Climbing the Influence Tiers on TikTok: A Multimodal Study

Pier Paolo Tricomi[1*], **Saurabh Kumar**[2*], **Mauro Conti**[1], **V.S. Subrahmanian**[2]

[1]University of Padua, Padua, Italy
[2]Northwestern University, USA
pierpaolo.tricomi@phd.unipd.it, kumar.saurabh@northwestern.edu, mauro.conti@unipd.it, vss@northwestern.edu

Abstract

Corporate social media analysts break influencers into five tiers of increasing importance: Nano, Micro, Mid, Macro, and Mega. We perform a comprehensive study of TikTok influencers with two goals: (i) what factors distinguish influencers in each of these tiers from the adjacent tier(s)? (ii) of the features influencers can directly control ("actionable" features), which ones are most impactful to reach the next tier? We build and release a novel TikTok dataset featuring over 230K videos from 5000 influencers—1000 from each tier. The dataset includes video details such as likes, facial action units, emotions, and music information derived from Spotify. Access to the videos is facilitated through provided URLs and hydration code. To find the most important features that distinguish influencers in a tier from those in the next tier up, we thoroughly analyze traditional features (e.g., profile information) and text, audio, and video features using statistical methods and ablation testing. Our classifiers achieve F1-scores over 80%. The most impactful actionable features are traditional and video features, including enhancing video pleasure, quality, and emphasizing facial expressions. Finally, we collect and release a YouTube Shorts dataset to conduct a comparative analysis, aiming to identify similarities and differences between the two platforms.

Introduction

Unlike most social media platforms, TikTok pays $0.02-0.04 for every 1000 views of a post.[1] This provides a huge incentive to create videos that garner lots of views. While academia abounds with "models of influence" (Kempe, Kleinberg, and Tardos 2003), industry and influencers on TikTok pay more attention to "tiers" of influence which are based on follower counts. As with Instagram (Tricomi et al. 2023) and YouTube (Rohit Shewale 2024), TikTok creators are divided into Nano influencers who have [1K, 10K) followers, Micro [10K, 50K), Mid [50K, 500K), Macro [500K,1M), and Mega [1M,+∞] influencers (Naseer et al. 2022; Tian, Dew, and Iyengar 2023). In addition to payments, influencers often get marketing opportunities from companies (Haenlein et al. 2020). A partnership with Charlie D'Amelio, one of the most followed people on TikTok, is estimated to cost more than $100,000 per post (Amal Moursi 2022). On the other hand, Nano and Micro-influencers tend to have a more homogeneous follower base (Haenlein et al. 2020). There is evidence that influencers differ in behavior and engagement depending on their tier (Jiang, Jin, and Deng 2022; Paksoy et al. 2023; Naseer et al. 2022), but it is unclear to what extent.

Thus, influencers have a huge incentive to "move up" from their current tier to the next tier. We study the factors that distinguish TikTok influencers in one tier from those in the next one up. Our findings may help influencers increase their reach and revenues.

To achieve this, we created the TikTok Influencer Dataset for Exploratory Study (or TIDES) containing data on 5,000 influencers, 230,406 videos, and 10,294 audio clips, by combining data from TikTok and Spotify.[2] To our knowledge, TIDES is the first publicly available dataset that analyzes TikTok influencers. Next, we developed a multimodal set of features associated with each of our 5,000 influencers. We trained classifiers to predict, for each tier t other than Mega, whether an influencer would belong to tier t or tier $(t + 1)$. We also performed ablation studies to find: (i) the most important (of all) features that separate influencers in tier t from those in $(t + 1)$, and (ii) the most important "actionable" features, i.e. features whose values can be directly modified by an influencer in order to reach the next tier.

We distinguished Nano from Micro influencers with an F1-score of over 0.84, Micro from Mid with an F1-score of over 0.83, Mid from Macro with an F1-score over 0.80 and Macro from Mega with an F1-score over 0.82. This suggests that influencers across adjacent tiers are relatively easy to separate. Interestingly, using only the top 50 most relevant features only drops these F1-scores by 1-2%.

As influencers cannot directly change features like the number of followers they have and the number of likes/views their posts receive, we introduce *actionable* features, e.g., they can choose music with a higher danceability score (ac-

---

[1]https://influencermarketinghub.com/how-much-does-tiktok-pay/

---

[2]The TIDES dataset and associated code are available at https://github.com/skmtr1/TIDES-TikTok-Influencer-dataset. Access to the videos is facilitated through provided URLs and hydration code (videos are not directly provided).

cording to Spotify[3]) or look happier in their videos (according to a video emotion classifier (Serengil and Ozpinar 2021)). These are *actionable* features. We addressed the following research questions:

RQ1: Do multimodal features (*audio*, *video*, *text*) differ significantly from *traditional* features (e.g., likes, profile information) when separating users in tier t from $(t + 1)$? Our ablation study looked at each of the four classification problems. We saw a 3% improvement in F1 for Nano-Micro, 1.6% for Micro-Mid, 0.4% for Macro-Mega, but no improvement for Mid-Macro.

RQ2: When considering only *actionable* features (*traditional*, *audio*, *video*, *text*), which ones make the most difference in separating tier t from $(t + 1)$? Our ablation study showed that *video* features are most significant (5.6% and 4.2%) in F1 for Nano-Micro and Mid-Macro, respectively, while *traditional* features are (3.4% and 6.7%) for Micro-Mid and Macro-Mega respectively. *Text* features only have an impact (1.3%) in the Macro-Mega case, while *audio* features only have a 1-2% impact in all cases except for Macro-Mega. This suggests that *video* and *traditional* features are the most important actionable features. *Audio* features are important. *Text* features are the least important.

RQ3: Which *actionable* features that an influencer can change are most important for them to move up to the next tier? Do these actionable features vary by tier? We found that influencers should primarily focus on *traditional* features (e.g., publishing videos regularly and frequently) and *video* features (e.g., producing more pleasant and high-quality videos). How influencers should change their behavior is linked to their current tier and the next one. For Macro influencers to become Mega influencers, having a verified profile is roughly eight times more important than increasing their total number of videos.

Finally, we created and released a new YouTube Shorts dataset containing ∼30K videos from 2,500 influencers. We applied our methodology to this dataset and found both similarities and differences between the two platforms. Unlike TikTok, views are far more critical than likes. However, facial action units are relevant on both platforms. This dataset is also available on our repository.

Related Work

We examine both general and influence studies of TikTok.

General TikTok Studies. TikTok use rose dramatically during the pandemic (Feldkamp 2021) leading to several studies of information spread on TikTok during that time (Li et al. 2021; Ostrovsky and Chen 2020; Southwick et al. 2021). Researchers studying virality found that close-up or medium-shot videos and videos containing text have a higher chance of going viral (Ling et al. 2022). Additionally, high user engagement metrics (e.g., likes, comments, shares) are pivotal in propelling videos to popularity. Surprisingly, joining a "trending" hashtag bandwagon seems less relevant (Klug et al. 2021). Further efforts study education (Fiallos, Fiallos, and Figueroa 2021;

Khlaif and Salha 2021), politics (Vijay and Gekker 2021), privacy (Neyaz et al. 2020), cyberbullying (Anderson 2020), and hate speech (Weimann and Masri 2023).

TikTok Influence. Influence maximization in social networks (Kempe, Kleinberg, and Tardos 2003; Banerjee, Jenamani, and Pratihar 2020) has been studied extensively. Custom models on how influencers should behave have been developed for Twitter (Lahuerta-Otero and Cordero-Gutiérrez 2016), Facebook (Arora et al. 2019; Hughes, Swaminathan, and Brooks 2019), YouTube (Sokolova and Kefi 2020), and Instagram (Tafesse and Wood 2021; Casaló, Flavián, and Ibáñez-Sánchez 2021), but not for TikTok. As a platform focused on fun, TikTok influencers' sense of humor, entertainment, and happiness increase the spread of messages (Barta et al. 2023; Yang and Ha 2021). Furthermore, TikTok influencers must continuously communicate with their audience and foster parasocial relationships to gain more followers. (Yang, Zhang, and Zhang 2021) developed an influence marketing algorithm to predict the increase in sales due to sponsored videos, and showed that disclosure of sponsorships (usually concealed) does not affect brand results.

