

Temporal Analysis of User Engagement on Instagram

Elin Thorgren

Supervisor : Alireza Mohammadinodooshan
Examiner : Niklas Carlsson

Upphovsrätt

Detta dokument hålls tillgängligt på Internet - eller dess framtida ersättare - under 25 år från publiceringsdatum under förutsättning att inga extraordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnd som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida <http://www.ep.liu.se/>.

Copyright

The publishers will keep this document online on the Internet - or its possible replacement - for a period of 25 years starting from the date of publication barring exceptional circumstances.

The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/hers own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its www home page: <http://www.ep.liu.se/>.

Abstract

Social media platforms have during the last decade expanded immensely in both the number of users and posts. An extensive number of previous studies have observed how users engage on these platforms to study how social systems work. However, compared to other social media platforms, Instagram has been less extensively researched. This work analyses user engagement on the social media platform Instagram for different characteristics related to posts, uploaders, and media files, consisting of photos and videos. What differentiates this work from others is studying the temporal dynamics of users' engagement across albums, photos, and videos. The results show that album posts receive the highest number of interactions and have the longest engagement lifespan, which is followed by photo posts. Additionally, the most important characteristic that attracts users' interactions is related to the uploader and includes their social network size and uploading rate. Further, different categories of users are analysed with respect to the post type. Compared to other influencer groups, brands and other types of organisations receive fewer interactions and musicians tend to have more loyal followers on Instagram. The conducted analysis may influence brands' and influencers' marketing strategies on various social media platforms and the result can influence the creation of analytical models to predict the temporal dynamics of user engagement on Instagram.

Acknowledgments

Firstly, I would like to extend my sincere thanks to my examiner Niklas Carlsson for his support throughout the thesis process. I would also like to express my appreciation to my supervisor Alireza Mohammadinooshan for his guidance and for collecting the data which enabled this thesis to be performed. I am really thankful for the help I have received from both of you. Lastly, I want to thank my family and boyfriend for all their love and support throughout the thesis process.

Contents

List of Figures	vii
List of Tables	ix
1 Introduction	1
1.1 Motivation	1
1.2 Aim	2
1.3 Research Questions	2
1.4 Delimitations	2
1.5 Thesis Outline	2
1.6 Publication	2
2 Background	3
2.1 Object Detection	3
2.2 Metrics	4
2.3 Related Work	5
3 Method	7
3.1 Dataset	7
3.2 Analysis	8
4 Initial Analysis	11
4.1 Temporal Dynamics of Interactions	11
4.2 Statistical Correlation	12
5 Post Type Analysis	14
5.1 Initial observations	14
5.2 Posts' Characteristics	15
5.3 Uploaders' Characteristics	17
5.4 Video Views	18
6 Media Analysis	20
6.1 Comparison Between Datasets	20
6.2 Media Features	20
6.3 Content Features	22
7 Categorical Analysis	25
7.1 Post Types	26
8 Discussion	28
8.1 Results	28
8.2 Methodology	30
8.3 The Work in a Wider Context	31

9 Conclusion	32
9.1 Results	32
9.2 Future Work	32
Bibliography	33
A Top-100 Users Analysis	37
B Additional Post Type Analysis	44
C Additional Media Analysis	47

List of Figures

4.1	Cumulative fraction of interactions (d: days, h: hours, min: minutes).	11
4.2	Correlation matrices for the statistical data.	12
5.1	Distributions of total interactions	14
5.2	Temporal impact of post types.	15
5.3	Temporal impact of words in posts' descriptions (w: words).	15
5.4	Average impact of hashtags and mentions in posts' descriptions.	16
5.5	Temporal impact of uploaded posts 7 days prior (p: posts).	17
5.6	Average impact of uploaded posts 7-days prior.	17
5.7	Average impact of users' number of followers.	17
5.8	Temporal and average impact of video views (v: views).	18
5.9	Average impact for video views.	18
6.1	Average impact of media files in albums.	21
6.2	Temporal impact of video duration (s: seconds).	21
6.3	Average impact of video duration (s: seconds).	21
6.4	Average impact of people in media files.	22
6.5	Temporal impact of genders.	22
6.6	Average impact of emotions.	23
6.7	Average impact of age.	23
7.1	Temporal impact of categories.	25
7.2	Average impact of post types for each category.	26
A.1	Cumulative fraction of interactions (d: days, h: hours, min: minutes).	37
A.2	Correlation matrices for the statistical data.	38
A.3	Distributions of total interactions	38
A.4	Temporal impact of post types.	39
A.5	Average impact of words in posts' descriptions.	39
A.6	Temporal impact of words in posts' descriptions (w: words).	39
A.7	Average impact of hashtags and mentions in posts' descriptions.	40
A.8	Temporal impact of hashtags in posts' descriptions).	40
A.9	Temporal impact of mentions in posts' descriptions.	40
A.10	Average impact of uploaded posts 7 days prior.	41
A.11	Temporal impact of uploaded posts 7 days prior (p: posts).	41
A.12	Average impact of followers.	41
A.13	Temporal impact of followers (f: followers).	42
A.14	Temporal and average impact of video views (v: views).	42
A.15	Average impact for video views.	43
B.1	Average impact of post type	44
B.2	Average impact of words in posts' descriptions.	44
B.3	Temporal impact of hashtags in posts' descriptions.	45

B.4	Temporal impact of mentions in posts' descriptions.	45
B.5	Average impact of uploaded posts 7 days prior.	45
B.6	Temporal impact of followers (f: followers).	46
C.1	Temporal impact of people in media files.	47
C.2	Average impact of gender.	47
C.3	Temporal impact of emotions.	48
C.4	Temporal impact of age.	48

List of Tables

3.1	Summary of the number of users, posts, and engagement for each category of users.	10
4.1	The most important variables extracted from the dataset.	12
A.1	Summary of the number of users, posts, and engagement for each category of users.	37



1 Introduction

As social media platforms such as Twitter and Facebook grow, there is a simultaneous growth in both the user base and the volume of content shared on these platforms. Multiple posts are uploaded each second on social media platforms but only a small fraction of them become popular and receive significantly more interaction such as likes, shares, or views compared to the average post. This thesis intends to analyse what impacts those posts to receive more interactions.

1.1 Motivation

The social media platform Instagram is intended for media sharing. Users, also referred to as uploaders, can upload photos, videos, and albums consisting of photos and/or videos, along with an optional description of each post. Users can interact with posts through likes, comments, and shares. Additionally, users can follow each other creating a social system. Studying the social system and identifying how and why users engage on Instagram can be beneficial in various aspects. Uploaders can learn how to gain and maintain a larger social network to make a profit. These uploaders are mainly referred to as influencers in their respective domain. In marketing, businesses can utilise studies to reach out to new consumers and promote brands and products. For research purposes, it is beneficial to study how people interact on social media platforms to understand human behaviour in general and how social systems on specific social media systems work. Previous studies have, for example, identified cyberbullying [6] and drug-related posts [51] on Instagram.

A large number of previous works conducted on Instagram have studied how users' engagements are impacted by post- or uploader-related characteristics such as followers, total uploads, or the use of hashtags [12, 21, 45]. Others have focused on what is present in media files of posts [1, 22, 28], which consists of uploaded photos or videos. Various models have also been developed to predict posts' popularity on Instagram [14, 52, 53]. For other photo- and video-sharing platforms, the temporal dynamics of users' engagement have been studied extensively [2, 29, 47]. For Instagram however, the literature is scarce.

1.2 Aim

This thesis aims to analyse users' temporal engagement on Instagram. The temporal impact of characteristics, which consists of posts, uploaders, and media files, are analysed. Unlike previous works, the analysis will also be performed with respect to the three post types: albums, photos, and videos. Additionally, different categories of uploaders are analysed with respect to the post type to further observe how users engage on Instagram.

1.3 Research Questions

Given the above objectives, the thesis set out to address the following research questions:

1. How do different characteristics influence users to engage in Instagram posts over time?
2. What are the most important characteristics that attract users' interactions?
3. How do users' engagements with Instagram posts change for different categories of uploaders?

1.4 Delimitations

In this thesis, only the top most followed users on Instagram are considered. This is due to the tool used to collect data, CrowdTangle, being limited in what data can be collected. CrowdTangle only collects data from verified Instagram accounts or public Instagram accounts with more than 50K followers [8].

1.5 Thesis Outline

The remainder of the thesis is organised as follows. Chapter 2 introduces background related to the methodology and previous work related to the analysis. In Chapter 3, the methodology and the collected dataset are presented. An initial analysis of the collected dataset is provided in Chapter 4, while Chapter 5 observes the impact of posts' type as well as users' and posts' characteristics. Chapter 6 presents an analysis of the media in each post and Chapter 7 concludes the analysis by observing how users engage differently among different categories of uploaders. A final discussion is presented in Chapter 8 and a conclusion of the analysis is drawn in Chapter 9. Additionally, Appendix A presents an analysis of the top-100 most followed users among the collected data. Further, Appendix B presents an additional analysis of the 1000 collected users and Appendix C presents an additional analysis of the media files.

1.6 Publication

This thesis resulted in our recently accepted research paper [44]:

Elin Thorgren, Alireza Mohammadinodooshan, and Niklas Carlsson, "Temporal Dynamics of User Engagement on Instagram: A Comparative Analysis of Album, Photo, and Video Interactions", In. Proc. ACM Web Science Conference (WebSci), Stuttgart, Germany, May 2024.

While the work and thesis were completed in 2023, to ensure double-blind submission (used by the conference), we held the publication of this thesis until the paper was accepted (in 2024).



2 Background

The following chapter presents detection models and metrics related to the methodology and previous works related to the conducted analysis.

2.1 Object Detection

Object detection is a technology used to locate and detect objects in images and videos. Object detection models can be divided into two separate methods: one-stage or two-stage methods, in which the former prioritises speed and the latter prioritises detection accuracy [26]. One example of a two-stage object detector is the Faster Region-based Convolutional Network (R-CNN) [36]. Faster R-CNN first uses a Regional Proposal Network (RPN) to generate a set of object proposals by, for example, differentiating between objects' features such as colour or texture. Then, the Faster R-CNN precursor, Fast R-CNN [15], is used to classify each proposal as foreground classes or background using a CNN. The RPN is a fully CNN, creating an end-to-end detector with Fast R-CNN.

Object detection also includes other forms of detection such as object tracking or face detection. Face detection is the task of detecting human faces in images and videos. One such one-stage model is RetinaFace [9]. RetinaFace uses a combination of object detection of faces, 2D landmark localisation, and 3D face reconstruction to detect and locate faces in images. The object detection model ResNet-50 [16] is used to detect faces and estimate bounding boxes for each detected face. For 2D landmark localisation, a Feature Pyramid Network (FPN) [25] is used to generate feature maps that can locate features in images. The maps RetinaFace extracts correspond to one of the five facial landmarks: nose, eyes, and the corners of the mouth. Lastly, through a 3D vertices regression, 1,000 vertices are predicted to capture more facial details. These are, along with the bounding boxes and five landmarks, used to calculate a joint accuracy score of each detected face. Face detection models such as RetinaFace can then be utilised in a pipeline to, for example, extract facial attributes, recognise peoples' identities, or compare similarities of peoples' faces [39, 40].

2.2 Metrics

The following section presents metrics that can be used to calculate the correlation or statistical significance between samples.

2.2.1 Pearson's Population Correlation Coefficient

In statistical analyses, it is common to measure the correlation between variables to observe the relationship between them easily. The correlation can be calculated for various dependencies and sensitivities to outliers, but the most common of them is Pearson's correlation coefficient. For a statistical population, Pearson's population correlation coefficient, from now on referred to as Pearson's coefficient, can be applied to measure the linear relationship between variables for a population [32]. The coefficient, sensitive to outliers, is denoted as

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y},$$

where cov is the covariance between X and Y , and σ is the standard deviation of X and Y respectively. The result of the equation is a number between -1 and 1, where a value of -1 indicates a perfect negative correlation and a value of 1 indicates a perfect positive correlation. Additionally, a value of 0 indicates no correlation.

2.2.2 Kruskal-Wallis Test

The Kruskal-Wallis test [23], also called one-way analysis of variance (ANOVA) on ranks, can be used to determine the statistical significance between three or more samples. Compared to one-way ANOVA, the Kruskal-Wallis test does not assume a normal distribution. The test questions the null hypothesis, that all samples have identical distributions, against the alternative hypothesis, that at least one sample has a different distribution, with respect to the median. It does so by calculating the H-value, denoted as

$$H = (N - 1) \frac{\sum_{i=1}^s n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^s \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2},$$

where N is the total number of observations for all samples, s is the number of samples, n_i is the number of observations for sample i , \bar{r}_i is the average rank of all observations in sample i , \bar{r}_{ij} is the rank of observation j in sample i , and \bar{r} is the average of all \bar{r}_{ij} . The p-value measures the probability of acquiring the obtained results, given that the null hypothesis is true. The value can be used to either reject the null hypothesis or not.

2.2.3 Dunn's Test

Since the Kruskal-Wallis test can only reject or retain the null hypothesis, Dunn's test [10] can perform pairwise comparisons across all samples to determine the statistical significance between each pair of samples for rejected null hypotheses. Dunn's test utilises the average rank W_s for each sample s from the Kruskal-Wallis test to calculate the z-value between two samples A and B

$$z_{A,B} = \frac{\bar{W}_A - \bar{W}_B}{\sqrt{\left(\frac{N(N+1)}{12} - \frac{\sum_{i=1}^r \tau_i^3 - \tau_i}{12(N-1)}\right) \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}},$$

where N is the total number of observations for both samples, r is the number of tied ranks for both samples, τ_i is the number of observations for both samples with the i^{th} rank.

2.2.4 Mann-Whitney U Test

Similar to Dunn's test, the Mann-Whitney U Test [27] can be used to determine the statistical significance between only two sample medians. The test also questions the null hypothesis against the alternative hypothesis but with the z-value calculated as

$$z_{A,B} = \frac{\min(U_A, U_B) - \frac{n_A n_B}{2}}{\sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}}},$$

where $U_A = n_A n_B + \frac{n_A(n_A+1)}{2} - R_A$ for sample A, n is the number of observations in each sample, and R is the sum of ranks of all observations in each sample.

2.2.5 Bootstrap

The bootstrap method [11] can be used to determine if there is a statistical significance between two sample means. It is useful when the underlying distribution of the data is unknown. Bootstrapping refers to the method of resampling which enables the confidence interval (CI) for a skewed distribution to be calculated. The CI for two samples with respect to the mean can be calculated as follows:

1. Bootstrap samples of size N are randomly generated from the original samples with replacement.
2. The means of the bootstrap samples are calculated and subtracted from each other.
3. The result is saved and step 1 and 2 is repeated X number of times.
4. The results from the iterations are ordered in magnitude and used to find the lower and upper limits of a set CI. For a 99% CI, the means corresponding to these limits are found at the 0.5% and 99.5% percentiles.

If the given CI does not include the null value, there is a statistical difference between the means of the two samples under the null hypothesis.

