Instagram Post Popularity Trend Analysis and Prediction using Hashtag, Image Assessment, and User History Features

Kristo Radion Purba, David Asirvatham, and Raja Kumar Murugesan
School of Computer Science and Engineering, Taylor's University, Malaysia

Abstract: *Instagram is one of the most popular social networks for marketing. Predicting the popularity of a post on Instagram is important to determine the influence of a user for marketing purposes. There were studies on popularity prediction on Instagram using various features and datasets. However, they haven't fully addressed the challenge of data variability of the global dataset, where they either used local datasets or discretized output. This research compared several regression techniques to predict the Engagement Rate (ER) of posts using a global dataset. The prediction model, coupled with the results of the popularity trend analysis, will have more utility for a larger audience compared to existing studies. The features were extracted from hashtags, image analysis, and user history. It was found that image quality, posting time, and type of image highly impact ER. The prediction accuracy reached up to 73.1% using the Support Vector Regression (SVR), which is higher than previous studies on a global dataset. User history features were useful in the prediction since the data showed a high variability of ER if compared to a local dataset. The added manual image assessment values were also among the top predictors.*

1. Introduction

Instagram is the fastest growing social network, with nearly 5% growth each quarter [19]. With 1m active users as of June [20], it becomes the best platform for brand marketing for a millennial audience [10] and the platform with the highest Engagement Rate (ER) [8]. The number of likes is commonly perceived as a social status [5], which is useful for marketing purposes. However, it can create a stressful experience for users [16].

Predicting popularity on Instagram is important to find the best possible influencer for brand marketing [15], and to help the general public to discover trends on Instagram. There were studies on Instagram's popularity prediction, with features such as image content [14, 27], hashtag [4], image aesthetic [17], time [6], and metadata. However, there were no existing studies that addressed the challenge of data variability in global dataset. Existing studies have either small data variance due to usage of local/specific dataset [28], or small variance of output due to discretized output [27].

This research used hashtag analysis, image content, image quality, user history, as well as metadata as the features for prediction. Several social studies have shown that hashtags [12, 18] and image content [4] have a vital contribution to popularity. The output of the prediction is the ER of a media within one month from the upload date. The ER is the number of likes divided by followers. Usually, there is a peak period for likes, where a post will get far fewer likes after a month [1]. Thus, this research predicted the post popularity during the peak period.

The following questions were addressed in this research, i.e., (R1) what is the relation between the features (hashtag, image analysis, user history, metadata) and the Engagement Grade (EG)? (R2) How do these features affect the prediction of ER? (R3) What is the best regression method for popularity prediction?

The following hypothesis was formulated in this research, i.e., (H1) Global dataset has more ER variability.

This research aims to analyze the ER and the factors that affect it, as well as create a machine learning regression model to predict them. The contributions of this research are as follows:

- This is the first study that analyzes popularity on Instagram using global dataset.
- Compared to previous studies, we added hashtags analysis, image assessment, and user history features to predict the popularity of a post.

The result of this study is beneficial for both business and regular users. The extensive features set used in this study helps the general public to understand the complexity of post popularity on Instagram. The paper is organized as follows: related works, methodology

(data collection and features extraction), popularity analysis, and popularity prediction.

2. Related Works

There were studies on popularity on Instagram, with some more focused on the statistics, some on predictions. Most of the studies used metadata (number of followers, posts, etc.,) as a feature. There were additional features such as image content/category, image quality, text/sentiment analysis. There were four types of popularity metrics in the recent studies, i.e., likes [4, 6, 9, 14], engagement level [17], intrinsic image popularity [7], growth of likes [1], and categorized output, such as popular-unpopular [27].

There were two types of image analysis in the prediction of popularity in recent studies, i.e., content analysis and quality analysis. Image content analysis can be either categorization or featurization. Image category is a single field, such as brand image, people, selfie, etc., [14], whereas image feature is a field with multiple aspects, such as selfie, outdoor, food, girls, etc., [4]. Image aesthetic analysis was also used for popularity prediction [17]. In this study, we added manual image assessment values for popularity prediction, which hasn't been used in previous studies.

All existing studies that used hashtags were either a statistical study on the effect of hashtags in increasing popularity [4, 6] or using hashtag count for prediction [9, 17]. Existing popularity prediction studies only used hashtag counts for prediction, even though a recent study has found the usefulness of hashtag in terms of increasing likes [3]. Thus, we further exploit hashtag by extracting hashtag popularity and visibility values.