Our New TIDES Dataset

We created a new TikTok Influencer Dataset for Exploratory Study (TIDES), a central contribution of our paper that is available in our repository. To the best of our knowledge, TIDES is the first publicly available dataset on TikTok influencers. We collected information and videos from 5,000 TikTok Influencers, 1,000 for each tier (Nano, Micro, Mid, Macro, Mega) using lists of influencers from HypeAuditor[4] and StarNgage[5], two famous influencer analytics sites. To limit statistical bias in our experiments, we first collected all the influencers listed on these websites, and then randomly sampled 1,000 from each tier. As websites may have out-of-date information, we verified the follower counts of each profile by cross-checking them with TikTok's data, correctly categorizing each influencer with their tier. To produce more generalizable findings, we included 1,000 influencers from each of the five major continents: America, Asia, Africa, Europe, and Oceania,[6] without restricting to categories based on age, race, or gender. In the future, influencers' domains or geographical locations could be taken into account to produce more specific results. After retrieving influencer lists, we used a TikTok scraper on Apify[7] to profile details (e.g., the number of videos) and the information of their (up to) 50 most recent videos, including the URLs. We used these URLs to download videos without the TikTok watermark[8] through a Python web scraper we wrote using Selenium library.[9] We collected a total of 230,406 videos in June 2023.

---

[3]https://developer.spotify.com/documentation/web-api/reference/get-audio-features

[4]https://hypeauditor.com/top-tiktok/

[5]https://starngage.com/plus/en-us/influencer/ranking/tiktok

[6]The actual distribution may differ, but we have no access to official data

[7]https://apify.com/clockworks/tiktok-scraper

[8]The TikTok logo is added to every video by the platform, potentially influencing the analyses.

[9]https://selenium-python.readthedocs.io/

Thus, we obtained two datasets: the Influencer dataset $\mathcal{I}$ ($|\mathcal{I}| = 5,000$) and the Video dataset $\mathcal{V}$ ($|\mathcal{V}| = 230,406$).

Feature Extraction

Let $\mathcal{I}$ denote the set of profile-related features for each influencer, and $\mathcal{V}$ be the set of meta-features per video (e.g., the number of likes, the music used). We augmented $\mathcal{I}$ with behavioral features and $\mathcal{V}$ by extracting content-related features from audio, video, and text (the caption) via several Deep Learning models. In total, we have 73 features, of which 68 are actionable. A complete list is available in Table 1. We now provide an overview of each set of features.[10] More discussions and evaluations of model performance are reported in the repository.

Traditional Features

Traditional features originate from the platform and do not include content-specific information. $\mathcal{I}$ contains data derived from the Influencer profile, e.g., follower and following counts, total number of likes and videos, geographical location, whether the user has a bio or a URL. We augmented $\mathcal{I}$ by calculating behavioral features, such as the distribution of videos by day of the week, videos published per day, and the inter-posting time (i.e., average time and standard deviation between each video). The video dataset $\mathcal{V}$ includes, for each video, whether it is sponsored, and engagement metrics such as the number of likes, views, shares, and comments.

Audio Features

Audio is a fundamental part of TikTok videos. Many videos use trending music, dances, lip-syncing, or interactions where the influencer communicates with their followers. We extracted features to understand how influencers use audio in their videos, using two channels: Spotify and the raw Audio Channel.

Spotify Features. The audio can either be original or from another artist. In the latter case, we used Spotify to extract additional features. Using the Spotify API's `search` endpoint[11], we first checked if the track existed on Spotify. If yes, we got the Spotify handle (id) and called `tracks/id` API to retrieve general track information like the popularity (in Spotify) and whether it is explicit. We then invoked the `audio-features/id` API to obtain audio features like danceability, speechiness, or instrumentalness, which are calculated by proprietary Spotify algorithms. These features serve as effective descriptors of the music, allowing us to gain deeper insights into the types of videos influencers share.[12] In total, we collected information from 10,294 tracks, often shared among the collected videos.

Emotion Features. We wondered whether emotions (e.g., angry, happy, sad, neutral, VAD emotional state (Russell and Mehrabian 1977) attributes such as valence, arousal, and dominance) were important in distinguishing between

influencer tiers. We extracted basic emotions by implementing a fine-tuned wav2vec2 model (speechbrain 2021) using SpeechBrain (Ravanelli et al. 2021). We also implemented the VAD model (Wagner et al. 2023).

Video Features

We extracted video features at the image level and then aggregated the results. As in past work (Shang et al. 2021), we extracted two frames per second of video, obtaining a total of 13,929,447 frames. We extracted the percentages of red, green, and blue channels. Next, using the HSV (Hue, Saturation, Value) representation, we calculate the percentages of luminance, warm and cool colors (Mehrabian and Russell 1974), as well as pleasure, arousal, and dominance scores (Valdez and Mehrabian 1994). We used several state-of-the-art deep learning models to analyze human subjects. Starting by detecting people's faces through Retinanet (Deng et al. 2020), we extracted Facial Action Units (Ekman and Friesen 1978) using Py-Feat (Cheong et al. 2023), Age and Gender through Dlib (Mowshon 2022), and race and emotion through DeepFace (Serengil and Ozpinar 2021). For each video, we extracted its definition, width, height, duration, and if it contains text (e.g., subtitles). This last feature was proven useful in recent studies (Ling et al. 2022): we extracted it via Tesseract.[13]

Text Features

These features are derived from the video caption. We extracted the caption length, the number of Emojis, and their sentiment (Kralj Novak et al. 2015) using the Emosent library[14] which returns a sentiment score ranging from -1 (negative) to $+1$ (positive), and three values representing the probabilities of being positive, negative, and neutral. We also extracted the number of hashtags and mentions.

Descriptive Statistical Exploration of Features

Based on the set of features discussed in the preceding section (see Table 1), we came up with several interesting hypotheses and tested them through Unpaired Student's t-test, a widely used statistical method to determine if there is a significant difference between the means of two groups. If the p-value is below a chosen significance level (commonly 0.05), the test reveals that there is a significant difference between the means of the two groups. In this section, we discuss these hypotheses, focusing first on the non-actionable features and then the actionable ones.

Non-Actionable Features

We wondered whether engagement by other users is a proxy for the tier to which an influencer belongs. We examined four hypotheses that model this general intuition.

Hypothesis 1 *We hypothesize that the total number of likes that a user gets is linked to the tier to which the influencer belongs.*

---

[10]The complete set of features is available in our repository.

[11]https://developer.spotify.com/

[12]Spotify features are fully explained at https://developer.spotify.com/documentation/web-api/reference/get-audio-features.

[13]https://tesseract-ocr.github.io/

[14]https://github.com/omkar-foss/emosent-py

Categories and Features

TR-has_bioLink*	SP.tempo*	A-arousal*	V-age*	V-RGB_percent_cold*
TR-has_signature*	SP-acousticness*	A-dominance*	V-race*	V-RGB_pleasure*
TR-commerceUser*	SP.danceability*	A-valence*	V-surprise*	V-RGB_arousal*
TR-likes	SP.energy*	A-neutral*	V-sad*	V-RGB_dominance*
TR-comments	SP.instrumentalness*	A-happy*	V-fear*	V-hasText*
TR-views	SP.liveness*	A-sad*	V-neutral*	T-n_hashtags*
TR-shares	SP.speechiness*	A-angry*	V-angry*	T-n_mentions*
TR-total_likes	SP.valence*	V-height*	V-happy*	T-CaptionLength*
TR-total_videos*	SP.mode_0*	V-width*	V-disgust*	T-EmojisLength*
TR-continent*	SP.explicit*	V-heightXwidth*	V-[n]_faces*	T-EmoNeutral*
TR-weekdays*	SP.time_signature*	V-duration*	V-RGB_r*	T-EmoPositive*
TR-videos_per_day*	SP.key*	V-definition*	V-RGB_g*	T-EmoNegative*
TR-interposttime*	SP.popularity*	V-n_effects*	V-RGB_b*	T-EmoSentScore*
TR-following*	SP.loudness*	V-FAUs (video_AU[XX])*	V-RGB_luminance*	
TR-videos_liked*	A-is_original*	V-gender*	V-RGB_percent_warm*	

Table 1: List of features expressed as category-feature(s). Actionable features are highlighted with a "*" symbol. The categories are TR = Traditional, SP = Spotify, A = Audio, V = Video, T = Text.

(a) Total number of likes.

(b) Average number of likes.