2.2.6 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test [7], also referred to as the KS test, can be used to determine the statistical significance between two samples with respect to their distributions. The test calculates the difference between the empirical cumulative distribution functions (eCDF) of two samples as

$$D_{A,B} = \max_x |F_A(x) - F_B(x)|,$$

where $F_A(x)$ and $F_B(x)$ are the eCDF's of sample A and B.

2.3 Related Work

The following section presents previous work related to the analysis. What differentiates this analysis from previous works is partly that the posts' type is considered. However, it should be noted that albums on Instagram were not introduced until February 2017 [48].

For Instagram, there have been many prior works analysing users' engagements. Jang et al. [21] studied some post and uploader-related characteristics but had their main focus on users' social networks and their use of hashtags. They found, through the use of uploaders' hashtag behaviours, that users with narrow interests were generally more popular than users with broader interests. The result contradicts a similar analysis done by Ferrara et al. [12] which presented that unpopular users tend to have narrow interests. However, they also

found that popular users received user engagement whether they had extremely narrow or broad interests.

Temporal analyses of users' engagements on platforms such as Facebook [37], Youtube [3], and Twitter [13] have been studied a fair amount. For Instagram, the number of previous works is fewer. Vassio et al. [45] analysed users' temporal engagement on Facebook and Instagram considering uploaders' and followers' activities. For both platforms, they found that the popularity of posts is diverse, even for the same uploader. Additionally, a high uploading rate among uploaders decreased the popularity of previous posts. However, no further characteristics were taken into account. Huang et al. [20] studied posts' and uploaders' characteristics and analysed the publish time of posts further. They found that users that upload posts after noon receive more interactions, but the number of received interactions is similar each day of the week. However, the dynamics of users' engagement over time were never considered.

A majority of the previous work studying Instagram has created models to predict the popularity of current or future posts. Such models included [14, 52, 53] or excluded [4] the impact of media content. Zhang et al. [52] created a model to predict the popularity of posts based on only the image and description of each photo post. They found that the unpopular images contained posters and food. However, descriptions were seen as a more reliable estimate for their prediction model. Gayberi and Oguducu [14] created a model to predict the popularity of posts based on posts' and uploaders' characteristics. Additionally, the media characteristics were accounted for by extracting features such as people, vehicles, and if the environment was indoor or outdoor. They, compared to previous works, also accounted for the type of post: album, photo, or video. However, unlike their model, this analysis presents a deeper understanding of individual aspects that may affect how users engage over time on Instagram.

Other than observing if people are present in images, several works have studied the facial attributes of people present in images on Instagram to further understand what drives users to engage with different kinds of content. Bakhshi et al. [1] found that photo posts on Instagram receive over 30% more interactions if a face was present in the image. However, the number of faces, as well as the age and gender of the people present in images did not have a significant impact on the number of user interactions. Jang et al. [22], explored other perspectives and extracted the ages from profile images and descriptions on Instagram. They studied how teens and adults (25-39 years old) engage on Instagram and found that teens use more interactions than adults. Additionally, they observed that teens upload fewer posts compared to adults. Lastly, Mazloom et al. [28] studied post-popularity among brands' media files. They found that images containing both people and the brand's product received more engagement compared to images including people or the product separately. Unlike their work, this analysis takes into account different categories of uploaders, other than brands, to observe how users engage among them.



3 Method

The following chapter presents how the dataset used for the analysis was gathered and what it consists of, as well as the analysis approach.

3.1 Dataset

The dataset can be divided into two separate categories: statistical data and media content.

3.1.1 Statistical Data

The statistical data consists of information about 1000 Instagram users and nearly all of their posts collected in the one-year period between December 15, 2021, and December 14, 2022. The selected users were among the most followed accounts on Instagram according to top lists for influencer and brand accounts found at starngage.com [42]. Note that Starngage do not provide complete lists of the most followed accounts on Instagram but they are publicly available to use [34, 46, 43]. Additionally, only accounts that existed while the data was gathered, as well as accounts that had uploaded posts in the targeted timeline were accounted for.

To collect statistical data for the selected Instagram users, CrowdTangle was used. CrowdTangle is a public insights tool that tracks public content on social media platforms such as Instagram and Facebook [8]. The data of the 1000 users collected from CrowdTangle consisted of the following:

- **User data:** Information about the users' account such as name, username, and the number of followers at the time it was last collected by CrowdTangle.
- **Post data:** Information about each post such as the type (album, photo, or video), description, date of upload, and total number of likes, comments, and views at the time it was last collected by CrowdTangle. Additionally, information regarding the user such as the number of followers when the post was uploaded was available.
- **Temporal data:** Information for each post regarding the number of likes and comments at specific time instances during the first 20 days after upload. A maximum of 75 instances for each post was gathered by CrowdTangle, mimicking a logarithmic curve

with the interval of each captured time instance varying from 15 minutes to 24 hours between one another.

Approximately 450K posts were collected, with a distribution of around 1 post to 25K posts per user. Note that CrowdTangle is not able to collect all posts of the accounts it covers and therefore not all the posts of the targeted accounts were indexed in their database.

3.1.2 Media content

Other than statistical data, the content of each post can also contribute to users' engagement. To analyse the media content, a subset consisting of the top-100 most followed users on Instagram, according to Starngage [42] and the criteria mentioned in Section 3.1.1, were selected. Nearly all media files for each post in the sample were gathered. Approximately 45K media files were collected, with a distribution of 1 file to 4,600 files per user. The collected media files consisted of photos and/or videos depending on the type of post that was uploaded.

To make an equal analysis among media files in photo, video, and album posts, only the first image users see of each post was considered, like in [14]. Hence, the first frame in videos and the first media file in each album. The limitation omits important insights that could contribute to engagement among users. However, it was deemed necessary to make fair comparisons between different types of media files and posts.

3.2 Analysis

To get a deeper understanding of users' temporal engagement on Instagram and to distinguish this work from previous analyses of Instagram, the analysis was divided into 4 sub-steps as presented below.

Initial analysis

An initial analysis of the top-1000 users was performed to get an overview of the collected dataset and how users engage over time. The correlation between variables extracted from the dataset was calculated across all posts using Pearson's population correlation coefficient to measure the linear relation between different variables. Additionally, the correlation was calculated across users to understand the importance of users' identities among posts.

Post type analysis

The correlation between the extracted variables and users' engagement over time was further analysed. The analysis was made with respect to the post types to observe if and how users engage differently among them. The distribution of user interactions was observed for the post types through cumulative distribution functions (CDF) and complementary cumulative distribution functions (CCDF). Additionally, the extracted variables were divided into post and uploader characteristics and analysed with respect to the post type. The statistical significance of the results was also calculated. The statistical significance was calculated with respect to the median using the Kruskal-Wallis test along with Dunn's test. The Mann-Whitney U test was used to confirm the result to prevent false discovery of statistical significance. Bootstrapping with 10K iterations and a 99% CI was used to determine the statistical significance between sample means, while the Kolmogorov-Smirnov test looked further at the distributions. The threshold for statistical significance was set to the p-value of 0.005 for all tests but bootstrapping. The statistical significance was also calculated for all results in the following analyses.

Media analysis

The temporal dynamics were further analysed with respect to the media files for each post type. However, as the dataset containing media content only contained a subset of the total amount of collected users, a comparison between the top-100 and top-1000 most followed users was first made. The comparison established differences to consider in further analysis.

To analyse the content in each post's media file, people and facial attributes were extracted from the media files in each post. To detect people in images, the library Detectron2 [50] was used. Detectron2, created by Facebook, is an open-source framework that provides a large number of pre-trained state-of-the-art object detection and segmentation models. This analysis implemented their Faster R-CNN model with a Feature Pyramid Network (FPN) [25] backbone built on ResNet 101 [16]. The model and backbone were chosen for their good box average precision and fast predictions compared to other models [19, 33]. The model was implemented with default settings and can detect objects such as people, cars, cats, and dogs. However, for this analysis, only people were detected.

To extract facial attributes in images, the library DeepFace [38] was used. DeepFace is an open-source framework that provides state-of-the-art facial recognition models and a facial attribute analysis module for the attributes age, gender, emotion, and race [39]. This analysis only analysed the facial attributes age, gender, and emotion. For the facial attribute analysis module, a variety of face detection models could be selected. This analysis used a re-implementation of the face detection model RetinaFace [9] for its performance compared to the other available models [39]. The models for all facial attributes, except emotion, were built on the face recognition model VGG-Face [31] for its over-performing scores [18, 49]. The number of output layers of the models was however dependent on the number of facial features that could be predicted. The gender model could be used to predict either the genders man or woman, while the age model could be used to predict ages from 0 to 100. Note that the results of the predictions are based on what the model detects as, for example, female or male features. Therefore, the results do not necessarily have to be the actual gender, age, or emotion of each detected person. For the emotion model, a CNN was instead used to detect one of 7 emotions: angry, disgust, fear, happy, sad, surprise, or neutral. However, only happy, neutral, and sad were analysed, thus they consisted of the majority of predicted emotions.

Categorical analysis

The top-1000 most followed users were categorised into 5 different categories. The analysis performed on these different categories consisted of a comparison of how users engage between each category, as well as the post types for the different categories. To categorise the collected users, ChatGPT 3.5 series was used. ChatGPT is a public pre-trained natural language processing model able to answer questions using an extensive knowledge base obtained from the internet up until September 2021 [5]. Note that ChatGPT is a machine learning model and cannot answer questions with entire certainty. The generated categories were accurate and up-to-date for a majority of users. However, some manual adjustments were made for users that were noticed to belong to the wrong category to diminish the number of misclassified users.

Each user was categorised according to their foremost profession except for accounts that belonged to, for example, brands, organisations, sports clubs, and magazines. These types of users had their own category: brands, thus a majority of them were brands. The rest of the users were categorised into one of the following categories: actors, musicians, athletes, or others. Others include professions such as influencer, comedian, and politician. Table 3.1 presents the number of users and posts in each category, as well as the total engagement within a 20-day period divided by the number of users in each category. The table is sorted based on the number of users for each category.

Category	Number of users	Number of posts				Engagement per user		
		Album	Photo	Video	Total	Likes	Comments	Views
Musicians	252	16,664	15,126	8,769	40,559	94,695,705	687,304	47,822,288
Others	249	24,847	36,268	7,731	68,846	75,405,771	832,286	45,528,958
Actors	236	15,240	9,672	4,158	29,070	57,138,957	364,897	27,983,632
Brands	172	90,828	157,820	47,780	296,428	224,458,563	2,234,431	238,912,556
Athletes	91	4,366	7,787	1,641	13,794	100,104,093	641,083	36,529,477
Total	1,000	151,945	226,673	70,079	448,697	551,803,089	4,760,001	396,776,911

Table 3.1: Summary of the number of users, posts, and engagement for each category of users.

4 Initial Analysis

This chapter presents an initial analysis of the temporal dynamics and statistical correlation of the collected dataset.

4.1 Temporal Dynamics of Interactions

For the collected dataset, the last interval in the temporal data has a minimum threshold of 20 days after a post was uploaded. However, no maximum threshold exists. Therefore, when studying the temporal dynamics, a maximum of 20 days after upload, including day 20, is considered. Hereafter when analysing the temporal dynamics of posts, the total number of interactions achieved 20 days after each upload is referred to as the total amount of interactions.

Figure 4.1 presents the average cumulative fraction of likes and comments for all users over the first 20 days after the posts were uploaded. The figure illustrates the average time posts acquire 25%, 50%, respectively 75% of the total interactions in the 20-day interval. The time thresholds were calculated by taking the average time each post obtained the respective percentage of the total likes. As observed in the figure, Instagram posts are short-lived. The majority of likes are received during the first 24 hours after upload and the majority of comments are received during the first 12 hours.

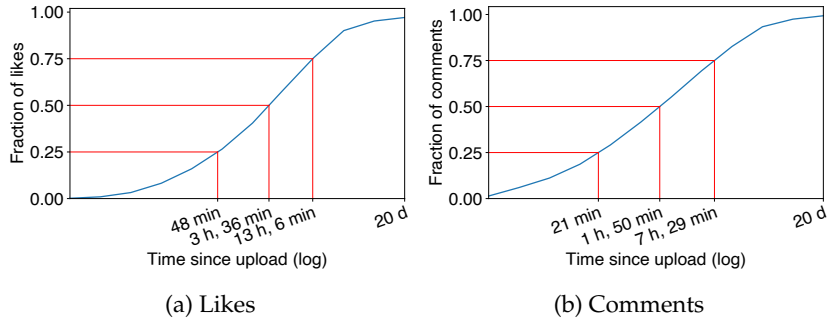


Figure 4.1: Cumulative fraction of interactions (d: days, h: hours, min: minutes).

Name	Description	Characteristics	Scale
Likes	Number of total likes after 20-days	User engagement	Log
Comments	Number of total comments after 20-days	User engagement	Log
Views	Number of views for video posts	User engagement	Log
Words	Number of words in the posts' description	Post	Log
Hashtags (#)	Number of hashtags in the posts' description	Post	Linear
Mentions (@)	Number of mentions in the posts' description	Post	Linear
Followers	Number of followers for the user at upload	Uploader	Log
Posts past week	Number of posts 7-days prior to upload	Uploader	Log

Table 4.1: The most important variables extracted from the dataset.

4.2 Statistical Correlation

To further analyse users' aggregated engagement on Instagram, dependencies between users' interactions and the characteristics of uploaders and posts are further looked upon. Table 4.1 presents the characteristics of the most important variables extracted from the collected data, as well as a description of them. The removed variables were either performance metrics produced by CrowdTangle, or not sufficiently covered. One example of the latter cases is if a post was sponsored or not. Note that the number of words in posts' descriptions includes both hashtags and mentions.

Figure 4.2 presents the correlation matrices for most of the variables calculated using Pearson's population correlation coefficient. Figure 4.2a presents an aggregated calculation over all posts, while Figure 4.2b presents the average correlation seen across the users. To ensure a linear relationship between variables, their characteristics and CDF's were accounted for when deciding their transformation. The scale of each variable is presented in Table 4.1. As observed in Figure 4.2a, the correlation among the variables is moderate to weak. The strongest correlation exists between likes and comments, which is followed by the correlation between user interactions and the characteristics of uploaders. The weak and moderate correlation may suggest that the relation between the variables is other than linear, or the calculated correlation is affected by outliers. Compared to Figure 4.2b, the correlation among the aggregated variables, in regards to users' interactions, is stronger for uploaders' characteristics, similar for posts' characteristics, and weaker between the two types of interactions. The result is in line with previous work which shows that uploaders' identities do not have a large impact on posts' popularity although uploaders' characteristics can [45]. For the remainder of this analysis, the identity of the user will therefore not be considered.