In terms of the dataset, some studies used local or region-specific datasets or global datasets. Iranian Instagram business accounts (consisted of 3 users and 281 posts) were used in a study [28], with prediction accuracy of up to 90.77% for three output labels. Single lifestyle magazine data was used in a study [6], with prediction accuracy of up to 88%. The global dataset is more challenging and was expected to produce lower prediction accuracy, only up to 71.19%, utilizing two output labels (popular, unpopular) [27].

Overall, existing studies lacked prediction on Instagram posts' popularity using the global dataset. While there was a study using a global dataset [27], it was focused on how to raise the popularity of a specific user. Thus, we used the dataset collected from global hashtags for popularity analysis and prediction.

There were other studies on various social media platforms, such as Flickr [9], Twitter [2], YouTube, and Facebook [23], which are different from Instagram in terms of features and ER.

3. Research Methodology

There were four phases in this research, i.e., data collection, data filtration, analysis, and popularity prediction, as shown in Figure 1.
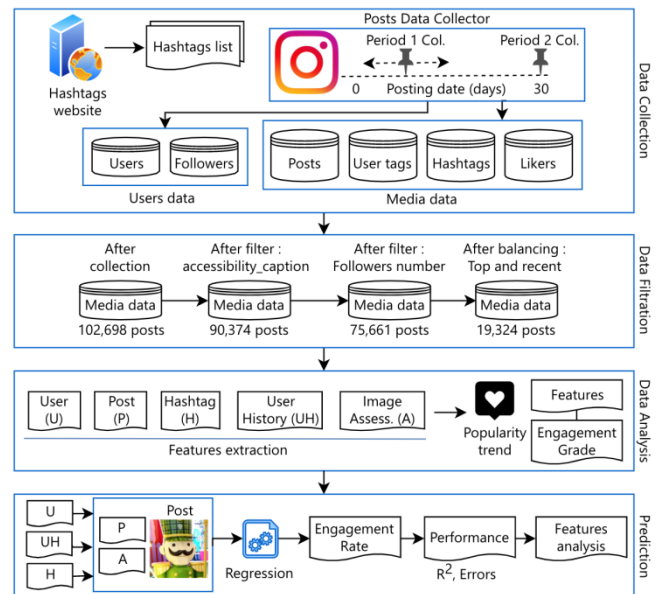

Figure 1. Research methodology.

3.1. Data Collection

Data collection was started by collecting 2,000 top hashtags from Top-Hashtags [22]. From the hashtags list, posts were collected using the Instagram Application Programming Interface (API). Data collection was done in two periods. The first period was used to get the posts listed under a hashtag at that time. Video posts and posts older than 30 days were removed.

The second period of data collection was used to collect the data of those posts exactly after 30 days of the upload date. Thus, a scraper application was built to constantly check the lifetime of each post available from period 1, and then it re-scraped the data of a post once it reached 30 days since the posting date. There were 6% of posts that were not available in the second period, due to these posts being taken out by Instagram or taken out by the owner.

3.2. Data Filtration

The number of posts data with age less than 30 days was 102,698 posts. From a hashtag, there are posts listed under top posts and most recent. One important variable that must be available from Instagram API is the accessibility-caption, to extract the image content. There were 12% of posts that didn't have this value.

From the remaining data, a filter was further done, based on the number of followers, i.e., removal of (assumed-to-be) new users and mega influencers. This was because the number of followers profoundly affect the ER [22]. Thus, posts from users with the following

criteria were removed, i.e., (1) Posts from users with < 100 followers, (2) Posts from users with total posts < 10, and (3) Posts with the number of likes < 5, and (4) Posts from users with > 1 million followers. Finally, the next step is to balance the posts, to get 50% posts from top posts and 50% from most recent. The final number of posts was 19,324 (from 16,804 unique users), which was used throughout this research.

4. Features Extraction

The source of the features was from the following table, i.e., user, post, hashtag. Other sources are from NIMA, annotation, and users' history. The list of features group and the features is as follows:

- User Features (U)

  - *flg*: Number of following (from 0 to 7,500).
  - *flr*: Number of followers (from 100 to 1,000,000).
  - *pos:* Number of posts (from 21 to 49,957).
  - *flgc*: Number of following category, i.e., 0 (for 0-500), … 9 (for 4,501-5,000), 10 (for > 5,000).
  - *flrc*: Number of followers category (0 to 10), according to Figure 5.
  - *posc*: Number of posts category, i.e. 0 (for 0-500), 1 (for 501-1,000), … 10 (for > 5,000).
  - *bl:* Biography length (characters).
  - *lin*: Link availability, i.e., 0 (no) or 1 (yes).