(c) Average number of comments.
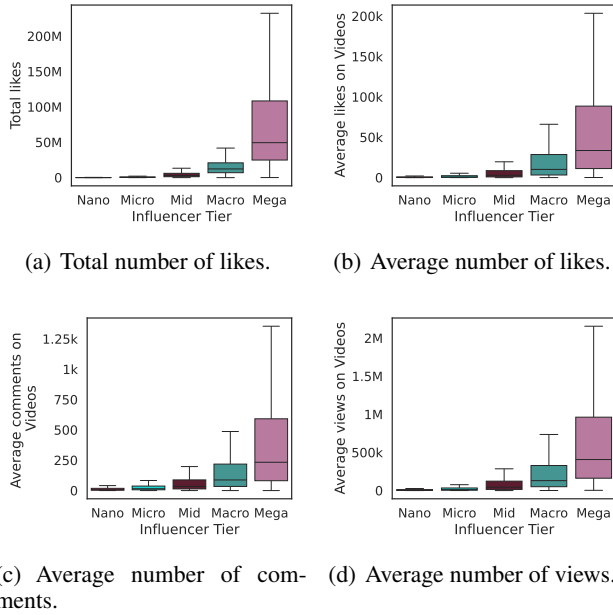
(d) Average number of views.

Figure 1: Box plot of the non-actionable features per tier.

Figure 1(a) shows a box plot whose x-axis represents the tiers and y-axis shows the total number of likes an influencer got from all her videos. We see that influencers at high tier have more likes. For each adjacent pair of tiers, the hypothesis that influencers at the higher tier get more likes is statistically valid via a t-test with $p < 10^{-36}$. One reason for this could be that influencers in higher tiers post more videos than those at lower tiers. To check this, we explored a derivative hypothesis.

Hypothesis 2 *We hypothesize that the average number of likes per video that a user gets is linked to the tier to which the influencer belongs.*

Figure 1(b) shows a box plot whose y-axis depicts the

mean number of likes garnered by each video posted by an influencer, while the x-axis reflects the influencer's tier. Again, influencers in higher tiers tend to receive a greater average number of likes per video they post. The big discrepancy in the number of followers between higher tiers and the lower tiers supports this observation. Hence, it is reasonable to conclude that higher tiers exhibit a greater mean number of likes. For each adjacent pair of tiers, the hypothesis that influencers at the higher tier get more likes per video is statistically valid via a t-test with $p < 10^{-5}$.

Hypothesis 3 *We hypothesize that the total number of comments that a user gets is linked to the tier to which the influencer belongs.*

Figure 1(c) shows a box plot showing the average number of comments per user by tier. The hypothesis that users in higher tiers receive more comments is true for all pairs of tiers with $p < 10^{-7}$.

Hypothesis 4 *We hypothesize that the total number of views a user gets is linked to the tier to which the influencer belongs.*

Figure 1(d) shows a box-plot of the average number of views per user by tier. The reader can readily see that users in higher tiers receive more views — this is validated by a t-test with $p < 10^{-17}$

Simply put, Figures 1(a), 1(b), 1(c) and 1(d) show that greater engagement by the audience (other users) is linked to the tier to which an influencer belongs. This is not very surprising.

Actionable Features

The preceding subsection looks at features that an influencer cannot directly influence because she cannot (at least honestly) directly increase the number of likes her videos get, the number of comments made on her videos, and so forth.[15]

---

[15]In the real world, influencers might do unethical things (e.g., create fake accounts to artificially inflate the number of views,

In this section, we look at *actionable* features — these represent features that an influencer can directly change through his actions.

Hypothesis 5 *We hypothesize that the total number of videos posted can directly influence the tier to which the influencer belongs.*
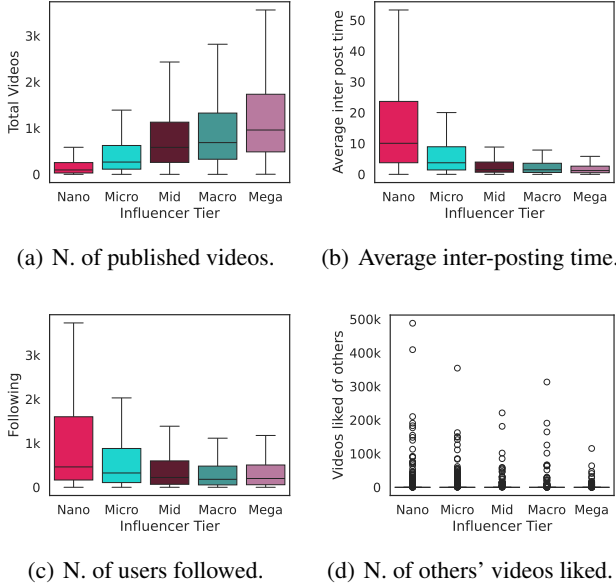

(a) N. of published videos.   (b) Average inter-posting time.

(c) N. of users followed.   (d) N. of others' videos liked.

Figure 2: Box plot of the *actionable* features per tier.

Figure 2(a) shows a box plot whose x-axis represents the tiers and whose y-axis shows the number of videos posted by the influencers in each tier. Higher-tier influencers post more videos than lower-tier influencers. One possible reason could be that higher-tier influencers have been in the network for a long time and, hence, have more videos. This finding holds for all pairs of adjacent tiers with $p < 0.02$.

Hypothesis 6 *We hypothesize that the average inter-posting time between videos posted can directly influence the tier to which the influencer belongs.*

Figure 2(b) shows a box plot whose x-axis shows the tiers and whose y-axis shows the average inter-post time between videos posted by influencers in that tier. A Student t-test confirmed statistical significance for all pairs of adjacent tiers ($p < 10^{-4}$) except for Mid-Macro. As reported in the related hypothesis, this feature is negatively correlated with the 'total videos' (Pearson correlation: -0.235). The same can be observed by seeing Figure 2(a) and Figure 2(b).

Hypothesis 7 *We hypothesize that the total number of users that an influencer follows is linked to the tier to which the influencer belongs.*

Figure 2(c) shows the number of users that an influencer follows. The x-axis shows the tiers and the y-axis denotes the

likes, and comments on their videos). In this paper, we assume that influencers are honest and do not take such unethical actions.
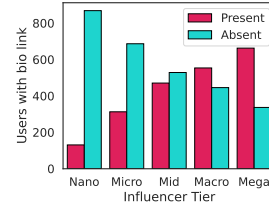

Figure 3: Presence of bio link info.

number of users that an influencer follows. The figure shows that lower-tier influencers follow more users than those in higher tiers. This suggests that higher-tier influencers may cultivate an aura of exclusivity by limiting the number of accounts they follow. This hypothesis is statistically significant for Nano-Micro, Micro-Mid, and Mid-Macro pairs with $p < 0.003$.

Hypothesis 8 *We hypothesize that the total number of others' videos that an influencer likes is linked to the influencer's tier.*

Figure 2(d) shows a box plot whose x-axis shows the tiers and y-axis represents the number of videos by other users that an influencer has liked. The results are inconclusive. This hypothesis is valid only for Macro-Mega with $p < 0.04$.

Hypothesis 9 *We hypothesize that influencers who post their bios publicly influence are more likely to belong to higher tiers.*

Figure 3 shows a bar chart whose x-axis shows tiers and whose y-axis shows the number of influencers that have (or do not have) a public bio link in their profile for each tier. The figure shows that influencers in higher tiers are more likely to have publicly posted bios. This finding is verified for all pairs of tiers with $p < 0.0003$.

Hypothesis 10 *We hypothesize that influencers tend to post videos associated with their respective tiers on specific days of the week.*

Figure 4 illustrates the mean proportion of videos an influencer uploads on weekdays. During weekends, influencers from the Nano tier exhibit a higher posting frequency than influencers from other tiers. Conversely, on workdays (not weekend), influencers across all tiers maintain a similar frequency of video posting. The Student t-test shows that this finding is valid for some pairs of tiers on specific days, and for others, it is not. For example, on Saturday and Sunday, it is valid for Nano-Micro and Micro-Mid ($p < 0.003$) but not for others.

Influencer Tier Classification

To answer our research questions, we trained multiple ML classifiers on pairs of consecutive tiers. We created binary classifiers (rather than a multiclass classifier) for two reasons: i) except for influencers migrating from other platforms, it is very unlikely that an influencer goes from a tier to a non-adjacent one (e.g., from nano to mid), given that several thousands of followers separate them, and ii) related to
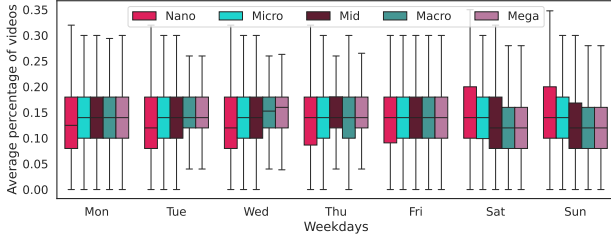
Figure 4: Box-plot of the average percentage of videos an influencer posts per tier per week of the day.

the previous point, binary classifiers make it easier to understand the important features separating two adjacent tiers. We started by aggregating video data for each influencer. We rigorously trained and tested several classifiers to identify the best one. We also conducted a set of ablation studies using the optimal classifier (for each of the four classification problems, i.e., Nano vs. Micro, Micro vs. Mid, Mid vs. Macro, Macro vs. Mega) to provide conclusive answers to our research questions. We remind the reader that the primary objective of this paper is to identify the most valuable (*actionable*) features for advancing through influencer tiers. Therefore, we are building high-quality classifiers but not striving for perfect predictions, which could be a future research direction.

