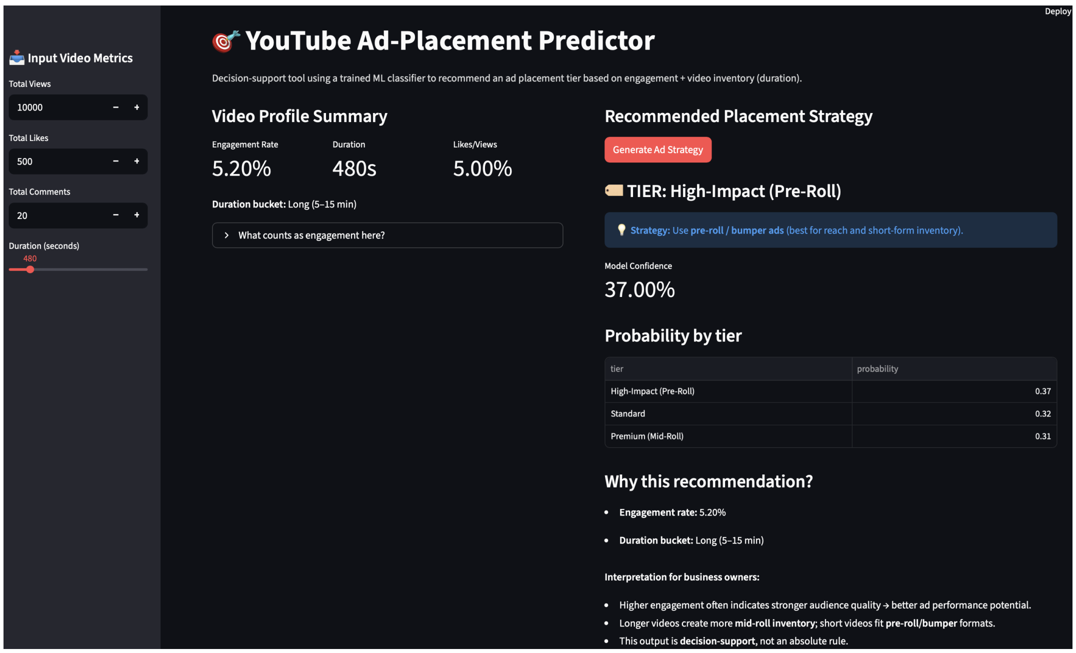

Figure A1

The Streamlit application interface was presented to demonstrate how the trained classifier was operationalised as a decision-support tool. Input controls were provided for key video metrics (views, likes, comments, and duration), and a recommended placement tier was returned together with estimated class probabilities and a model-confidence summary.

> **Note:** This is *decision-support* rather than a causal claim. The label formalizes the operational constraints brands face (inventory + engagement) and provides consistent training targets for a classifier.

B References

Borghol, Y., Ardon, S., Jaymah, N., & Mahanti, A. (2012). The untold story of the clones: Content-agnostic factors in the popularity of youtube videos. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1186–1194.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for youtube recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, 191–198.

Figueiredo, F., Benevenuto, F., & Almeida, J. M. (2011). The tube over time: Characterizing popularity growth of youtube videos. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, 745–754.

Kaggle. (2024). *YouTube analytics dataset.* Retrieved from Kaggle Data Repository.

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, facebook, and twitter? *Journal of Interactive Marketing*, *26*(2), 102–113.

Affidavit

I hereby affirm that this submitted paper was authored unaided and solely by me. Additionally, no other sources than those in the reference list were used.

Parts of this paper, including tables and figures, that have been taken either verbatim or analogously from other works have,in each case, been properly cited for their origin and authorship.

This paper, either in parts or in its entirety, be it in the same or similar form, has not been submitted to any other examination board and has not been published.

Location, Date

Cologne, 02/02/2026

Signature

Abhishek Verma