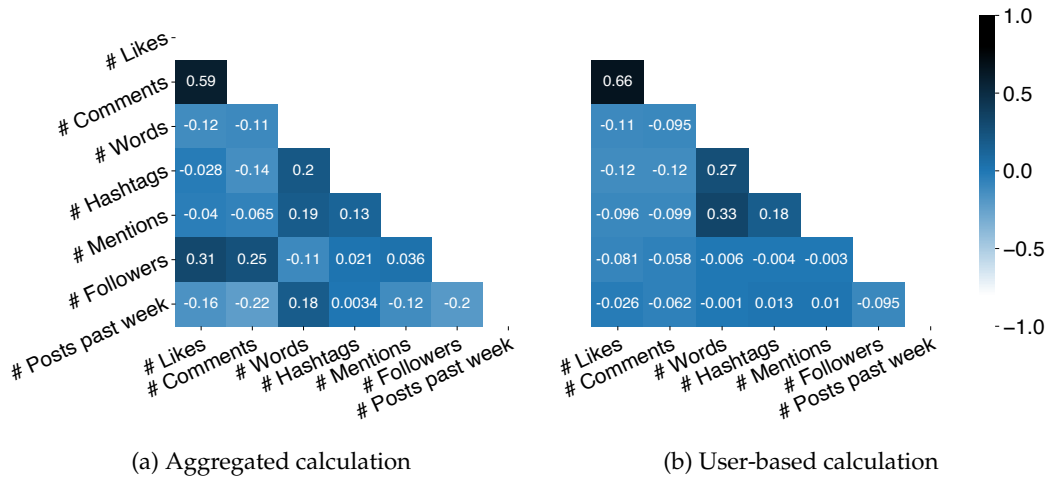


Figure 4.2: Correlation matrices for the statistical data.

For videos, the aggregated correlation between the number of views and likes is 0.92, comments 0.83, followers 0.42, and posts past week -0.24. However, even if the correlation is stronger compared to Figure 4.2a, the correlation between the number of likes and comments for videos is 0.82. The result indicates that the stronger correlation between views and other variables compared to previously observed correlations, is due to the post type. Further, the strong correlation between the number of views and likes may indicate that the number of received likes can be used as a proxy for views.

Videos on Instagram receive additional views if they have been seen for a total of 3 seconds or almost their entire length for videos under 3 seconds, excluding rewatches [8]. Hence, the number of acquired views is a good measure of what users observe. As Instagram only collects views for video posts, a proxy would be beneficial for the remaining post types. However, using likes as a proxy for views should be done with caution. Holmström et al. [17] analysed the temporal relation between retweets and the number of users that opened and read news article links on Twitter. They found that a considerable amount of tweets had more retweets than opened links. Hence, links were shared but not read. Therefore, the number of likes can not be trusted to be a true proxy for views without further research. However, due to the nearly perfect positive correlation between the number of likes and views, likes can, for this analysis, be seen as a proxy for views. The following chapters present the number of views in relation to different characteristics when available. Otherwise, the number of acquired likes is considered a measurement of users' observations.

5 Post Type Analysis

The following chapter presents differences in user engagement among the three types of posts: albums, photos, and videos. An initial analysis first compares the three types. Then, the post types are analysed with respect to post and uploader characteristics. Either the temporal or average impact is presented dependent on the results.

5.1 Initial observations

To capture the distribution of users' interactions on Instagram, Figure 5.1 presents the cumulative distribution function (CDF) and complementary cumulative distribution function (CCDF) for the total number of likes and comments with respect to the post types. The functions show the likelihood of a random variable being higher or lower than a certain value presented on the x-axis. Note that both axes in Figure 5.1b and Figure 5.1d are of logarithmic scale. As observed in the figures, a majority of albums receive more interactions compared to the other post types. On the other hand, most videos receive the least number of interactions. As illustrated by the CCDFs, all post types have a heavy right-tail distribution. Hence, a limited number of users and posts receive an excessive number of interactions. Additionally, photos seem to have a heavier tail compared to albums, which shows that photos include some additional outliers with many interactions.

To capture the temporal increase of interactions between the post types, the posts' temporal data was divided into 8 buckets representing the rounded logarithmic time in seconds after upload. The number of interactions for the largest time instance in each bucket was

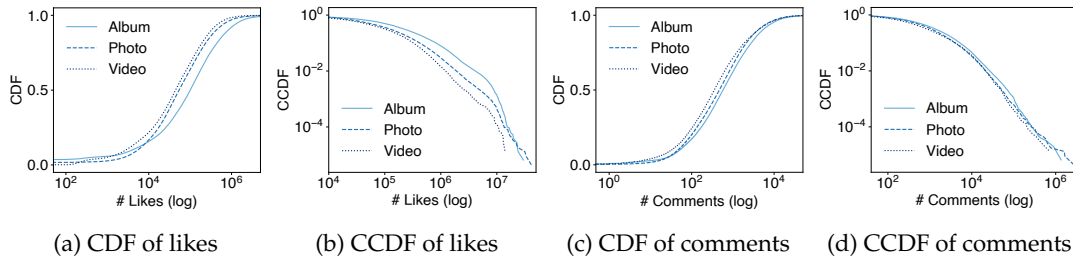


Figure 5.1: Distributions of total interactions

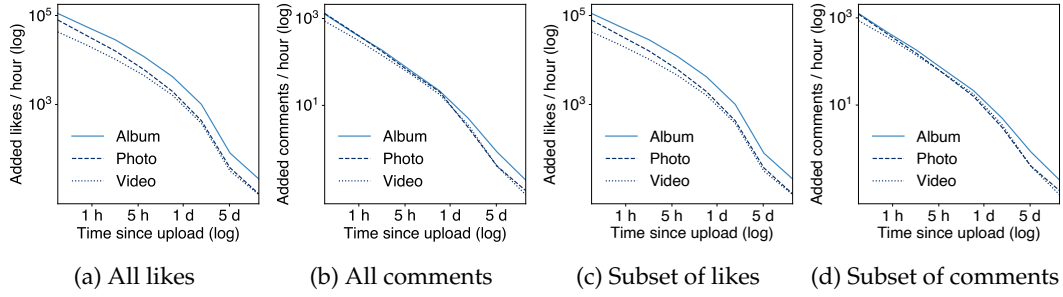


Figure 5.2: Temporal impact of post types.

then subtracted from the adjacent time instances in each bucket for each post. This resulted in the logarithmic temporal increase of interactions for each post. Furthermore, the added interactions for each logarithmic time were divided by the number of hours passed, to give more weight to the initial interactions. The logarithmic result of this calculation is presented in Figure 5.2, but with the two types of interactions separated into two figures.

Figure 5.2 presents the temporal impact of post types for the entire dataset but also for a subset of the data. As observed in Figure 5.2b, photos receive an abundance of comments around 24 hours after upload. The abundance of comments is due to one single post in the data that received 1000 times more comments than the average post within 24 hours. The abnormality was better observed for the upcoming figures and intervened with both the results and interpretations of the results. Therefore, the abnormality was removed from the data to create comprehensive and accurately represented figures for further analyses. The result of the new subset is presented in Figure 5.2c and Figure 5.2d.

As illustrated in Figure 5.2c and Figure 5.2d, users follow a similar pattern among the three post types for both types of interactions, even if the number of interactions is not the same. Albums receive the highest number of added interactions, while videos receive the least number of added interactions during the 20-day time period. Photos receive nearly as many interactions initially as albums but later decline. Although albums receive more total interactions than photos, as illustrated in Table 3.1, more photos are published among the 1000 collected users.

Statistical tests were calculated to determine the significance between post types. The tests were not applied to the temporal data but to data containing the total number of interactions. With a threshold of 0.005, the test indicated statistical significance between the three post types with respect to the median, mean, and distribution (p-values $< 10^{-147}$).

5.2 Posts' Characteristics

Simple characteristics of posts, such as asking a question in the description of a post, can contribute to increased engagement among users. Figure 5.3 presents the average number

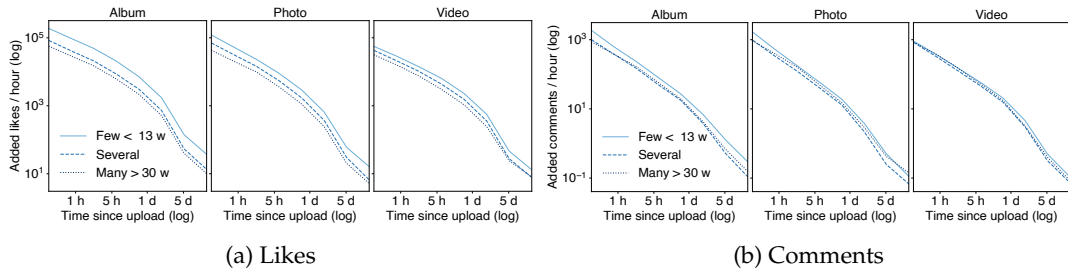


Figure 5.3: Temporal impact of words in posts' descriptions (w: words).

of added interactions received divided by the number of hours passed since upload, with respect to the number of words in the description. Words, digits, hashtags, mentions, and emojis are each counted as one word. The three categories: few, several, and many, are determined to contain roughly the same number of posts. As observed in the two figures, posts with few words in their descriptions receive the highest number of added interactions for all post types. As seen in Figure 5.3a, posts containing several words receive more likes than posts with many words in their description. However, posts containing many words can be noted to receive more comments than posts containing several words. This may be due to uploaders' long descriptions encouraging users to comment.

The statistical tests were applied to the total number of interactions for each post type to determine the statistical significance between the three categories. The results of the tests with respect to the medians and distributions, indicated significant differences for all pairwise comparisons. For the means, bootstrapping showed no significant difference in comments between album posts with several and many words in their descriptions. However, the result does not affect made observations.

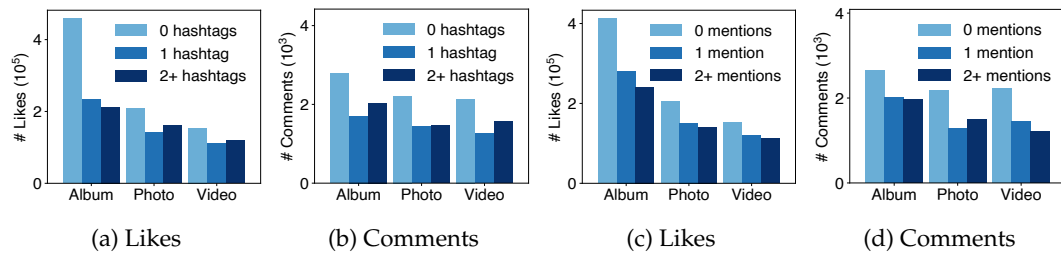


Figure 5.4: Average impact of hashtags and mentions in posts' descriptions.

To further analyse the characteristics of uploaders' descriptions, Figure 5.4 presents the average impact that hashtags and comments have on interactions. Figure 5.4a and Figure 5.4b illustrate the average amount of interactions, that albums, photos, and videos receive based on the number of hashtags included in posts' descriptions. As illustrated by the two figures, posts that receive the most interactions are posts that do not include any hashtags. A similar pattern among the occurrence of mentions in the description of posts can be observed in Figure 5.4c and Figure 5.4d. The figures illustrate the average impact of mentioned users in posts' descriptions. As seen in the figures, posts that do not mention any users receive more user interactions compared to posts that do mention users in their descriptions. Note that a correlation among posts that mentioned users or included hashtags in their description could not be found other than receiving fewer interactions than posts that did not mention users or include hashtags.

The statistical tests showed a significant difference between posts that included mentions and posts that excluded mentions. For hashtags, the Kruskal-Wallis test showed statistical significance (with the highest p-value being $1.97 \cdot 10^{-57}$ for likes and $1.73 \cdot 10^{-104}$ for comments). However, Dunn's test and Mann-Whitney test both showed no difference in likes between videos with hashtags and videos without hashtags (p-values of 0.73 and 0.52 respectively). The result could possibly indicate some additional popularity among videos with more hashtags. The statistical tests were only considered between posts that included hashtags or mentions and posts that excluded hashtags or mentions. Thus, no trend could be observed between posts that mentioned users or included hashtags in their descriptions.

Hashtags and mentions on Instagram have been commonly used in posts to reach out to new users and help uploaders attract a broader audience. Either via the hashtag or mention itself or to guide Instagram's algorithm. The lack of hashtags observed in Figure 5.4 might be due to hashtags being uncommon to use among the most followed users today, unlike in previous years [12, 21]. Additionally, users can tag other users not only through mentions in

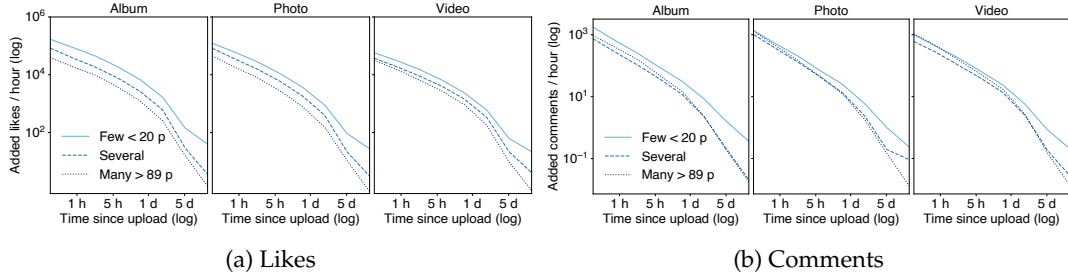


Figure 5.5: Temporal impact of uploaded posts 7 days prior (p: posts).

descriptions but also by tagging them directly in the media files of each post. This may be a reason why mentions are not common in descriptions.

5.3 Uploaders' Characteristics

Besides the characteristics of posts, the activity of uploaders can attract or repel users' interest. Figure 5.5 presents the temporal impact of the number of uploaded posts up to 7 days before the upload of each post. The three categories are determined to contain roughly the same number of posts. As illustrated by the two figures, users who upload fewer posts weekly (less than 3 posts a day on average) receive more interactions on their posts than users who upload more posts. Furthermore, users that upload many posts receive the least amount of added likes which may suggest that users acquire a similar amount of likes which are distributed among the number of uploaded posts. In Figure 5.5b it can be observed that users who upload many posts receive more comments initially, but later receive fewer or a similar amount of comments compared to posts with several prior uploads. Note that for all types of posts and interactions, users who upload 3 posts or more a day on average, seem to receive proportionally fewer interactions at a later stage compared to users who upload less than 3 posts a day on average.

The statistical tests showed significant differences between all pairwise comparisons of medians and distributions. However, bootstrapping indicated no significant difference in comments for photo posts with several and many previously uploaded posts. The result may be due to the observed trend among photos in Figure 5.5b, which cannot be observed for the other post types.

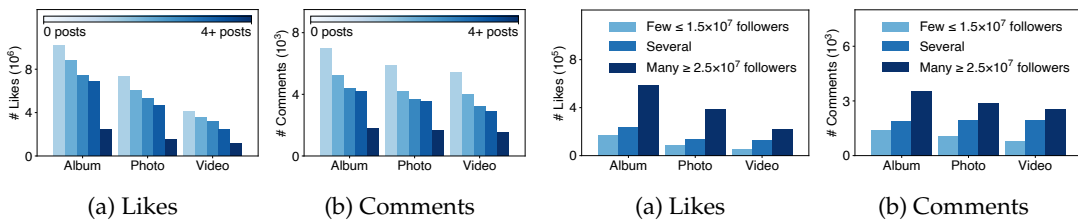


Figure 5.6: Average impact of uploaded posts 7-days prior.