Data Aggregation

We conducted our experiments at the user level. Each influencer corresponds to a row in our dataset. For each influencer, we aggregated all the information about their videos into a single entry. We used different aggregation types depending on the variable type (e.g., boolean, categorical, float). For boolean variables (e.g., if the audio is original), we computed the percentage of true values, i.e., the percentage of videos posted by the user that used original audio. Likewise, for categorical variables, we calculated their respective percentages (e.g., for the posting day, we calculated the percentages of videos published on Monday, Tuesday, etc.). When dealing with numeric values, we computed several statistics, including the mean, standard deviation, min, and max. We also used distribution features associated with a feature f: we found the minimum and maximum values of f across all influencers and divided the $[min, max]$ range into ten bins. We then computed the probability of a user's video having a value within each of the ten bins. These ten bins were computed only on training data to prevent data leakage. Two additional bins were used to accommodate values below the minimum and above the maximum. For likes, comments, views, and shares, which exhibit vast value ranges, we opted for a six-bin approach (instead of the 10-equidistant bins) inspired by the box plots shown earlier in the paper (excluding outliers). The boundaries were $-\infty$, minimum, first quartile, median, third quartile, maximum, $+\infty$. As our *video* features are extracted from frames, we aggregated video values for each influencer by considering all frames from all their videos. For instance, to calculate the video pleasure feature of an influencer i, we extracted all
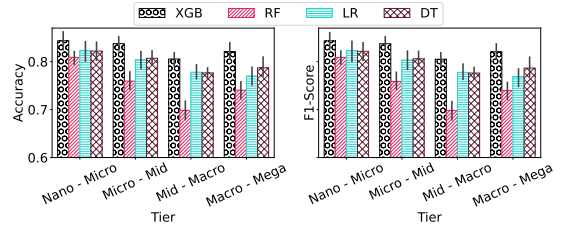

Figure 5: Classification results to predict the influencer tier on TikTok using four best classifiers.

frames of the videos made by i, calculated the pleasure value on all these frames, and aggregated the results using the ten bins approach, defined through the minimum and maximum video pleasure of all the other influencers in the training set. Last, we calculated minimum, maximum, mean, and standard deviation video pleasure values related to the influencer i only. In total, we had 869 features per influencer.

All Features Analysis

To differentiate between consecutive tiers of influencers, we trained eight machine learning classifiers: XGBoost (XGB), Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), Naive Bayes (NB), Support Vector Machine (SVM), Deep Neural Network (DNN), and K-Nearest Neighbors (KNN). We validated our results using nested 10-fold cross-validation (CV). We found the best hyperparameters via grid-search (validation values are reported in our repository).

Figure 5 shows the results of the classification using the best four classifiers (for results of all classifiers, see our repository). XGBoost outperformed all the other classifiers, reaching an over 80% F1-score in every tier. In particular, the easiest task is to differentiate between the Nano and Micro tiers (84.3% F1-Score), while the most challenging is Mid vs. Macro (80.5% F1-Score).

Given its superior performance, we continued our experiments using XGBoost. We repeated the experiments by selecting the 10, 20, and 50 most important features at training time (out of 869), using Anova F-values[16] (Table 2). Interestingly, using 10, 20, or 50 features only drops the F1-Score by 1-3%, except in the case of 10 features for Macro-Mega, where we have a drop of 11%. This suggests that unlike the other tiers, substantially more than 10 factors distinguish Macro and Mega influencers.

N Feat.	N-MC	MC-MD	MD-MA	MA-ME
10	83.7±2.3	80.4±4.0	78.1±2.5	70.9±3.7
20	83.6±2.5	80.0±3.4	78.1±1.8	79.0±2.0
50	83.1±2.9	80.7±2.8	79.2±2.5	79.3±2.6
All	84.3±3.1	83.7±2.5	80.5±2.4	82.1±2.9

Table 2: F1-Score of XGB Classifier using top N Features. N=Nano, MC=Micro, MD=Mid, MA=Macro, ME=MegA-

---

[16]https://scikit-learn.org/stable/modules/generated/sklearn. feature_selection.f_classif.html

To inspect these factors, we combined and explored feature importance from our 3 best classifiers, XGBoost, Logistic Regression, and Decision Trees, improving the stability of the results. Given our 10-fold CV approach, we have 10 best models for each classifier result. We combined the feature importance of these 10 models to calculate the final feature importance for each classifier (e.g., XGB). We grouped the features based on their typology (e.g., video dimension and definition are under video quality), and ranked the groups based on the value of their most important feature. This allows us to maintain the original feature-importance ranking, giving a general idea of the type of features, and preventing some groups from being disadvantaged by considering many irrelevant aggregation features.[17] Obviously, a graph containing all the features would be more accurate — space constrain us from doing that. Still, our representation allows the reader to understand how features' importance changes within and across the tiers. The following figures report the weighted feature importances across our best three classifiers (XGB, LR, DT). For each feature f and classifier j, we calculated the rank $R_j(f)$:

$$R_j(f) \quad b = \quad 1 - \frac{Rank_j(f) - 1}{N - 1},$$

where 1 is the maximum importance and 0 the least. The starting rank $Rank_j(f)$ is calculated by creating a descending order list by feature importance returned by Sklearn[18]. For DT, we used Gini Importance. For XGB, we used Gain information and for LR, we used the absolute values of the coefficients associated with each feature.[19] We then calculated the feature importance of f as the weighted average rank of the three classifiers:

$$Imp(f) \quad = \quad \frac{\sum_{j=1}^{3} Rank_j(f) * c_j}{\sum_{j=1}^{3} c_j}$$

where c_j is the F1-score for classifier j. Hence, the more accurate the classifier, the more impact its feature importance will have on the final result. In the graphs, a value of 1 means that (group of) feature was in the first rank for all three classifiers, while 0 indicates the feature was always the least important.

Figure 6 shows the feature importances in each tier. We see that likes, and views help distinguish influencers between each pair of tiers, as also verified by Hypotheses 2 and 4. Video quality (e.g., size, definition) helps in low tiers, since when the tier increases, the quality has already reached high standards. In contrast, profile information helps more in distinguishing influencers in high tiers. Indeed, most Mega influencers have a verified profile with a bio link.

These results show that *traditional* features such as likes and views play a prominent role in distinguishing influencer tiers. Returning to our first research question (**RQ1**): "Do multimodal features (*audio*, *video*, *text*) make any difference
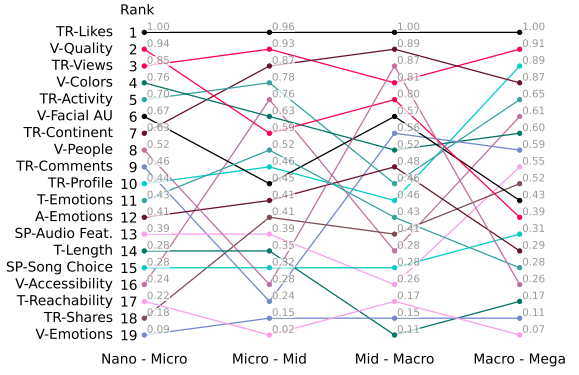
Figure 6: Feature importance per Tier - All features.

when separating users in one tier from the next higher tier?". To answer this question, we performed an ablation study. We divided our features into *traditional*, *video*, *audio*, and *text* sets, and performed the classification by removing, in turn, each possible combination of feature sets. Table 3 reports the mean increase in F1-score when that combination of feature sets is added on top of the remaining sets. Therefore, the row "video + audio + text" answers our RQ1. When these features are used, our classifier improves by 2.95% F1-Score for Nano-Micro problem, 1.6% for Micro-Mid, and 0.4% for Macro-Mega, but loses 0.1% F1-Score for Mid-Macro. Therefore, it is clear that *traditional* features, in particular likes, views, or the influencer activity, are the most useful during classification, while the others are less important, except in the case of Nano-Micro.

Features	N-MC	MC-MD	MD-MA	MA-ME
A	0.4±1.8	0.8±2.1	−0.5±1.3	0.5±1.6
A+Tr	10.4±2.7	18.8±4.1	22.1±5.3	25.6±3.5
A+T	1.2±2.0	0.8±2.8	−1.0±2.5	0.4±0.9
A+T+Tr	12.6±4.1	20.3±4.7	21.5±4.3	24.3±4.5
Tr	10.4±2.7	18.5±4.0	22.8±5.4	23.6±2.5
T	0.7±2.4	0.2±1.8	−0.2±1.5	−0.1±1.9
T+Tr	11.5±3.4	19.5±4.2	22.8±5.9	25.1±3.3
V	2.0±2.6	0.6±1.6	1.0±1.5	1.2±2.7
V+A	2.8±2.4	1.2±2.8	1.0±1.7	1.0±1.3
V+A+Tr	20.8±4.5	26.9±3.4	26.3±2.4	27.2±3.7
V+A+T	3.0±3.0	1.6±3.4	−0.1±1.1	0.4±1.4
V+Tr	20.3±4.8	24.6±4.5	25.9±4.2	27.0±3.5
V+T	3.0±2.4	1.3±1.7	0.8±1.4	1.1±2.2
V+Tr+T	27.0±5.2	27.5±5.0	26.9±3.5	30.9±4.6

Table 3: Impact, in percentage, of feature groups when considering all features - Increase of F-1 Score.

Answer RQ1: When considering all features, multimodal features (*audio*, *video*, *text*) only play a minor role in distinguishing users from one tier to the next, with the highest benefit between Nano-Micro. Indeed, *traditional* features such as likes, views, and user activity play a leading role.

Actionable Feature Analysis

Though *traditional* features are more critical than multimodal features when distinguishing between users in consecutive tiers, a user cannot directly modify these features — for instance, she cannot take actions that directly increase the number of likes or views of her videos. Table 1 has previously defined the set of *actionable* features, i.e., those the user can directly control to improve her status. So we repeated our ablation study with only *actionable* features to learn how an influencer can climb to the next tier.

Table 4 reports the results of the new ablation study. Unlike before, multimodal features now bring improvements of 7.1% for Nano-Micro, 2.1% for Micro-Mid, 3.5% for Mid-Macro, and 1.1% for Macro-Mega. Interestingly, *video* features are now more critical than *traditional* features for Nano-Micro (5.6% vs 1.3%) and Mid-Macro (4.2% vs 2.8%), suggesting that in these tiers, influencers should focus more on the content of their videos. In the other tiers, *traditional* features are most impactful. *Audio* features impact around 1-2% in each category, while *text* features have the most negligible impact, often below 1%.

> **Answer RQ2:** When evaluating *actionable* multimodal features, *video* features emerge as the most influential, followed by *audio* features. *text* features, except for Macro-Mega, exhibit minimal relevance. Meanwhile, *traditional* features retain their significance across all tiers, although they take a backseat to *video* features in the Nano-Micro and Mid-Macro tiers.

Feature	N-MC	MC-MD	MD-MA	MA-ME
A	1.1±2.1	1.0±2.9	2.0±3.7	0.7±2.6
A+Tr	1.4±3.3	3.7±3.6	2.1±1.9	8.5±2.8
A+T	0.6±2.0	1.6±4.1	1.9±2.9	1.7±3.6
A+T+Tr	3.5±3.4	5.3±3.3	1.9±3.7	7.1±3.7
Tr	1.3±3.6	3.4±2.8	2.8±3.9	6.7±3.1
T	0.3±2.0	0.8±3.4	0.7±2.6	1.3±2.7
T+Tr	2.4±3.3	4.4±2.2	2.8±3.3	8.0±5.2
V	5.6±3.0	1.8±3.7	4.2±4.7	0.5±2.0
V+A	6.6±4.5	3.0±4.1	3.0±4.4	0.9±2.3
V+A+Tr	11.7±5.1	11.9±5.3	6.3±5.1	10.0±6.1
V+A+T	7.1±3.8	2.1±4.5	3.5±2.1	1.1±3.3
V+Tr	11.2±5.1	9.5±3.1	5.9±4.1	9.9±3.0
V+T	7.2±3.1	1.1±4.2	4.4±3.8	1.6±2.6
V+Tr+T	17.9±5.6	12.5±5.2	5.9±3.4	13.8±3.8

Table 4: Impact, in percentage, of feature groups when considering actionable features - Increase of F-1 Score.

Which Features to Improve?

Thus far, we have focused on high-level feature groups to quantify the significance of multimodal content. Yet, influencers are keen to gain insights into the specific adjustments required to ascend to the next tier, encompassing global strategies and granular details. To support this, we first inspect feature importance using all *actionable* features, similar to the preceding section. We then inspect the feature importance of the classifiers we used in the ablation study, which considered only a particular group of features (*traditional*, *audio*, *video*, *text*).

All Actionable Features Figure 7 shows the feature importance of all the *actionable* features. It immediately stands out that each classification considers different features. For example, emotions in the *text* are important for Nano-Micro classification, but not important in Micro-Mid, and again important for Mid-Macro and Macro-MegA- This is because, on average, Nano influencers use many positive emoticons, while Mega influencers use less. However, we still see some common patterns. For example, the video quality, profile, and Facial Action units features are in the top-5 groups for each category. Similarly, the video emotions are always in the last positions, suggesting that emotions are less important.
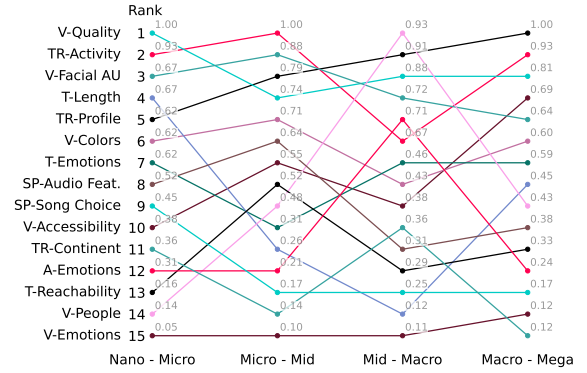

Figure 7: Feature importance - All actionable features.

Obviously, these features are important in different ways and with different weights. For example, the video quality is significantly more impacting than Facial AU for Nano-Micro. To understand how algorithms consider these features, we should examine how influencers behave in their tiers with respect to these features. For instance, from Figure 4, we know that Nano influencers post more on the weekend, while Macro and Mega post constantly throughout the week. However, the Nano-Micro and Macro-Mega algorithms trend in the opposite direction. We now explore each feature set.

Traditional Features Figure 8 shows the importance of *traditional* features. When we look at the top-5 *actionable* features across the four classification problems, we note that whether the user has a publicly available bio is in the top-5 in 3 out of 4 cases, and the total number of videos posted is always in the top-3. As tiers increase, influencers tend to set their bio link more frequently (Figure 3) and have many more videos (Figure 2(a)). When we look at the top-10 most important features, the time between posts and the number of videos posted daily also become significant. In fact, influencers should behave differently depending on their tiers. Figure 2(b) shows that the inter-posting time should decrease when the tier increases: the average inter-posting time is ten days for Nano influencers, but less than one day for Mega influencers. Interestingly, the significance of geographical continents varies across tiers, implying that influ-

encers might also tailor their behavior based on it. Furthermore, being verified is not important for Nano-Micro, but is extremely important for Macro-MegA- Indeed, only one of our profiles is verified in the nano category, contrary to 171 and 474 in the Macro and Mega categories, respectively.
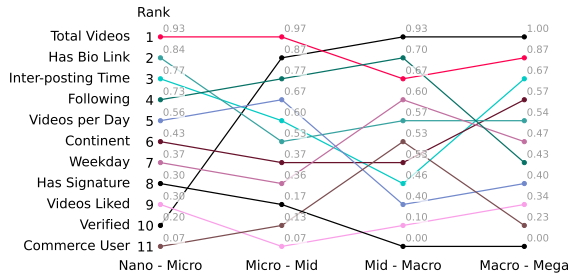

Figure 8: Feature importance - Traditional Actionable.