Figure 5.7: Average impact of users' number of followers.

To examine a less frequent upload rate, Figure 5.6 presents the average impact of a lower quantity of uploaded posts 7 days prior. As observed in the figures, posts with more prior uploads generally receive fewer interactions than posts with fewer prior uploads. The negative correlation can be observed for larger quantities of prior uploaded posts as well. This may further suggest that users receive a similar amount of interactions that are then distributed among the uploaded posts. The trend may otherwise suggest that users who receive more interactions tend to upload fewer posts.

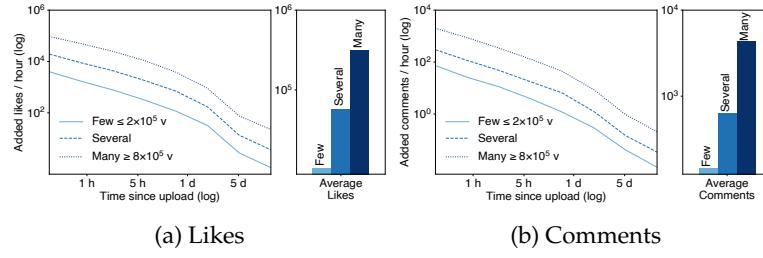


Figure 5.8: Temporal and average impact of video views (v: views).

Another characteristic of uploaders is the size of their social network. Figure 5.7 illustrates the average amount of interactions that albums, photos, and videos receive based on the number of followers the uploader of each post has at the time of upload. The three categories are determined to contain roughly the same number of posts. As seen in the two figures, the more followers a user has, the more interactions they receive. The observed trend may emphasise the importance of possessing a larger social network. Further, the results were supported by the statistical tests which indicated a statistical significance for both types of interactions with respect to the median, mean, and distribution (p-values $< 10^{-62}$).

5.4 Video Views

To further observe the strong correlation between views and other user interactions, Figure 5.8 presents the temporal and average impact of video views. The three categories are determined to contain roughly the same number of posts. As observed in the two figures, the more views videos have, the more interactions they receive. However, the total number of views does not seem to affect the rate at which videos receive interactions. Instead, each video seems to have a similar rate of attracting likes, despite how many views each video has. Note that Figure 5.8 presents an average result and temporal views may yield different results.

Figure 5.9 presents the average number of views considering the impact of post characteristics in Figure 5.9a and uploader characteristics in Figure 5.9b. The three categories for each type of characteristic have the same limits as the previous figures. As observed in Figure 5.9a, posts with few words in their descriptions receive the highest amount of views, followed by posts containing several words. The trend can also be found among likes in Figure 5.3a, but not among comments as posts with several words in their description receive the least amount of comments. For all three types of user engagement, posts without hashtags and mentions in their descriptions seem to receive the highest amount of engagement. The observed results are supported by the statistical test which indicated significant differences (with the highest p-value being $3.14 \cdot 10^{-66}$ for words, $6.81 \cdot 10^{-33}$ for hashtags, and $3.06 \cdot 10^{-220}$ for mentions).

A positive correlation between views and the number of followers can be observed in Figure 5.9b. The trend occurs for all user interactions. Further, a negative correlation can be

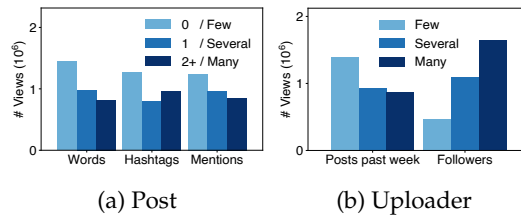


Figure 5.9: Average impact for video views.

observed between views and the number of uploaded posts 7 days prior. The trend is similar for likes, but in Figure 5.7b it can be observed that posts with several prior uploads receive the least amount of comments. The results in Figure 5.9b are supported by the statistical tests which indicated significant differences (with the highest p-value being $1.55 \cdot 10^{-52}$ for posts past week and $3.81 \cdot 10^{-235}$ for followers).



6 Media Analysis

The following chapter analyses media files of posts that belong to the top-100 most followed users in the collected dataset. First, a comparison between the top-100 and top-1000 most followed users is made.

6.1 Comparison Between Datasets

Since the dataset consisting of media content only contained a subset of the total amount collected users, a comparison between the two datasets is made. For the top-100 users, approximately 46K posts were collected, which corresponds to approximately 10% of the total amount of posts collected. Further analysis of the top-100 users is found in Appendix A. The following paragraph reports similarities and differences between the datasets.

For both quantities of users, posts are short-lived and receive a majority of their interaction within the first 24 hours. Additionally, it was promising to possess many followers, to upload fewer posts a week, and to include fewer words and no hashtags or mentions in each post's description. The observed difference was that the top-100 users generally received more interactions and used fewer words in their descriptions compared to the total number of collected users. Hence, the two datasets provided similar trends for all characteristics and would likely yield similar trends for the media analysis presented in the following sections.

6.2 Media Features

This section presents features that are related to the uploaded media but not the content of the media files. One such feature is the number of media files an album consists of. Figure 6.1 presents the average impact of media files in albums in which the bottom and top of the box show the 30th and 70th percentile respectively and the median line and average cross are also illustrated. An album can contain between 2 and 10 files. As seen in Figure 6.1a, the more media files an album has the more likes it generally receives. The trend is also confirmed by the whiskers representing the 20th and 80th percentile respectively. For comments, the trend can also be observed but it is not as strong. However, for both types of interactions, it seems as if albums with 2 media files generally receive the least amount of interactions.

The statistical tests revealed significant differences in medians, means, and distribution between albums containing 2 media files and 10 media files (p-values $< 2.11 \cdot 10^{-44}$). How-

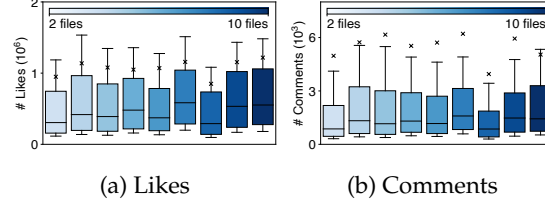


Figure 6.1: Average impact of media files in albums.

ever, not all pairwise comparisons with respect to the number of media files in albums showed a significant difference. With statistical support, the observed differences among the number of media files in albums may indicate that popular users are more likely to upload large albums. The differences can also indicate that large albums have the potential to reach out to more users via the Instagram algorithm or shares.

Another post-specific feature is the duration of video posts. Figure 6.2 presents the temporal impact of video duration. The three intervals are determined to contain roughly the same amounts of posts. As illustrated in Figure 6.2a, short videos receive the highest number of added likes initially, followed by videos with intermediate duration. At a later stage, long videos instead receive more likes, followed by videos with intermediate duration. The latter trend can also be observed for comments in Figure 6.2b.

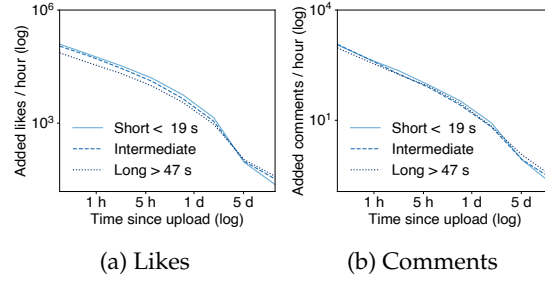


Figure 6.2: Temporal impact of video duration (s: seconds).

The temporal results are expected as it might be time-consuming for most users to watch longer video posts. Additionally, the content of longer videos may be more intriguing to come back and watch later compared to shorter videos. Although longer videos receive more added interactions at a later stage, it can be seen in Figure 6.3 that videos with a short duration receive the highest number of average likes, comments, and views. Furthermore, the whiskers in Figure 6.3, representing the 20th and 80th percentile respectively, indicate that shorter videos generally receive more total interactions compared to longer videos. The results are confirmed by statistical tests which indicated statistical significance for almost all pairwise comparisons. However, the tests of medians and distributions showed no significant difference in comments between intermediate and long video lengths (p-values of 0.03, 0.03, and 0.01). Additionally, no difference in means was seen in comments between all three

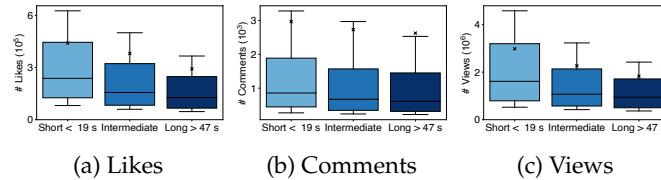


Figure 6.3: Average impact of video duration (s: seconds).

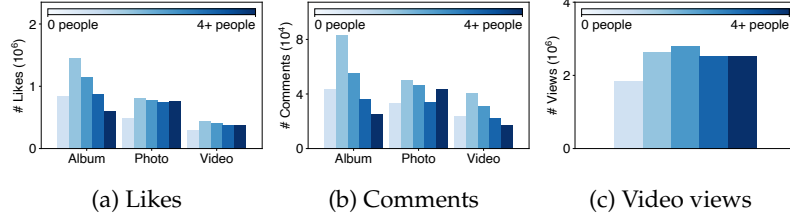


Figure 6.4: Average impact of people in media files.

video lengths. Still, the results observed in the figures may indicate that popular users are more likely to upload shorter videos or shorter videos have the potential to reach more users.

6.3 Content Features

This section observes the content in media files. As previously mentioned, only the first media file in each album and the first frame in each video were analysed. Figure 6.4 presents the average number of interactions acquired, with respect to the number of detected people in each media file. The first column represents images in which no people were present or people were detected with an accuracy under 85%. The threshold was determined by calculating the average accuracy given across all objects detected as people by the model. As illustrated in the figures, albums that contain more people receive fewer interactions. Additionally, it can be observed that albums containing more than two people generally receive fewer interactions than albums in which no people are present. Statistical tests could confirm both observed trends by indicating significant differences with respect to the median and distributions. However, the means of comments indicated no significant difference between albums with no people present and two people present. The trends can also be observed among videos in Figure 6.4b but it is not present for videos in Figure 6.4a nor Figure 6.4c. Similarly to albums, the means of comments only indicated statistical significance between albums with no people present and one person present. However, the acquired video comments did not show a significant difference among all pairwise comparisons for people present in media files. Hence, the negative correlation observed for videos in Figure 6.4b is not statically supported.

For the other instances, the number of people present in media files does not seem to affect the number of received interactions for those figures and photo posts. However, those posts receive more interactions if people are present in the media files compared to posts in which no people are present. These trends observed in the figures are present for larger quantities of detected people as well.

For faces detected by the pre-trained RetinaFace model, Figure 6.5 presents the temporal impact of detected genders. Each detected gender had an accuracy of at least 86%. The threshold was obtained by calculating the average accuracy given across all genders detected

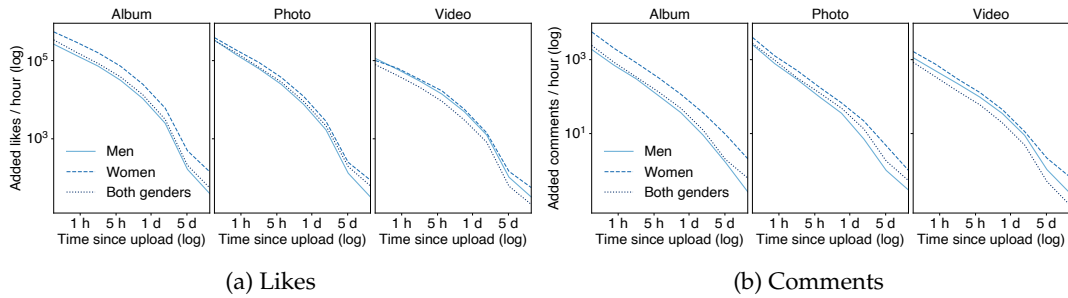


Figure 6.5: Temporal impact of genders.

by the model. The columns men and women include media files in which at least one predicted man or woman was detected. The column both genders includes images in which both genders were detected at least once. As seen in the two figures, files with only women present receive the highest number of interactions for all post types. This is, for albums and photos, followed by files in which both genders are present. For videos, it is followed by files with only men detected. Media files that include women (with or without men) seem to follow a similar trend over time. However, files that only contain men seem for some post types and interactions, such as photos and videos in Figure 6.5a, to receive more interactions initially but later decline.

The Kruskal-Wallis test indicated no significant differences among the categories for videos (p-values of 0.01, 0.06). Additionally, both Dunn’s test and the Mann-Whitney test showed no significant difference in likes between media files in albums in which men and both genders appeared (0.12 and 0.05). For the means, there was no statistical significance between men and both genders for albums and for photos among likes. For videos, there was no significant difference in means between men and women, but those categories were the only pair for videos that had different distributions according to the KS test. The tests’ results may be due to a less significant difference in received interactions among posts with different detected genders.

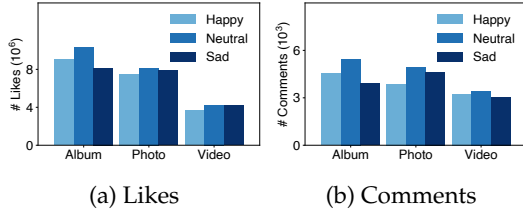


Figure 6.6: Average impact of emotions.

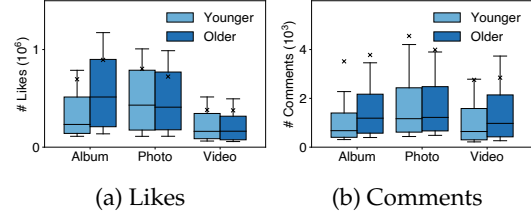


Figure 6.7: Average impact of age.

Figure 6.6 presents the average impact of the predicted emotions: happy, neutral, and sad. Each detected emotion had an accuracy of at least 80%. The threshold was obtained by calculating the average accuracy given across all emotions detected by the model. As seen in the figures, faces expressing neutral emotions receive the highest number of likes. For albums, this is followed by happy expressions, and for photos it is followed by sad expressions. For videos, happy expressions receive more likes but fewer comments than sad expressions. However, statistical tests indicated that the means and distributions of video posts had no significant difference. Additionally, the Kruskal-Wallis test indicated no statistical significance in comments between emotions (p-value of 0.25). For photos, no statistical significance was shown between medians of likes (0.17), means of likes, and the distribution of likes (0.17 between happy and neutral, 0.05 between happy and sad, and 0.08 between neutral and sad). Other pairwise comparisons also indicated no significant difference for albums, photos, and videos. Hence, the observed differences between emotions have no to little statistical support, which may indicate that facial expressions in media files do not contribute to the acquired interactions.