Audio Features Figure 9 shows the importance of *audio* features, where we see a mix of common and conflicting trends. For instance, Spotify loudness is crucial for Nano-Micro classification, but not for distinguishing between other tiers. Nano influencers use more Spotify tracks with a loudness value close to 0, contrary to the other tiers, which tend to avoid that on average. Similarly, Spotify acousticness is essential only in Macro-Mega classification, where Mega influencers tend to use more acoustic tracks in their videos. On the other hand, Spotify popularity and audio valence stay in the top-10 across all four classification problems, but are important in different ways. Further inspecting popularity, Nano and Macro tiers often use less popular songs, with Mega influencers avoiding them.
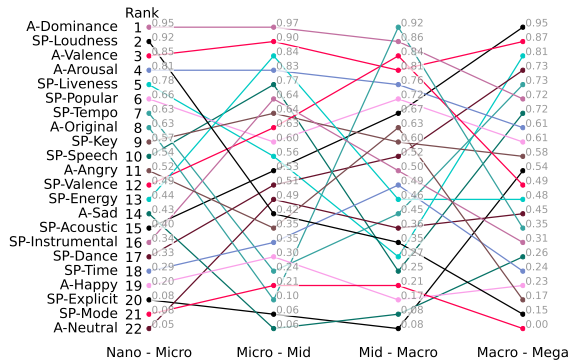

Figure 9: Feature importance - Audio Actionable.

Video Features Figure 10 shows the importance of *video* features, which are among the most important features for Nano-Micro classification, and also important across the other three classifications. We noticed that lower-tier influencers shot their videos using many different sizes, while high-tier influencers tend to use more uniform sizes. This might indicate that most high-tier influencers own the same devices (e.g., top-notch smartphones with outstanding cameras). Facial Action Units are also very important in all four

classification problems. For instance, Nano-Micro classifications show that Nano influencers exhibit more lip-related AUs, likely related to the higher presence of lip-syncing video in the Nano tier.[20] The video pleasure is also significant, which increases and stabilizes in the Mega tier. Neutral and angry emotions appear to be less significant. Text on videos (under video accessibility) becomes less necessary for higher tiers.
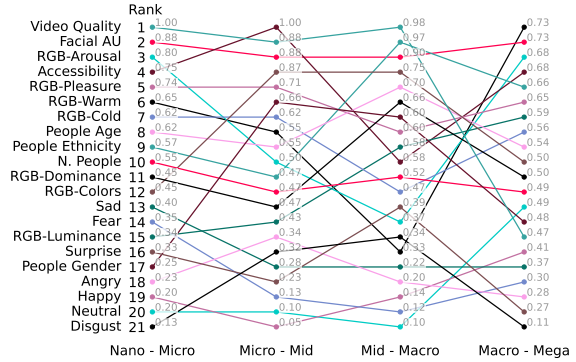

Figure 10: Feature importance - Video Actionable.

Actionable Text Features Figure 12 shows the importance of *text* features. Here, we witness more linear patterns. As stated above, *text* features do not have much impact. Still, the caption length is the feature that best separates Nano-Micro and Micro-Mid. Nano, has the shortest captions, Micro is in the middle, and Mid captions are the longest. Positive and neutral emotions are important in all classifications except Macro-Mega, in which negative ones take over. Mega influencers tend to use more negative emotions in their captions. Past work (Klug et al. 2021) suggests that hashtags and mentions are not important.

> **Answer RQ3:** Influencers should primarily focus on *traditional* and *video* features. Important features vary greatly within each feature type depending on the classification problem considered. How influencers should change their behavior is strictly related to their current tier and the next one up.

---

[20]An example of FAU exhibiting lip-related AU (e.g., AU 23, 24, 25) is visible in Figure 11.
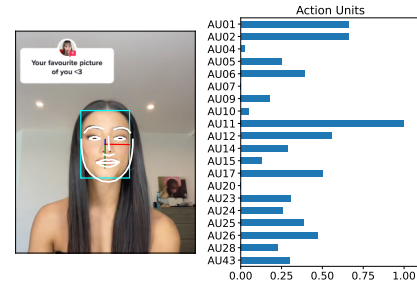

Figure 11: Example of Facial Action Units probabilities. The subject is the popular Mega Influencer @charlidamelio.
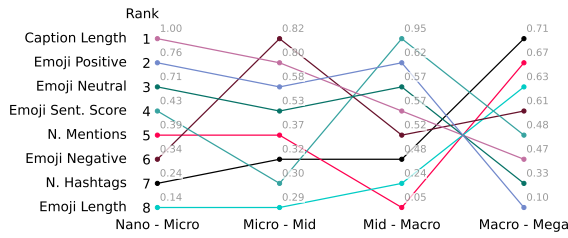
Figure 12: Feature importance - Text Actionable.

YouTube Shorts Comparison

We now study whether our TikTok findings apply to other social video-sharing platforms. We focused on YouTube, the most popular platform for video sharing, with more than 2.7 billion monthly active users (Rohit Shewale 2024). In 2021, YouTube introduced YouTube Shorts to compete with Tik-Tok, a feature that allows content creators to upload short videos (60 seconds max). As with TikTok, we collected such short videos for different tiers of influencers, and conducted a multi-modal analysis to understand the factors that separate an influencer in one tier from those in the next one up.

Data Collection

Just like TikTok, YouTube influencers are placed in tiers based on the number of subscribers (followers) (Lillie 2023). With a procedure similar to TikTok, we gathered information about 2,500 influencers (500 per tier) using HypeAuditor and StarNgage websites. We selected influencers randomly from the five major continents to have general results. Using Apify scrapers for shorts[21] and channel info,[22] we collected information about the latest shorts for each influencer (up to 20) and profile information. We used the video URLs and the Python library Pytube[23] to download a total of 29,318 videos in December 2023. Similar to our TikTok experiments, we extracted Traditional, Audio, Video, and Text features, keeping only those in common with TikTok for a fair comparison. As we could not access Spotify information on YouTube, we excluded these features. The YouTube dataset, containing video information along with URLs and code for hydration, is available in our repository.

Results

Like our TikTok experiments, we built binary classifiers to predict the tier of an influencer. Figure 13 reports the F1-score of the best classifiers. Predictive performance has F1-scores in the 70-85% range. Unlike TikTok, classifying Nano vs. Micro tier is more difficult than Macro vs. Mega, with the best F1-scores of 0.77 and 0.83, respectively.

Figure 14 shows the feature importance derived from the best three models (XGB, DT, LR). Compared to Figure 6, we immediately see that on YouTube, views are far more important than Likes (the top-1 feature on TikTok), and that comments play a crucial role (but are of limited importance in TikTok). Moreover, profile information is less important

[21]https://apify.com/newbs/youtube-shorts
[22]https://apify.com/streamers/youtube-channel-scraper
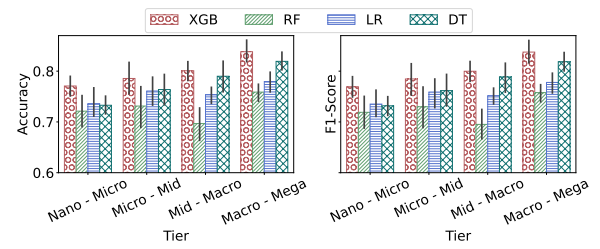[23]https://pytube.io/

Figure 13: Classification results to predict the influencer tier on YouTube using the four best classifiers.
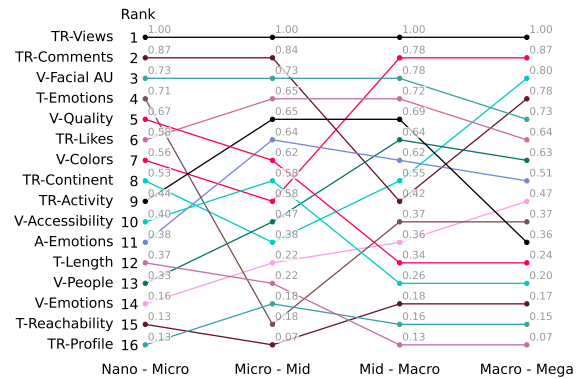

Figure 14: YouTube all feature importances.

than on TikTok, and Video Emotions play a more impactful role on YouTube, especially for Macro vs. Mega influencers. In terms of similarities between the two platforms, video quality is critical in distinguishing lower tiers, while for higher tiers it becomes irrelevant, since producing high-quality videos is a given. Facial Action Units are important in both platforms, with a prominent role in YouTube, while text reachability is not important on both platforms.