Figure 6.7 presents the average impact of predicted ages. Ages between 0 and 100 years old could be detected and the threshold was determined by the pre-trained model. To diminish the number of misclassified ages, detected people either belonged to the group younger (0-34) or older (35+), thus a majority of users were predicted to be in their early thirties. As illustrated in the figures, older people receive more average interactions on albums and videos, while younger people receive more average interactions on photos. Furthermore, the whiskers representing the 20th and 80th percentile respectively, show for albums and videos that older people are more popular. For photos, the whiskers are denser which may indicate some additional popularity among posts that contain older people. However, it should be noted that the statistical tests indicated no significant difference in medians of likes for pho-

tos (0.08) and videos (0.19). Additionally, it was seen for videos that the distributions of likes had no statistical significance (0.04), as well as the means of comments. For albums, only the means of comments indicated no significant impact between the groups younger and older. Still, the tests do not affect the made observations.

7 Categorical Analysis

The following chapter analyses how users engage among different categories of uploaders. As mentioned in Section 3.2, the chosen categories are actors, athletes, brands, musicians, and others. Figure 7.1 presents the average temporal impact of each category. As seen in both figures, brands receive the least amount of interactions overall. Additionally, it can be observed that brands receive proportionally fewer interactions at a later stage compared to the other categories that follow a trend similar to each other. It should be noted that brands have the highest number of uploaded posts compared to the other categories. Hence, a similar pattern as presented in Figure 5.5 occurs in which users with many prior uploaded posts receive fewer interactions at a later stage.

Further, the category brands include accounts that belong to, for example, high- and low-end brands, organisations, sports clubs, and magazines. The result presented in Table 3.1, in which brands receive the most likes, comments, and views per user among the categories, may be explained by the variation of accounts that are included in the category brands. Thus, some accounts may receive more engagement than others. A distinguishment between these types of accounts might yield a different result, but such distinguishment could not be performed for the number of accounts collected. Still, the result presented in Figure 5.5 may indicate why it is common for brands to use other users on social media platforms such as influencers and celebrities to promote their products or services.

Moreover, it can be observed in Figure 7.1a that athletes receive the highest number of likes, followed by users that belong to the category musicians. On the contrary, it can be ob-

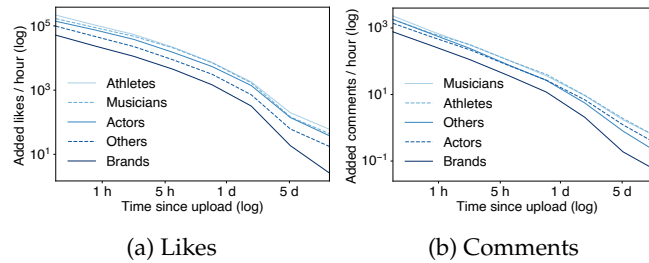


Figure 7.1: Temporal impact of categories.

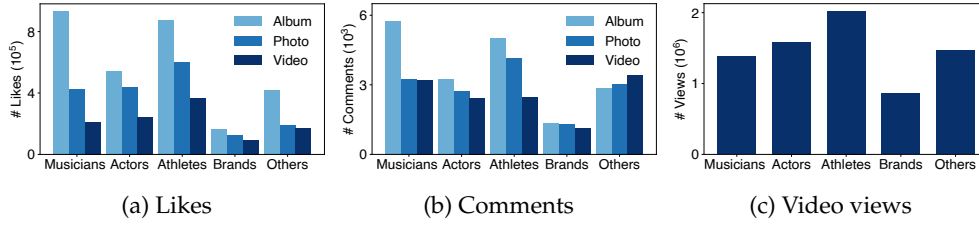


Figure 7.2: Average impact of post types for each category.

served in Figure 7.1b that musicians receive more comments than athletes. Note that athletes have the least number of uploaded posts which might affect the results.

The Kruskal-Wallis test indicated statistical significance between the categories. For Dunn's test, athletes and others had no significant difference in comments (with a p-value of $4.10 \cdot 10^{-2}$). However, the Mann-Whitney test indicated no significant difference in comments between musicians and actors (with a p-value of 0.34). As the results are contradictory, the null hypothesis cannot be rejected nor retained. The KS test indicated statistical significance between all pairwise comparisons while bootstrapping showed no significant difference in means of comments between musicians and athletes, as well as others and actors. The results are reasonable considering that the category ranking for comments differs from likes.

The following section will take a closer look at how users engage differently among the three post types for each category of users.

7.1 Post Types

To compare how users engage differently for each post type among the categories, Figure 7.2 presents the average impact of post types for each category of uploaders. As observed in Figure 7.2a, brands and others receive the least number of likes for all post types. On the other hand, athletes and actors receive the highest number of likes on photos and videos, which is negatively correlated with the number of uploaded posts for each category as seen in Table 3.1. For albums, musicians instead receive the highest amount of likes which may indicate that albums are good for musicians to promote the release of new songs or concerts.

Further, it can be observed in Figure 7.2b that albums receive the highest number of comments, followed by photos for all categories but others, in which the contrary is illustrated. The category others was further observed to gain insights into why the opposite trend occurs. It was seen that models and influencers received the highest number of comments on videos among the category others. Although, the result may be due to models and influencers being the majority of professions in the category others. Albums follow the same trend in Figure 7.2b as observed in Figure 7.2a, while photos and videos follow different trends. Other than the category others, it can be seen for both post types that musicians receive more comments than expected by the like count. This may indicate that musicians' followers are more faithful compared to the other categories as their interactions, in the form of comments, involve more effort than liking a post.

Figure 7.2c presents the average number of video views considering the categories of uploaders. As seen in the figure, the observed trend is similar to Figure 7.2a in which athletes receive the highest amount of interactions and brands the lowest amount. What differentiates the figures is the large number of views for the category others, which may be influenced by a large number of comments. However, compared to Figure 7.2b, the category others, surprisingly, do not receive the highest number of views. The results may indicate that those users that comment may leave multiple comments, thus the category others receive more views per like compared to the other categories.

The Kruskal-Wallis test indicated statistical significance between the post types for all categories of uploaders. However, Dunn's test indicated no significant difference in likes

between photos and videos for the category others (p -value of $3.58 \cdot 10^{-2}$), in comments between photos and videos for musicians and athletes ($7.43 \cdot 10^{-1}$ and $5.95 \cdot 10^{-2}$), and in comments between albums and photos for the category others ($5.54 \cdot 10^{-3}$). The Mann-Whitney test made similar indications for all pairwise comparisons except for the category others in which a significant difference in comments was shown between albums and photos (p -values of 0.02, 0.71, 0.05, and 0.003 respectively). For the means, musicians and actors had no significant difference in comments between photos and videos, while athletes, brands, and others had no difference in comments between albums and photos. Some occurrences of similarity may be explained by the difference in distributions. The KS test indicated no significant difference in comments between photos and videos for musicians, actors, and athletes (0.27, 0.02, 0.03). Still, the results do not affect a majority of the observations made.



8 Discussion

The following chapter discusses the methodology and results, as well as the ethical and societal aspects of the analysis.

8.1 Results

The following section discusses the results of the conducted analysis with respect to the presented research questions. It should be noted that this analysis was performed on some of the most followed users on Instagram. Hence, the results may not be applicable to all users on Instagram.

8.1.1 How Do Different Characteristics Influence Users to Engage in Instagram Posts over Time?

From the initial analysis of post types in Chapter 5, it was observed that all post types had a heavy right-tailed distribution. Further, album posts received the highest number of received interactions, followed by photo posts. However, more photos were uploaded compared to albums. The results with respect to the post type of the temporal or average impacts of different characteristics are summarised below.

Uploader Characteristics It was found that a large social network was a great contributor to users' engagement. Additionally, the number of uploaded posts 7 days prior had a negative correlation with the number of acquired interactions for posts. The results are in line with other works analysing uploaders' characteristics on Instagram [1, 20, 21, 45]. Further, it was observed how users who on average upload 3 posts a day or more, receive proportionally less interaction at a later stage compared to users who upload less than 3 posts a day on average.

Post Characteristics For characteristics related to the post, it was observed that short descriptions (containing few words) were related to high engagement among users. The result may however be due to the most popular users using fewer words in their descriptions compared to other users, as observed in the comparison between the top-100 and top-1000 users in Section 6.1. Further, it was observed that posts with long descriptions received more comments compared to intermediate descriptions.

Incorporating hashtags and mentions in posts' descriptions were seen to reduce users' engagement. As previously discussed, the result contradicts works from previous years that studied hashtag behaviours on Instagram [12, 21], which may indicate that a new trend has emerged across all users or only the most popular ones.

Media Characteristics Features related to the media but not the content of media files showed that longer albums receive more likes. The result may be due to them having a higher possibility of being uploaded by popular users. Additionally, it was seen that shorter video posts received more interactions initially and may be more likely to be uploaded by popular users, while longer videos received more interactions at a later stage.

For features related to what was present in posts' media files, it could be observed that a lack of people present in media files generally produced less engagement for photos and videos. For albums, media files with more than 2 people present generally received fewer interactions than media files in which no people were present. Additionally, the number of people present in media files had a negative correlation for album posts but did not matter for photo posts.

For posts in which people were detected, it was observed that posts that included women received more interactions than posts that only had men. People with a predicted age of at least 35 years old received more interactions generally than people below that age for albums and videos. Lastly, neutral expressions on faces received the highest number of interactions which was followed by happy expressions for albums and sad expressions for photos.

8.1.2 What are the Most Important Characteristics that Attract Users' Interactions?

Foremost, it can be observed in the correlation matrices in Figure 4.2 and other previous figures that uploaders' characteristics seemed to have the greatest impact on user engagement. These results are in line with previous works analysing users' interactions on Instagram that found that the size of social networks had the greatest impact [1, 20, 21]. Additionally, it was observed that the three types of interactions: likes, comments, and views had a strong correlation to each other. Whether the number of likes and comments is dependent on the number of views or vice versa, could however not be further analysed as the temporal view count could not be extracted.

Between posts and media characteristics, it seemed as if media characteristics had less statistical support for the observed differences compared to post characteristics. Therefore, post characteristics are deemed more important in attracting users' interactions than media characteristics.

8.1.3 How do Users' Engagements with Instagram Posts Change for Different Categories of Uploaders?

As observed in Figure 7.1, the different types of categories: actors, athletes, brands, musicians, and others, all seemed to have a similar trend of receiving likes over time except for brands. The category brands received the least amount of total interactions compared to the other categories and were also seen to receive proportionally fewer likes at a later stage after upload. The large number of uploaded posts might have an impact on the results, as well as the variation of accounts included in the category brands. However, the results also explain why influencers are commonly used to promote brands and their products on various social media platforms.

Further, it was noticed that differences in user engagement among the categories could be observed between the types of posts. Albums received the highest number of interactions, followed by photos, for all categories but others in which the contrary was observed for the number of received comments. Video posts thus received the highest number of comments

for the category others which was reflected in the number of acquired video views but not likes. As indicated by the results, users that comment might therefore leave multiple comments on posts uploaded by the category others. Musicians, on the other hand, were seen to receive the highest number of interactions on albums and more comments on photos and videos than expected by the number of likes and views received. The results show that Instagram is a good platform for musicians to promote their music, especially using album posts, and it may also indicate that their followers are more faithful compared to the other categories.

8.2 Methodology

The following section discusses the analysed characteristics and limitations of the selected models. Additionally, the section criticises the used sources.

8.2.1 Characteristics

A number of characteristics of posts and uploaders were taken into consideration when analysing the temporal engagement of users on Instagram. A majority of the characteristics were gathered by CrowdTangle. However, a few were extracted manually, one of which was the number of words in posts' descriptions. It was extracted by considering the number of words, digits, hashtags, mentions, and emojis in each description. However, different languages and their semantics were not considered during the analysis which could have affected the results.

Some characteristics used in previous works could have been interesting to analyse such as users' social networks [12, 21]. There were also characteristics among the statistical data that was available, but not covered in the analysis. One such characteristic was whether a post contained sponsored content or not. The feature was not sufficiently covered as only a small portion of posts were sponsored. Additionally, many posts on social media platforms have been shown to contain sponsored content without or with only a subtle indication of it [24]. Partly it may deceive customers to buy products but more prominent to this analysis, sponsored posts have been shown to receive less user engagement compared to non-sponsored posts [30]. Therefore, it would have been valuable to consider features such as hashtags and mentions in the description along with logos in the media to predict whether a post was sponsored or not.

When analysing the media files of posts, only facial attributes and the number of people present in posts were accounted for. Less general features such as if a dog was present in the image or a person wore glasses seemed too specific and not properly studied to narrow the scope further. Other general features were considered but not pursued due to interfering images or a lack of existing pre-trained state-of-the-art models. For example, previous work showed that text posters received less engagement among users on Instagram [52]. This feature was not pursued as text on jerseys or buildings in images could confuse the detection of text posters. However, incorporating more features would have given a more detailed analysis. Similarly, more characteristics of posts or uploaders may have had an impact on the analysis.

8.2.2 Detection Models

Instead of developing detection models from scratch, current state-of-the-art models were implemented from different frameworks. The limitation was due to time restrictions and since the analysis was the central part of this work. Additionally, developing your own model does not necessarily yield a better result. The models used in this analysis to extract the number of people in each image and the facial attributes: age, gender, and emotion, were carefully selected. However, it should be noted that the chosen models do not have perfect

accuracy and detection models in general have shown to be biased against certain subgroups [41]. Either the model is unable to generalise across different subgroups or certain features in the data are underrepresented due to uniform or inaccessible data. For example, only a limited quantity of pictures containing children are available due to legal restrictions [35]. Still, the chosen models are currently state-of-the-art. Models more fitted to the collected data could however yield better results.

For both models, the thresholds were calculated by taking the average predicted accuracy over all detected attributes or people. This limitation was made due to time restrictions, thus it is time-consuming to manually annotate such image features and may be difficult to find a classification dataset that could correctly represent the collected dataset. The models might have improved their results if additional metrics such as precision, recall, and F1-score had been used to decide the thresholds. However, it was deemed that the limitation would not have a significant impact on the final results of the analysis.

Further, ages between 0 and 100 years old could be predicted for people detected by the chosen model. To diminish the number of misclassified ages, it was determined to divide detected people into the group younger (0-34) or older (35+). Thus, a majority of users were predicted to be in their late twenties or early thirties. Other divisions of ages could have been made. One such example is dividing users into three groups; under 25, 25-35, and over 35. However, it was noticed that the number of people in each division could never be both evenly distributed and reasonable due to the immense amount of people detected to be around 30 years old. Hence, the two chosen groups were seen as a better alternative although it may give a limited result.

The use of both models in this analysis is thoroughly demonstrated. Although the dataset has limited replicability due to confidentiality, the models could be applied to similar or extended datasets to further analyse user engagement on social media platforms.

8.2.3 Source Criticism

Efforts have been made to find relevant and trustworthy sources related to the analysis and methodology. Primarily, papers from highly ranked proceedings and journals have been cited. These papers were still read with caution to prevent any biased beliefs. Other sources include fundamental concepts or models and websites have also been used to reference repositories or to add trivial information.

8.3 The Work in a Wider Context

In this work, the societal aspects of Instagram were analysed regarding user engagement. The results can be profitable for both individual users and brands as they can expand their social network and gain more engagement on their posts. Brands can utilise the growth in popularity by marketing their products and additionally, use individual users for marketing. Individual users do, however, have an ethical responsibility in what products and brands they promote due to their influence among consumers. For example, important factors are the brands' ideology, conditions in the workplace, or view on sustainability.