Discussion

We analyzed the features that influencers should focus on to climb to the next tier. Our findings can enable digital marketers to collaborate with influencers who follow these trends to create successful campaigns. Although our study is the first in this direction, others have studied the influence and virality of TikTok videos. Humor and authenticity play an important role in influencer marketing (Barta et al. 2023; Yang, Zhang, and Zhang 2021). Although automated humor detection is challenging (Kayatani et al. 2021), studying how different tiers of influencers use humor may help increase their influence. To boost video engagement, close-up and medium-shot scales are preferred, as well as a second point of view (Ling et al. 2022). As these features were manually extracted by (Ling et al. 2022), incorporating them automatically in the future may improve our analyses. The authors also found that using text on videos reduces virality, which is consistent with our results: higher tiers tend to use less text on videos. More hashtags usually bring more likes (Li et al. 2021), a feature that we found to be very important in separating Macro and Mega influencers. Recent studies also showed that joining a trend bandwagon contributes to

higher engagement (Su et al. 2020; Wang et al. 2023). Therefore, future influence studies could consider whether influencers in different tiers should join a trend bandwagon, launch new trends, or strive to generate unique content as suggested in (Chu, Deng, and Mundel 2022). Lastly, (Wu et al. 2022) showed how people of different ages, genders, and races can be more influential than others in marketing campaigns depending on the subject. We found those features to be of medium-high importance, especially in distinguishing between Mid and Macro influencers.

As the first large data-driven study of TikTok influencers, we considered influencers across diverse domains. However, as reported in past studies of TikTok (Wu et al. 2022) and Instagram (Tricomi et al. 2023), studying influencers by separating them by domain can be insightful. Moreover, as demonstrated in our comparison with YouTube Shorts, some important factors are common (e.g., a preference for higher video quality), while there are also differences (e.g., views are more important than likes on YouTube). For instance, high visual dominance is critical for Instagram Fashion influencers (Tricomi et al. 2023) but moderately for TikTok. Therefore, each social media platform should be studied separately since findings could be different. In this paper, we proposed a general methodology that could be ideally applied to any social media, as we did for TikTok and YouTube. Similarly, new TikTok features could be introduced in the future after a proper examination.

Ethical Issues. When attempting to climb the influence tiers, ethical considerations arise, especially regarding content authenticity and audience manipulation. This paper ethically analyzes the factors that might help influencers deliver better content for their audience with the ultimate goal of creating a more enjoyable platform for both influencers and users. Still, we emphasize that influencers should focus on transparency and authenticity in their content (Yang, Zhang, and Zhang 2021), and focus on genuine engagement rather than pursuing tactics solely to increase metrics. Manipulating engagement numbers or using deceptive practices may lead to short-term gains but can damage credibility in the long run. Similarly, influencers should prioritize authenticity over sheer quantity of content or followers. Building a genuine connection with the audience is more sustainable and ethical than resorting to shortcuts that may compromise the integrity of the influencer.

Conclusion and Future Works

We investigated the key factors contributing to influencers' tier progression. We built powerful tier prediction classifiers and explored features importance within the best models. While metrics such as likes and views may seem influential, they are not within influencers' control. Our experiments identified *actionable* features primarily within *traditional* features (e.g., posting frequency) and video-related attributes (e.g., video pleasure or facial expressions).

We encountered certain limitations that offer opportunities for future research. First, while we can identify features that influencers should enhance, we currently lack precise guidance on how much improvement is needed to maximize tier progression probabilities. Second, our current approach analyzes features individually, even though combining them often yields superior results. In future work, we will address these constraints, refine our feature sets, and explore trends on the platform, including how well influencers align with these trends. Last, a longitudinal examination, monitoring the progression of influencers and their content strategies over time, has the potential to yield further insights into the dynamics of advancing through the influence tiers.

References

Amal Moursi. 2022. The Most from TikTok Sponsored Posts? https://www.hopperhq.com/blog/2022-tiktok-rich-list/. Accessed: 2023-09-08.

Anderson, K. E. 2020. Getting acquainted with social networks and apps: it is time to talk about TikTok. *Library hi tech news*, 37(4): 7–12.

Arora, A.; Bansal, S.; Kandpal, C.; Aswani, R.; and Dwivedi, Y. 2019. Measuring social media influencer index-insights from Facebook, Twitter and Instagram. *Journal of retailing and consumer services*, 49: 86–101.

Bailey, M.; Dittrich, D.; Kenneally, E.; and Maughan, D. 2012. The menlo report. *IEEE Security & Privacy*, 10(2).

Banerjee, S.; Jenamani, M.; and Pratihar, D. K. 2020. A survey on influence maximization in a social network. *Knowledge and Information Systems*, 62: 3417–3455.

Barta, S.; Belanche, D.; Fernández, A.; and Flavián, M. 2023. Influencer marketing on TikTok: The effectiveness of humor and followers' hedonic experience. *Journal of Retailing and Consumer Services*, 70: 103149.

Casaló, L. V.; Flavián, C.; and Ibáñez-Sánchez, S. 2021. Be creative, my friend! Engaging users on Instagram by promoting positive emotions. *Journal of Business Research*, 130: 416–425.

Cheong, J. H.; Jolly, E.; Xie, T.; Byrne, S.; Kenney, M.; and Chang, L. J. 2023. Py-feat: Python facial expression analysis toolbox. *Affective Science*, 1–16.

Chu, S.-C.; Deng, T.; and Mundel, J. 2022. The impact of personalization on viral behavior intentions on TikTok: The role of perceived creativity, authenticity, and need for uniqueness . *Journal of Marketing Communications*, 1–20.

Deng, J.; Guo, J.; Ververas, E.; Kotsia, I.; and Zafeiriou, S. 2020. Retinaface: Single-shot multi-level face localisation in the wild. In *IEEE/CVF CVPR*, 5203–5212.

Ekman, P.; and Friesen, W. V. 1978. Facial action coding system. *Environmental Psychology & Nonverbal Behavior*.

Feldkamp, J. 2021. The rise of TikTok: The evolution of a social media platform during COVID-19. *Digital responses to Covid-19: Digital innovation, transformation, and entrepreneurship during pandemic outbreaks*, 73–85.

Fiallos, A.; Fiallos, C.; and Figueroa, S. 2021. Tiktok and education: Discovering knowledge through learning videos. In *Proceedings of IEEE ICEDEG '21')*, 172–176.

Fiesler, C.; Beard, N.; and Keegan, B. C. 2020. No robots, spiders, or scrapers: Legal and ethical regulation of data collection methods in social media terms of service. In *Proceedings of AAAI ICWSM*, volume 14, 187–196.

Haenlein, M.; Anadol, E.; Farnsworth, T.; Hugo, H.; Hunichen, J.; and Welte, D. 2020. Navigating the New Era of Influencer Marketing: How to be Successful on Instagram, TikTok, & Co. *California management review*, 63(1): 5–25.

Hughes, C.; Swaminathan, V.; and Brooks, G. 2019. Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns. *Journal of marketing*, 83(5): 78–96.

Jiang, Y.; Jin, X.; and Deng, Q. 2022. Short video uprising: how# BlackLivesMatter content on TikTok challenges the protest paradigm. *arXiv preprint arXiv:2206.09946*.

Kayatani, Y.; Yang, Z.; Otani, M.; Garcia, N.; Chu, C.; Nakashima, Y.; and Takemura, H. 2021. The laughing machine: predicting humor in video . In *Proceedings of the IEEE/CVF WACV*, 2073–2082.

Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD*, 137–146.

Khlaif, Z. N.; and Salha, S. 2021. Using TikTok in education: a form of micro-learning or nano-learning? *IJVLMS*, 12(3): 213–218.

Klug, D.; Qin, Y.; Evans, M.; and Kaufman, G. 2021. Trick and Please. A Mixed-Method Study On User Assumptions About the TikTok Algorithm. In *Proceedings of the 13th ACM WebSci '21*, 84–92.

Kralj Novak, P.; Smailović, J.; Sluban, B.; and Mozetič, I. 2015. Sentiment of Emojis. *PLOS ONE*, 10(12): 1–22.

Lahuerta-Otero, E.; and Cordero-Gutiérrez, R. 2016. Looking for the perfect tweet. The use of data mining techniques to find influencers on twitter. *Computers in Human Behavior*, 64: 575–583.

Li, Y.; Guan, M.; Hammond, P.; and Berrey, L. E. 2021. Communicating COVID-19 information on TikTok: a content analysis of TikTok videos from official accounts featured in the COVID-19 information hub. *Health Education Research*, 36(3): 261–271.

Lillie, L. 2023. How to Pay YouTube Influencers. https://www.lumanu.com/blog/how-to-pay-youtube-influencers. Accessed: 2024-01-10.

Ling, C.; Blackburn, J.; De Cristofaro, E.; and Stringhini, G. 2022. Slapping cats, bopping heads, and oreo shakes: Understanding indicators of virality in tiktok short videos. In *Proceedings of the 14th ACM WebSci '22*, 164–173.