The confidentiality of the collected data has previously been mentioned as it affects the replicability of the analysis. There could be ethical implications if confidential information, such as the collected identities, were to be exposed. Anonymisation of the collected data would not be possible as there would be a high probability of re-identification. The dataset used for the analysis is therefore not publicly available.



9 Conclusion

The following chapter concludes the result of the analysis and presents possible future work.

9.1 Results

This work has analysed different characteristics that impact users' temporal engagement among short-lived posts on the social media platform Instagram. Multiple characteristics related to posts, uploaders, and media files, consisting of photos and videos, were further observed with respect to the post type: album, photo, or video. It was observed that album posts received the highest amount of interactions, followed by photo posts. The most important characteristic that attracted users' interactions was related to the uploader and included their social network size and uploading rate. Statistical tests deemed post characteristics to be more important in attracting interactions than media characteristics. Further, different categories of users were analysed with respect to the post type. Compared to other influencer groups, it was found that brands and other types of organisations received fewer interactions and musicians tended to have more loyal followers on Instagram.

9.2 Future Work

Apart from performing the analysis on a larger dataset or including other characteristics that were not covered in this analysis such as the number of following, views, or shares, a significant impact of the analysis would be to include a greater diversity of users. The results presented in this analysis were performed on a dataset containing some of the most followed users on Instagram. Therefore, the observed trends might not apply to private users with small networks. Further, a more in-depth analysis of the media files could be performed to take into account other features that are more or less general. For people, such features could be a person's attractiveness, actions, or the relation between individuals in an image. Of the remaining content, such features could be if the environment is indoor, outdoor, or related to any sports activities. The categorical analysis could also be extended further to observe how users engage differently across categories of uploaders. Lastly, an analytical model could be developed to predict the temporal dynamics of user engagement dependent on the features used in this analysis or additional features.



Bibliography

- [1] A. Bakhshi, D. Shamma, and E. Gilbert. “Faces engage us: Photos with faces attract more likes and comments on instagram”. In: *Proceedings of the SIGCHI Conference on human factors in computing systems*. 2014, pp. 965–974.
- [2] Y. Borghol, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti. “The untold story of the clones: Content-agnostic factors that impact YouTube video popularity”. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge discovery and data mining*. Beijing, China, 2012, pp. 1186–1194.
- [3] Y. Borghol, S. Mitra, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti. “Characterizing and modelling popularity of user-generated videos”. In: *Performance Evaluation* 68.11 (2011), pp. 1037–1055.
- [4] S. Carta, AS. Podda, DR. Recupero, R. Saia, and G. Usai. “Popularity prediction of instagram posts”. In: *Information* 11.9 (2020), p. 453.
- [5] ChatGPT. <https://openai.com/blog/chatgpt>. Accessed: 2023-04-29.
- [6] L. Cheng, R. Guo, Y. Silva, D. Hall, and H. Liu. “Hierarchical attention networks for cyberbullying detection on the instagram social network”. In: *Proceedings of the SIAM international conference on data mining*. 2019, pp. 235–243.
- [7] William Jay Conover. *Practical nonparametric statistics*. Vol. 350. John Wiley & Sons, 1999.
- [8] CrowdTangle. <https://www.crowdtangle.com/>. Accessed: 2023-03-01.
- [9] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou. “Retinaface: Single-shot multi-level face localisation in the wild”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, pp. 5203–5212.
- [10] Olive Jean Dunn. “Multiple comparisons using rank sums”. In: *Technometrics* 6.3 (1964), pp. 241–252.
- [11] B. Efron. *Bootstrap methods: another look at the jackknife*. Springer, 1992.
- [12] E. Ferrara, R. Interdonato, and A. Tagarelli. “Online popularity and topical interests through the lens of instagram”. In: *Proceedings of the ACM Conference on Hypertext and social media*. 2014, pp. 24–34.
- [13] K. Garimella and R. West. “Evolution of Retweet Rates in Twitter User Careers: Analysis and Model”. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 15. 2021, pp. 1064–1068.

- [14] M. Gayberi and S.G. Oguducu. "Popularity prediction of posts in social networks based on user, post and image features". In: *Proceedings of the International Conference on Management of Digital EcoSystems*. 2019, pp. 9–15.
- [15] R. Girshick. "Fast R-CNN". In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1440–1448.
- [16] K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition". In: *Proceedings of the IEEE Conference on computer vision and pattern recognition*. 2016, pp. 770–778.
- [17] J. Holmström, D. Jonsson, F. Polbratt, O. Nilsson, L. Lundström, S. Ragnarsson, et al. "Do we read what we share? analyzing the click dynamic of news articles shared on Twitter". In: *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 2019, pp. 420–425.
- [18] G. Huang, M. Mattar, T. Berg, and E. Learned-Miller. "Labeled faces in the wild: A database for studying face recognition in unconstrained environments". In: *Workshop on faces in Real-Life Images: detection, alignment, and recognition*. 2008.
- [19] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, et al. "Speed/accuracy trade-offs for modern convolutional object detectors". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 7310–7311.
- [20] J. Huang, C. Wang, M. Su, Q. Dai, and MZA. Bhuiyan. "Inspecting influences on likes and comments of photos in instagram". In: *Proceedings of the IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*. 2018, pp. 938–945.
- [21] JY. Jang, K. Han, and D. Lee. "No reciprocity in" liking" photos: analyzing like activities in instagram". In: *Proceedings of the ACM conference on hypertext and social media*. 2015, pp. 273–282.
- [22] JY. Jang, K. Han, D. Lee, H. Jia, and P. Shih. "Teens engage more with fewer photos: temporal and comparative analysis on behaviors in instagram". In: *Proceedings of the ACM Conference on hypertext and social media*. 2016, pp. 71–81.
- [23] W.H. Kruskal and W.A. Wallis. "Use of ranks in one-criterion variance analysis". In: *Journal of the American statistical Association* 47.260 (1952), pp. 583–621.
- [24] Xueting Liao, Danyang Zheng, Yubao Wu, and Xiaojun Cao. "Should We Trust Influencers on Social Networks? On Instagram Sponsored Post Analysis". In: *Proceedings of the International Conference on Computer Communications and Networks*. 2021, pp. 1–8.
- [25] TY. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. "Feature pyramid networks for object detection". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 2117–2125.
- [26] TY. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. "Focal loss for dense object detection". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2980–2988.
- [27] H.B. Mann and D.R. Whitney. "On a test of whether one of two random variables is stochastically larger than the other". In: *The annals of mathematical statistics* (1947), pp. 50–60.
- [28] M. Mazloom, R. Rietveld, S. Rudinac, M. Worring, and W. Van Dolen. "Multimodal popularity prediction of brand-related social media posts". In: *Proceedings of the ACM International Conference on Multimedia*. 2016, pp. 197–201.

- [29] A. Mohammadinodooshan and N. Carlsson. "Effects of Political Bias and Reliability on Temporal User Engagement with News Articles Shared on Facebook". In: *Proceedings of the International Conference on Passive and Active Network Measurement*. Springer. 2023, pp. 160–187.
- [30] L.M. de Oliveira and O. Goussevskaia. "Sponsored content and user engagement dynamics on Instagram". In: *Proceedings of the ACM symposium on applied computing*. 2020, pp. 1835–1842.
- [31] O. Parkhi, A. Vedaldi, and A. Zisserman. "Deep face recognition". In: *British Machine Vision Association* (2015).
- [32] K. Pearson. "VII. Note on regression and inheritance in the case of two parents". In: *Proceedings of the royal society of London* 58.347-352 (1895), pp. 240–242.
- [33] V. Pham, C. Pham, and T. Dang. "Road damage detection and classification with detectron2 and faster R-CNN". In: *Proceedings of the IEEE International Conference on Big Data*. 2020, pp. 5592–5601.
- [34] L. Picazo-Sánchez, R. Domínguez-Martín, and D. García-Marín. "Health Promotion on Instagram: Descriptive–Correlational Study and Predictive Factors of Influencers' Content". In: *International Journal of Environmental Research and Public Health* 19.23 (2022), p. 15817.
- [35] I.D. Raji, T. Gebru, M. Mitchell, J. Buolamwini, J. Lee, and E. Denton. "Saving face: Investigating the ethical concerns of facial recognition auditing". In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 2020, pp. 145–151.
- [36] S. Ren, Kaiming He, R. Girshick, and J. Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks". In: *Advances in neural information processing systems* 28 (2015).
- [37] F. Sabate, J. Berbegal-Mirabent, A. Cañabate, and P. Lebherz. "Factors influencing popularity of branded content in Facebook fan pages". In: *European management journal* 32.6 (2014), pp. 1001–1011.
- [38] S. Serengil. *DeepFace: A Lightweight Face Recognition and Facial Attribute Analysis (Age, Gender, Emotion and Race) Library for Python*. <https://github.com/serengil/deepface>. Accessed 2023-05-22.
- [39] S. Serengil and A. Ozpinar. "Hyperextended lightface: A facial attribute analysis framework". In: *Proceedings of the International Conference on Engineering and Emerging Technologies*. 2021, pp. 1–4.
- [40] S. Serengil and A. Ozpinar. "Lightface: A hybrid deep face recognition framework". In: *Proceedings of the Innovations in Intelligent Systems and Applications Conference*. 2020, pp. 1–5.
- [41] R. Singh, P. Majumdar, S. Mittal, and M. Vatsa. "Anatomizing bias in facial analysis". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. 11. 2022, pp. 12351–12358.
- [42] *Starngage*. <https://starngage.com/plus/en-us>. Accessed: 2023-03-31.
- [43] W. Tafesse and M. Dayan. "Content creators' participation in the creator economy: Examining the effect of creators' content sharing frequency on user engagement behavior on digital platforms". In: *Journal of Retailing and Consumer Services* 73 (2023), p. 103357.
- [44] E. Thorgren, A. Mohammadinodooshan, and N. Carlsson. "Temporal Dynamics of User Engagement on Instagram: A Comparative Analysis of Album, Photo, and Video Interactions". In: *Proceedings of the ACM Web Science Conference (WebSci)*. 2024.

- [45] L. Vassio, M. Garetto, C. Chiasserini, and E. Leonardi. "Temporal dynamics of posts and user engagement of influencers on facebook and instagram". In: *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 2021, pp. 129–133.
- [46] C. Wallner, S. Kruber, SO. Adebayo, O. Ayandele, H. Namatame, TT. Olonisakin, et al. "Interethnic Influencing Factors Regarding Buttocks Body Image in Women from Nigeria, Germany, USA and Japan". In: *International Journal of Environmental Research and Public Health* 19.20 (2022), p. 13212.
- [47] K. Wang, P. Wang, X. Chen, Q. Huang, Z. Mao, and Y. Zhang. "A feature generalization framework for social media popularity prediction". In: *Proceedings of the ACM International Conference on Multimedia*. 2020, pp. 4570–4574.
- [48] Wikipedia. *Instagram*. <https://en.wikipedia.org/wiki/Instagram>. Accessed 2023-05-22.
- [49] Lior Wolf, Tal Hassner, and Itay Maoz. "Face recognition in unconstrained videos with matched background similarity". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2011, pp. 529–534.
- [50] Y. Wu, A. Kirillov, F. Massa, WY. Lo, and R. Girshick. *Detectron2*. <https://github.com/facebookresearch/detectron2>. 2019.
- [51] X. Yang and J. Luo. "Tracking illicit drug dealing and abuse on Instagram using multi-modal analysis". In: *Proceedings of the ACM Transactions on Intelligent Systems and Technology* 8.4 (2017), pp. 1–15.
- [52] Z. Zhang, T. Chen, Z. Zhou, J. Li, and J. Luo. "How to become Instagram famous: Post popularity prediction with dual-attention". In: *Proceedings of the IEEE International Conference on big data*. 2018, pp. 2383–2392.
- [53] A. Zohourian, H. Sajedi, and A. Yavary. "Popularity prediction of images and videos on Instagram". In: *Proceedings of the International Conference on Web Research*. 2018, pp. 111–117.

A

Top-100 Users Analysis

Category	Number of users	Number of posts				Engagement per user		
		Album	Photo	Video	Total	Likes	Comments	Views
Musicians	32	2,464	2,540	1,844	6,848	303,640,468	2,300,294	174,409,359
Actors	24	1,021	1,111	587	2,719	136,796,922	707,913	85,061,258
Brands	17	12,302	13,424	4,253	29,979	788,452,881	3,336,734	424,209,350
Others	15	2,239	1,355	752	4,346	308,158,882	1,280,570	101,361,249
Athletes	12	515	903	227	1,645	395,710,932	2,931,052	135,957,420
Total	100	18,541	19,333	7,663	45,537	1,932,760,085	10,556,563	920,998,636

Table A.1: Summary of the number of users, posts, and engagement for each category of users.

Table A.1 presents the number of users and posts in each category, as well as the total engagement divided by the number of users in each category for the top-100 most followed users among the collected data.

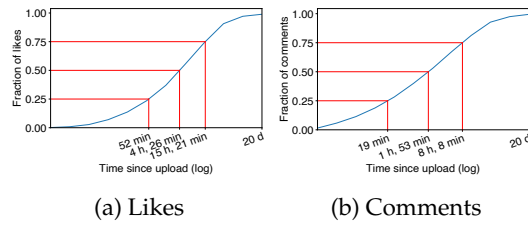


Figure A.1: Cumulative fraction of interactions (d: days, h: hours, min: minutes).

Figure A.1 presents the average cumulative fraction of interactions for all users over the first 20 days after the posts were uploaded. The figure illustrates the average time posts acquire 25%, 50%, respectively 75% of the total interactions in the 20-day interval. The time thresholds were calculated by taking the average time each post obtained the respective percentage of the total interactions. As observed in the figure, Instagram posts are short-lived. The majority of likes are received during the first 24 hours after upload and the majority of comments are received during the first 12 hours.

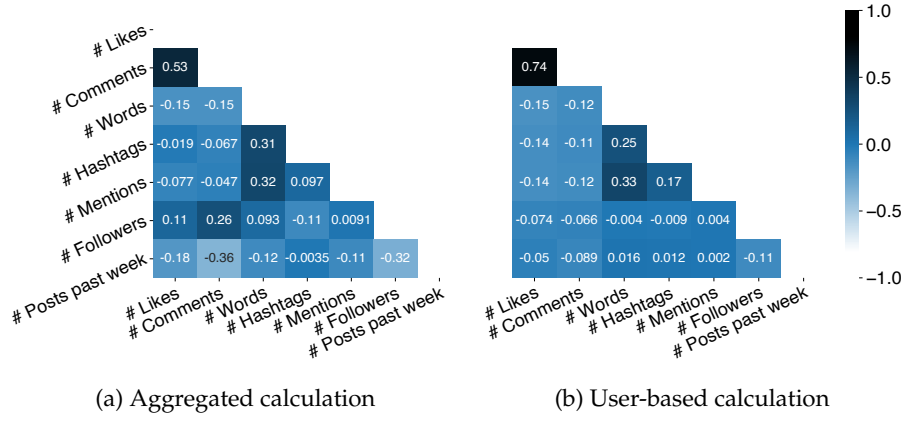


Figure A.2: Correlation matrices for the statistical data.

Figure A.2 presents the correlation matrices for most of the variables calculated using Pearson's population correlation coefficient. Figure A.2a presents an aggregated calculation over all posts, while Figure A.2b presents the average correlation seen across the users. To ensure a linear relationship between variables, their characteristics and cumulative distribution functions were accounted for when deciding their transformation. The scale of each variable is presented in Table 4.1. As observed in Figure A.2, the correlation among the variables is moderate to weak. The strongest correlation exists between likes and comments, which is followed by the correlation between user interactions and the characteristics of uploaders. The weak and moderate correlation may suggest that the relation between the variables is other than linear, or the calculated correlation is affected by outliers. Compared to Figure A.2b, the correlation among the aggregated variables, in regards to users' interactions, is strongest for uploaders' characteristics, similar or weaker for posts' characteristics, and weaker between the two types of interactions.

For videos, the correlation between the number of views and likes is 0.91, comments 0.75, followers 0.32, and posts past week -0.25. However, even if the correlation is stronger compared to Figure A.2a, the correlation between the number of likes and comments for videos is 0.72. This shows that the stronger correlation between views and other variables compared to previously observed correlations, is due to the post type. Further, the strong correlation between the number of views and likes may indicate that the number of received likes can be used as a proxy for video views.

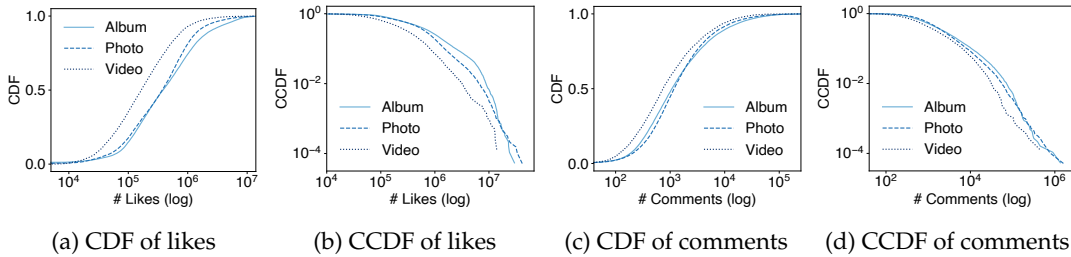


Figure A.3: Distributions of total interactions

Figure A.3 presents the CDFs and CCDFs for the total number of likes and comments with respect to the post types. Note that both axes in Figure A.3b and Figure A.3d are of logarithmic scale. As observed in the figures, a majority of albums receive more interactions but the distribution is similar to photos. On the other hand, most videos receive the least number of interactions. As illustrated by the CCDFs, all post types have a heavy right-tail

distribution. Additionally, photos seem to have a heavier tail compared to albums as seen in Figure A.3b, which shows that photos include some additional outliers with many likes.

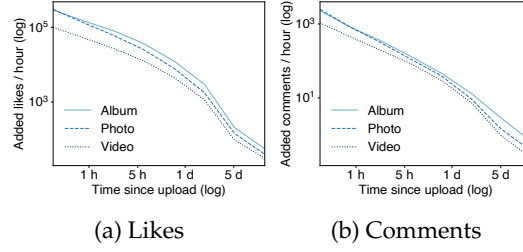


Figure A.4: Temporal impact of post types.

Figure A.4 presents the average number of added interactions acquired, divided by the number of hours passed since upload. As illustrated in the figures, users follow a similar pattern among the three post types for both types of interactions, even if the number of interactions is not similar for the two figures. As observed, albums receive the highest amount of added interactions, while videos receive the least amount of added interactions during the 20-day time period. Photos receive nearly as many interactions initially as albums but later decline. Although albums receive more total interactions compared to photos, more photos are published among the top-100 collected users as illustrated in Table A.1.

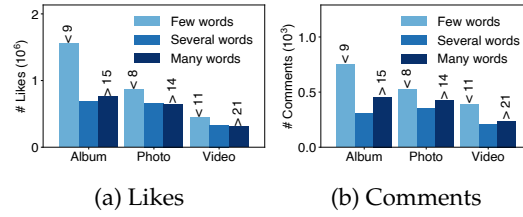


Figure A.5: Average impact of words in posts' descriptions.

Figure A.5 presents the average number of received interactions dependent on the number of words in posts' descriptions. The categories are determined to contain roughly the same amounts of posts. As illustrated in the figures, videos use on average more words in their descriptions compared to albums and photos. For all post types, posts with few words in their descriptions receive the highest number of interactions. Furthermore, posts that contain many words receive more comments than posts containing several words as seen in Figure A.5b. The trend can also be observed in Figure A.5a for albums.

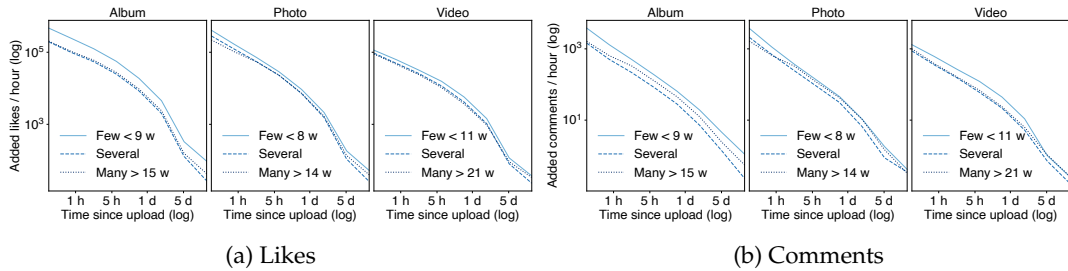


Figure A.6: Temporal impact of words in posts' descriptions (w: words).

Figure A.6 presents the average temporal impact of the number of words in posts' descriptions. The categories are determined to contain roughly the same amounts of posts. As observed in the figures, videos use on average more words in their descriptions compared to albums and photos. For all post types, posts with few words in their descriptions receive the

highest number of added interactions. Additionally, posts that contain many words receive more comments than posts containing several words as illustrated in Figure A.6b. The trend can also be observed in Figure A.5a for albums. Furthermore, it can be observed in Figure A.5 that posts which contain many words receive more added likes later on in the 20-day period, compared to posts with several words in their descriptions. This may correlate to the fact that posts with many words in their description receive more comments and are therefore more active at the later stage than posts with several words.

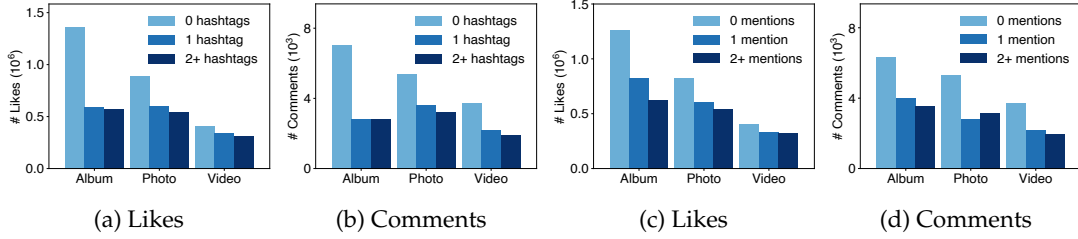


Figure A.7: Average impact of hashtags and mentions in posts' descriptions.

Figure A.7 present the average amount of interactions that albums, photos, and videos receive based on the number of hashtags or mentions included in posts' descriptions. As illustrated by the figures, posts that receive the most interactions are posts that do not use any hashtags or mentions. Note that a correlation among posts that included different amounts of hashtags or mentions in their description could not be found other than receiving fewer interactions than posts that did not include hashtags or mentioned other users.

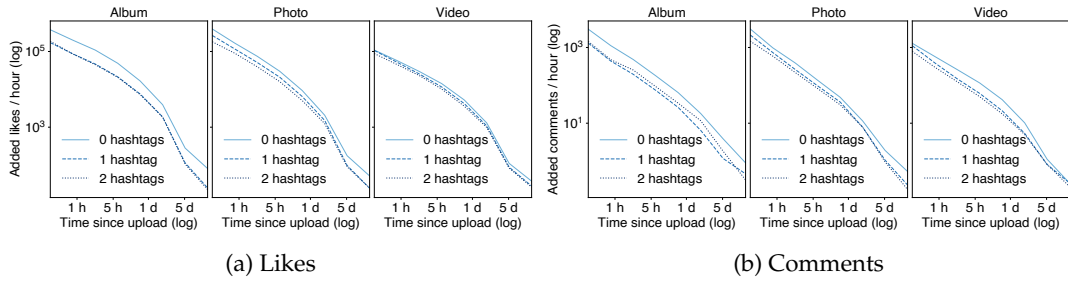


Figure A.8: Temporal impact of hashtags in posts' descriptions).

Figure A.8 presents the average temporal impact of the number of hashtags in the posts' descriptions. As illustrated by the two figures, posts that receive the most added interactions are posts that do not use any hashtags. Note that a correlation among posts that included different amounts of hashtags in their description could not be found other than receiving fewer interactions than posts that did not include hashtags.

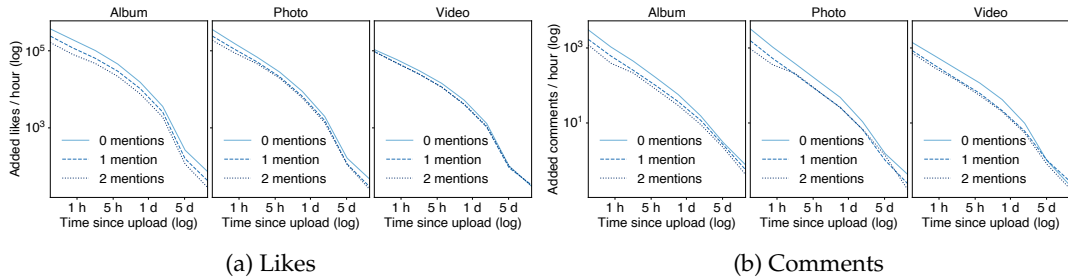


Figure A.9: Temporal impact of mentions in posts' descriptions.

Figure A.9 presents the average temporal impact of the number of mentions in the posts' descriptions. As illustrated by the two figures, posts that receive the most added interactions

are posts that do not use any mentions. Note that a correlation among posts that included different amounts of mentions in their description could not be found other than receiving fewer interactions than posts that did not include mentions.

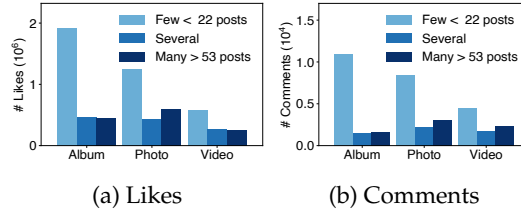


Figure A.10: Average impact of uploaded posts 7 days prior.

Figure A.10 presents the average amount of received interactions based on the number of uploaded posts up to 7 days prior to the upload of each post. The three categories are determined to contain roughly the same number of posts. As observed in the two figures, users that upload fewer posts per week (less than 3 posts a day on average) receive more interactions. Furthermore, users that upload many posts receive more comments than users that upload several posts as illustrated in Figure A.10b. This is, however, not reflected in the number of likes acquired by albums and video posts where the contrary is observed.

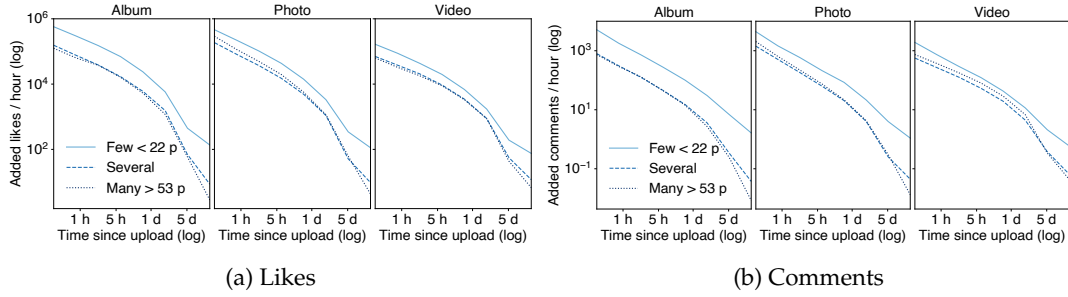


Figure A.11: Temporal impact of uploaded posts 7 days prior (p: posts).

Figure A.11 presents the average temporal impact of the number of uploaded posts up to 7 days before the upload of each post. The three categories are determined to contain roughly the same number of posts. As illustrated by the two figures, users that upload fewer posts per week (less than 3 posts a day) receive more interactions on their posts compared to users with more prior uploaded posts. For photo posts, it can be observed that users who upload many posts receive more interactions initially compared to users that upload several posts a week. At a later stage, a decrease in the number of interactions can be observed among users with many prior uploaded posts. The decrease occurs for all post types and interactions but is more noticeable among albums and photos. The decrease may suggest that the same amount of added interactions are acquired, but distributed among the large number of posts.

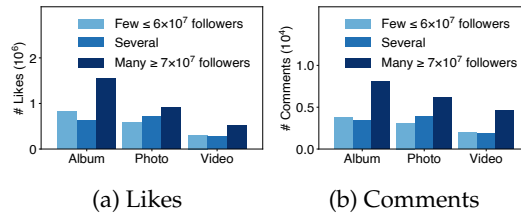


Figure A.12: Average impact of followers.

Figure A.12 presents the average amount of interactions that albums, photos, and videos receive based on how many followers the uploader has at the time of upload. The three

categories are determined to contain roughly the same number of posts. As observed in the two figures, users with many followers receive the most interactions on all their posts. For albums and videos, it seems as if users with few followers receive more interactions than users with several followers. On the contrary, users with few followers receive the lowest number of interactions for photos.

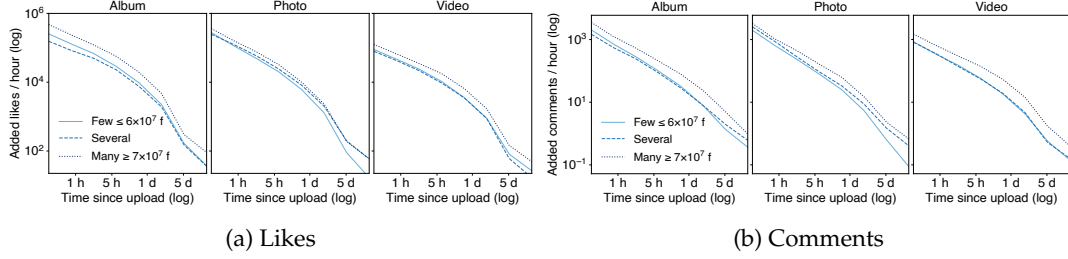


Figure A.13: Temporal impact of followers (f: followers).