Mehrabian, A.; and Russell, J. A. 1974. *An approach to environmental psychology*. the MIT Press.

Mowshon. 2022. Age and Gender. https://github.com/mowshon/age-and-gender. Accessed: 2023-07-10.

Naseer, S.; Hasan, S.; Bhuiyan, J.; and Prasad, A. 2022. Current public trends in the discussion of dry eyes: a cross-sectional analysis of popular content on TikTok. *Cureus*.

Neyaz, A.; Kumar, A.; Krishnan, S.; Placker, J.; and Liu, Q. 2020. Security, privacy and steganographic analysis of FaceApp and TikTok. *IJCSS*, 14(2): 38–59.

Ostrovsky, A. M.; and Chen, J. R. 2020. TikTok and its role in COVID-19 information propagation. *Journal of adolescent health*, 67(5): 730.

Paksoy, T.; Şen, S. C.; Ustaoğlu, G.; and Bulut, D. G. 2023. What do TikTok videos offer us about dental implants treatment? *Journal of Stomatology, Oral and Maxillofacial Surgery*, 124(1): 101320.

Ravanelli, M.; Parcollet, T.; Plantinga, P.; Rouhe, A.; Cornell, S.; Lugosch, L.; Subakan, C.; Dawalatabad, N.; Heba, A.; Zhong, J.; Chou, J.-C.; Yeh, S.-L.; Fu, S.-W.; Liao, C.-F.; Rastorgueva, E.; Grondin, F.; Aris, W.; Na, H.; Gao, Y.; Mori, R. D.; and Bengio, Y. 2021. SpeechBrain: A General-Purpose Speech Toolkit. ArXiv:2106.04624, arXiv:2106.04624.

Rohit Shewale. 2024. YouTube Statistics For 2024. https://www.demandsage.com/youtube-stats. Accessed: 2024-1-12.

Russell, J. A.; and Mehrabian, A. 1977. Evidence for a three-factor theory of emotions. *Journal of research in Personality*, 11(3): 273–294.

Serengil, S. I.; and Ozpinar, A. 2021. HyperExtended LightFace: A Facial Attribute Analysis Framework. In *Proceeding of IEEE ICEET '21*, 1–4.

Shang, L.; Kou, Z.; Zhang, Y.; and Wang, D. 2021. A multimodal misinformation detector for covid-19 short videos on tiktok. In *2021 IEEE BigData*, 899–908. IEEE.

Sokolova, K.; and Kefi, H. 2020. Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of retailing and consumer services*, 53: 101742.

Southwick, L.; Guntuku, S. C.; Klinger, E. V.; Seltzer, E.; McCalpin, H. J.; and Merchant, R. M. 2021. Characterizing COVID-19 content posted to TikTok: public sentiment and response during the first phase of the COVID-19 pandemic. *Journal of Adolescent Health*, 69(2): 234–241.

speechbrain. 2021. Emotion Recognition with wav2vec2 base on IEMOCAP. https://huggingface.co/speechbrain/emotion-recognition-wav2vec2-IEMOCAP. Accessed: 2023-07-10.

Su, Y.; Baker, B. J.; Doyle, J. P.; and Yan, M. 2020. Fan engagement in 15 seconds: Athletes' relationship marketing during a pandemic via TikTok . *International Journal of Sport Communication*, 13(3): 436–446.

Tafesse, W.; and Wood, B. P. 2021. Followers' engagement with instagram influencers: The role of influencers' content and engagement strategy. *Journal of retailing and consumer services*, 58: 102303.

Tian, Z.; Dew, R.; and Iyengar, R. 2023. Mega or Micro? Influencer Selection Using Follower Elasticity. *Influencer Selection Using Follower Elasticity (June 14, 2023)*.

Tricomi, P. P.; Chilese, M.; Conti, M.; and Sadeghi, A.-R. 2023. Follow us and become famous! insights and guidelines from instagram engagement mechanisms. In *Proceedings of the 15th ACM WebSci '23*, 346–356.

Valdez, P.; and Mehrabian, A. 1994. Effects of color on emotions. *Journal of experimental psychology: General*, 123(4).

Vijay, D.; and Gekker, A. 2021. Playing politics: How Sabarimala played out on TikTok. *American behavioral scientist*, 65(5): 712–734.

Wagner, J.; Triantafyllopoulos, A.; Wierstorf, H.; Schmitt, M.; Burkhardt, F.; Eyben, F.; and Schuller, B. W. 2023. Dawn of the Transformer Era in Speech Emotion Recognition: Closing the Valence Gap. *IEEE TPAMI*, 1–13.

Wang, Y.; Salim, N. A. M.; Subri, S.; Zhang, X.; and Zhu, M. 2023. The Features of TikTok Viral Video Advertising: A Systematic Review . In *2023 ICIMCIS*, 267–272. IEEE.

Weimann, G.; and Masri, N. 2023. Research note: Spreading hate on TikTok. *Studies in conflict & terrorism*, 46(5).

Wu, J.; Harlow, A. F.; Wijaya, D.; Berman, M.; Benjamin, E. J.; Xuan, Z.; Hong, T.; and Fetterman, J. L. 2022. The impact of influencers on cigar promotions: a content analysis of large cigar and swisher sweets videos on TikTok . *IJERPH*, 19(12): 7064.

Yang, J.; Zhang, J.; and Zhang, Y. 2021. Influencer video advertising in TikTok. *Mit Initiative on The Digital Economy*.

Yang, Y.; and Ha, L. 2021. Why people use TikTok (Douyin) and how their purchase intentions are affected by social media influencers in China: A uses and gratifications and parasocial relationship perspective. *Journal of Interactive Advertising*, 21(3): 297–305.

Ethics Statement

Similar to previous work (Ling et al. 2022), our study exclusively utilizes publicly available data and does not involve human subjects, thus exempting it from a formal review by our institution's IRB. Nonetheless, our experiments are all performed by adhering to the guidelines of the Menlo report (Bailey et al. 2012). For instance, we report data only in aggregated form and do not pose risks to TikTok users, e.g., attracting unwanted attention. Our released datasets (which incorporates FAIR principles) excludes the raw video data due to potential privacy settings or deletions. Instead, we provide the videos' URLs and code for hydration. We collected (public) data using automated methods, which is discouraged by the ToS. However, as discussed in a recent ICWSM paper (Fiesler, Beard, and Keegan 2020), "online data collection decisions should extend beyond ToS and consider contextual factors." TikTok's ToS permits manual collection, implying that automated collection is likely restricted to prevent server overload (Fiesler, Beard, and Keegan 2020). We took measures to ensure our data collection did not burden TikTok servers and received no warnings or bans from the platform.

Checklist

1. For most authors...

   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes

   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes.

   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes, we justified any choice regarding our methodological approach according to our claims.

   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes

   (e) Did you describe the limitations of your work? Yes.

   (f) Did you discuss any potential negative societal impacts of your work? Yes

   (g) Did you discuss any potential misuse of your work? Yes, in the Ethics statement.

   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes, in the Ethics statement.

   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes, we have read, and our paper conforms to them.

2. Additionally, if your study involves hypotheses testing...

   (a) Did you clearly state the assumptions underlying all theoretical results? Yes

   (b) Have you provided justifications for all theoretical results? Yes, justification is provided using statistical testing and experimental evaluation.

   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? Yes.

   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Yes, we have discussed alternative mechanisms in future works.

   (e) Did you address potential biases or limitations in your theoretical framework? Yes.

   (f) Have you related your theoretical results to the existing literature in social science? NA

   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? NA

3. Additionally, if you are including theoretical proofs...

   (a) Did you state the full set of assumptions of all theoretical results? NA

   (b) Did you include complete proofs of all theoretical results? NA

4. Additionally, if you ran machine learning experiments...

   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Yes, we provided a URL with source code and datA-

   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? Yes.

   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? Yes, we have added the error bars and standard deviation for such results.

   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? We specified the total amount of computing in our repository.

   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes, along with the evaluation, we have also performed statistical tests for the validation.

   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? NA

5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...

   (a) If your work uses existing assets, did you cite the creators? Yes

   (b) Did you mention the license of the assets? NA

   (c) Did you include any new assets in the supplemental material or as a URL? NA

   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? NA

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? Yes, we are not using any personally identifiable information.

(f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see **?**)? Yes, our dataset will adhere to FAIR principles.

(g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see **?**)? Yes.

6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...

(a) Did you include the full text of instructions given to participants and screenshots? NA

(b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA

(d) Did you discuss how data is stored, shared, and deidentified? NA