Figure A.13 presents the average temporal impact of how many followers the uploader has at the time of upload. The three categories are determined to contain roughly the same number of posts. As observed in the two figures, users with many followers receive the most added interactions on all their posts. For albums, users with few followers receive more interactions initially compared to users with several followers. However, at a later stage, the number of added interactions declines faster and users with fewer followers receive a similar amount or fewer added interactions. The latter trend can also be observed for photos. However, as opposed to albums, users with few followers acquire the least amount of received likes during the 20-day period compared to users with more followers. For videos, no noticeable pattern for the number of interactions between few and several followers can be found.

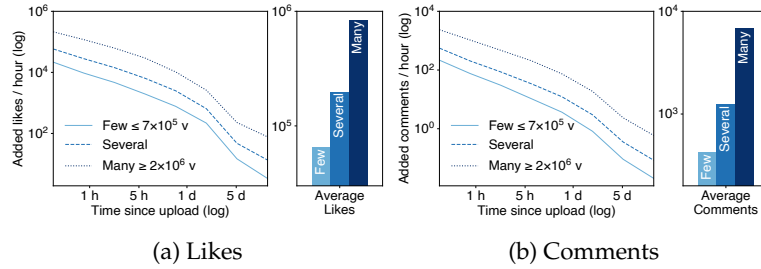


Figure A.14: Temporal and average impact of video views (v: views).

Figure A.14 presents the temporal and average impact that the number of video views has on interactions. The three categories are determined to contain roughly the same number of posts. As observed in the two figures, the more views videos have, the more interactions they receive. However, the total number of views does not seem to affect the rate at which videos receive interactions. Instead, each video seems to have a similar rate of attracting likes, despite how many views each video has. Note that Figure A.14 presents an average result and temporal views can yield different results.

Figure A.15 presents the average number of interactions considering the impact of post and uploader characteristics. The three categories for each characteristic have the same limits as the previous figures. As observed in Figure 5.9a, videos with few words in their descriptions receive the highest amount of views, followed by videos containing several words. The trend can also be found among likes in Figure A.5a, but not among comments as videos with several words in their description receive the least amount of comments. For all three types of user engagements, posts without hashtags and mentions in their descriptions receive the highest amount of engagement. Furthermore, all interactions have a similar trend in regard

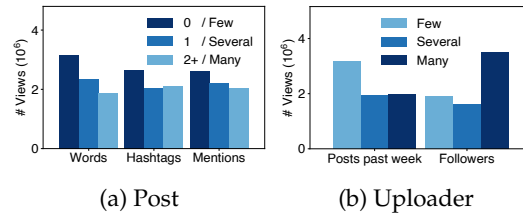


Figure A.15: Average impact for video views.

to the number of followers for the user. However, for the number of uploaded posts 7 days prior, the trend for views is more similar to comments than likes. Hence, posts with several prior uploaded posts receive the least amount of views.

B

Additional Post Type Analysis

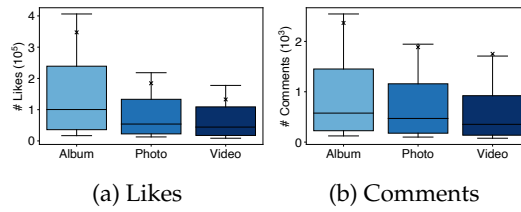


Figure B.1: Average impact of post type

Figure B.1 presents the average impact of posts' type in which the bottom and top of the box show the 30th and 70th percentile respectively and the median line and average cross are also illustrated. The whiskers show the 20th and 80th percentile respectively. As seen in both figures, album posts receive the highest number of likes, followed by photo posts. The top whiskers could indicate that more popular users more commonly upload albums. However, the bottom whiskers indicate that all three post types are common among all users.

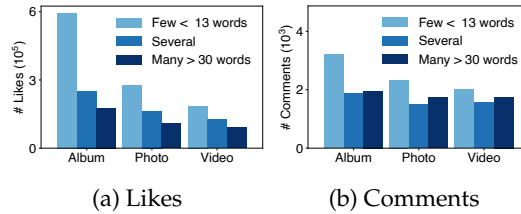


Figure B.2: Average impact of words in posts' descriptions.

Figure B.2 presents the average number of interactions received with respect to the number of words in each post. The columns are calculated to contain roughly the same amount of posts. As illustrated in both figures, posts with few words in their descriptions receive the highest amount of interactions. As seen in Figure B.2a, posts with many words in their descriptions receive the least amount of likes, while posts with several words receive the least amount of comments as illustrated in Figure B.2b. Unlike Figure A.5, there was no significant difference in how many words each post type included for the 1000 users.

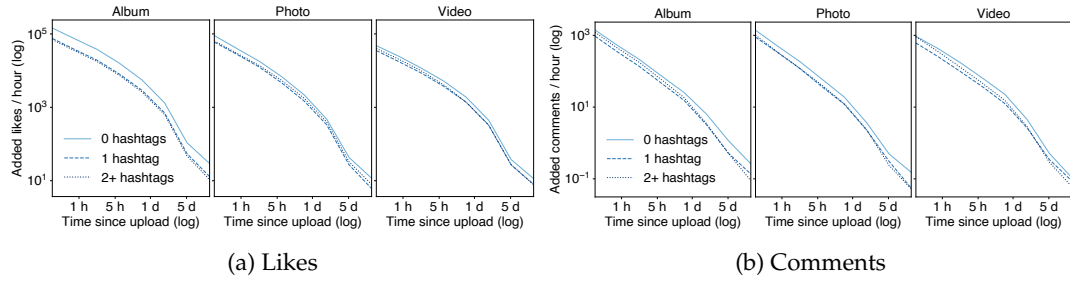


Figure B.3: Temporal impact of hashtags in posts' descriptions.

Figure B.3 illustrates the average temporal impact of hashtags included in posts' descriptions. As illustrated by the two figures, posts that receive the most interactions are posts that do not use any hashtags. Note that a correlation among posts that included different amounts of hashtags in their description could not be found other than receiving fewer interactions than posts that do not include hashtags.

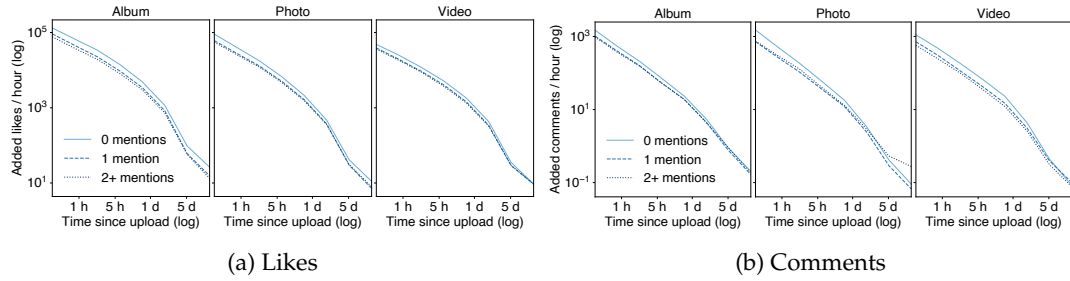


Figure B.4: Temporal impact of mentions in posts' descriptions.

Figure B.4 illustrate the average temporal impact of the number of users mentioned in posts' descriptions. As seen in the figures, posts that do not mention any users receive more user interactions compared to posts that do mention users in their descriptions. Note that a correlation among posts that mentioned users in their description could not be found other than receiving fewer interactions than posts that did not mention users.

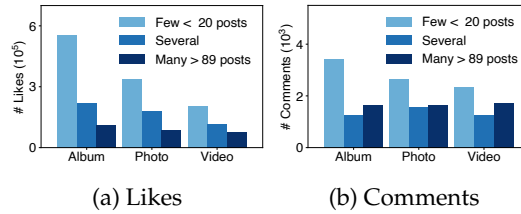


Figure B.5: Average impact of uploaded posts 7 days prior.

Figure B.5 presents the average impact of uploaded posts up to 7 days before the upload of each post. The three categories are determined to contain roughly the same number of posts. As observed in the two figures, posts with more prior uploaded posts receive fewer likes. On the other hand, posts with few prior uploaded posts receive the highest amount of comments, followed by posts with many prior uploaded posts.

Figure B.6 presents the average temporal impact of how many followers the uploader has at the time of upload of each post. The three categories are determined to contain roughly the same number of posts. As observed in the two figures, users with many followers receive the largest amount of interactions followed by users with several followers. Further, it can be seen in photos and videos how users with several followers have more interactions initially

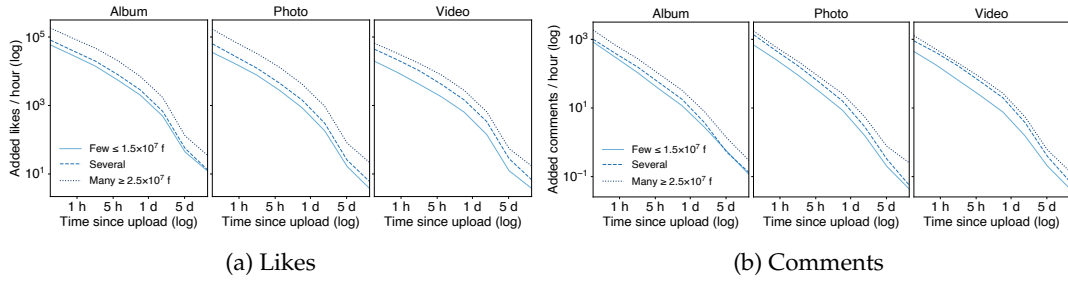
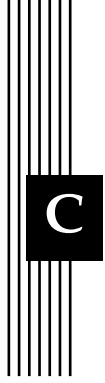


Figure B.6: Temporal impact of followers (f: followers).

compared to users with many followers. However, at a later stage, the difference is less significant. Lastly, it can be noted in Figure B.6 that posts uploaded by users with few and many followers receive likes at a similar rate and do not seem to be affected by the number of followers. Instead, each post seems to have a similar rate of attracting likes.



Additional Media Analysis

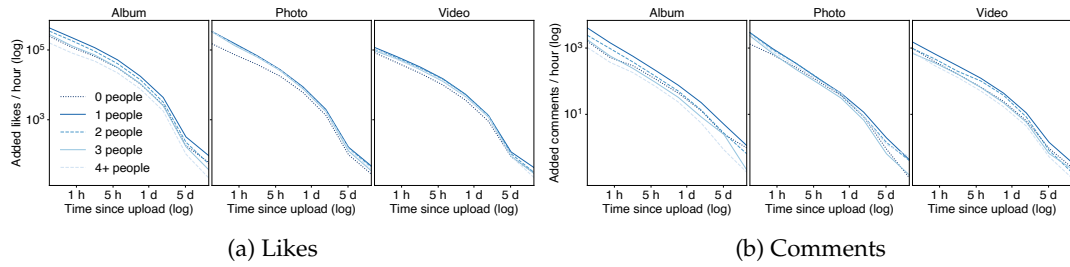


Figure C.1: Temporal impact of people in media files.

Figure C.1 presents the average temporal impact of people present in posts' media files. As seen in the two figures, they follow a similar trend as observed in Figure 6.4. Hence, media files without any people generally receive fewer interactions compared to files including 1 person. For albums, it can be observed that the more people are present in media files, the less interaction the posts receive. The trend can also be observed among videos in Figure C.1b, but it is not present for videos in Figure C.1a. For photos, the number of people present in the media file does not seem to affect the number of received interactions.

Figure C.2 presents the average impact of detected genders. Each detected gender had an accuracy of at least 86%. The threshold was obtained by calculating the average accuracy given across all genders detected by the model. The columns men and women include media files in which at least one predicted man or woman was detected. The column both genders

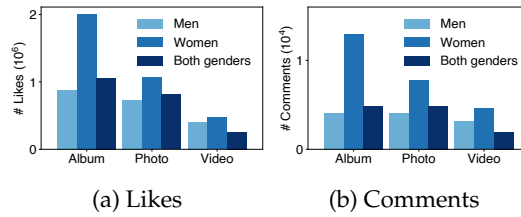


Figure C.2: Average impact of gender.

includes images in which both genders were detected at least once. As seen in the two figures, files with only women present receive the highest number of interactions for all post types. This is, for albums and photos, followed by files in which both genders are present. For videos, it is followed by files with only men detected.

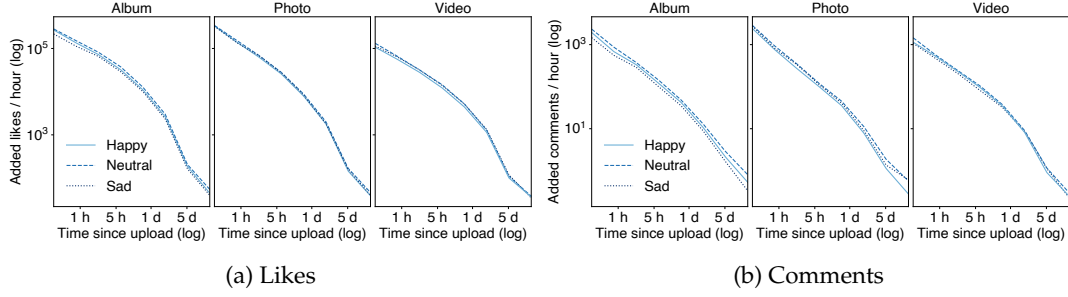


Figure C.3: Temporal impact of emotions.

Figure C.3 presents the temporal impact of the predicted emotions: happy, neutral, and sad. Each detected emotion had an accuracy of at least 80%. The threshold was obtained by calculating the average accuracy given across all emotions detected by the model. As seen in the figures, faces expressing neutral emotions receive the highest number of likes. For albums, this is followed by happy expressions, and for photos it is followed by sad expressions. For videos, happy expressions receive more likes but fewer comments than sad expressions.

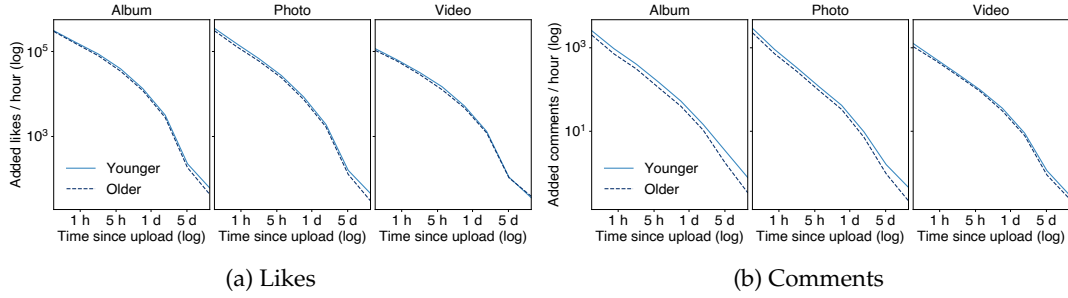


Figure C.4: Temporal impact of age.

Figure C.4 presents the average impact of predicted ages. Ages between 0 and 100 years old could be detected and the threshold was determined by the pre-trained model. To diminish the number of misclassified ages, detected people either belonged to the group younger (0-34) or older (35+) thus a majority of users were predicted to be in their early thirties. As illustrated in the figures, older people receive more added interactions on albums and videos, while younger people receive more added interactions on photos.