- Post Features (P)

  - *hc*: Hashtags count.
  - *lt*: Location tag, i.e., 0 (no location), 1 (using location).
  - *cl*: Caption length (characters)
  - *day*: Upload day, i.e., d1 (Sunday) to d7 (Saturday).
  - *tf*: Upload time, i.e., t1 (00.00-03.00)-t8 (21.00-24.00).
  - *ut*: Number of user tags.
  - *pic_**: Nine image content features from section 4.1.

- Hashtag Features (H)

  - *hp*: Hashtags popularity, i.e., sum of (global usage of each hashtag) in the caption.
  - *hv*: Hashtags visibility, i.e., sum of (growth rate of each hashtag). The growth rate is hours/post.
  - *hr*: Hashtag reachability, i.e., sum of (global usage * growth rate). This is a combination of *hp* and *hv*.

- Image Assessment (Auto) Features (Aa)

  - *aest*: Image aesthetic, i.e., 0.00 (bad)-1.00 (good).
  - *tech*: Technical quality, i.e., 0.00 (bad)-1.00 (good).

- Image Assessment (Manual) features (Am)

  - *beauty*: Beautiful/sexy woman, handsome man, beautiful scenery/object.
  - *artistic*: Artistic value (if drawing), photographic quality, or arrangements.
  - *emotion*: Emotional feeling of a human face or objects.
  - *unique*: Unique objects, or rarely seen image.

- User History Features (UH)

  - *ef*: Engaging followers, i.e., percentage of followers (unique users) who have liked any posts from a user.
  - *eo*: Engaging outsiders, i.e., same with *ef*, but for outsiders. Outsider is anyone outside the followers.
  - *av_erl*: Average of (number of likes/followers)
  - *av_erc: Average of (number of comments/followers).*
  - *st_erl: Standard deviation of (likes/followers)*
  - *st_erc: Standard deviation of (comments/followers).*
  - *min_er: Minimum of (likes+comments)/followers.*
  - *max_er: Maximum of (likes+comments)/followers.*
  - *av_er: Average of (likes+comments)/followers.*
  - *st_er: Std. deviation of (likes+comments)/followers.*

4.1. Post Features: Image Content

Image content is extracted from the keywords using the accessibility-caption variable. Then, the keywords were grouped into categories based on manual observation. These categories became image content features. If any of the keywords from a respective category match, the feature will be set to 1. Note that we have tried grouping the keywords using several keywords grouping API, but the results were either too many groups or most of the words grouped into "other".

For example, if the accessibility-caption is "Image may contain: one or more people, people sitting, screen and indoor," the features pic-activity, pic-wearable, pic-room will be set to 1. The image content features and the respective keywords are:

- *pic_text*: meme, text, that says.
- *pic_nature*: beach, sky, mountain, cloud, night, ocean, beach, snow, twilight, bridge, christmas, fireworks, outdoor, flower, plant, tree, grass, candle, cat, dog, horse, bird, nature, water.
- *pic_vehicle*: car, bicycle, motorcycle, boat, airplane, bus, road, flight.
- *pic_food*: eating, drink, fruit, food, dessert, icecream, sushi, pizza, smoking, coffee.

- *pic_talent*: sport, basketball, pool, swimming, riding, stadium, scraper, baseball, tennis, football, golf, chess, guitar, concert, musical, dancing, drawing, stage, makeup, wedding, stripes.
- *pic_room*: room, house, kitchen, office, indoor, table.
- *pic_person*: person, closeup, selfie, beard, child, baby.
- *pic_wearable*: shoes, suit, shorts, bros, boots, eyeglasses, hat, sunglasses, jewelry, ring, phone, screen, laptop, camera.
- *pic_activity*: people, crowd, sitting, standing, smiling, walking, sleeping.

4.2. Hashtag Features

The hashtag's popularity (hp) of a post is the sum of all hashtags used, as can be seen in Figure 2. The hashtag's visibility (hv feature) is based on the hashtag's growth rate (hg). The hg value of a hashtag is the average interval (in hours) between every two posts in the most recent posts in the hashtag. This hg value measures how quickly a post gets buried in the most recent list. The hv value of a post is the total of hg value of all hashtags contained in a post caption, as illustrated in Figure 3. This total value (instead of average) gives a perspective of how long a post will appear in the most recent list, contributed by all hashtags in a post.
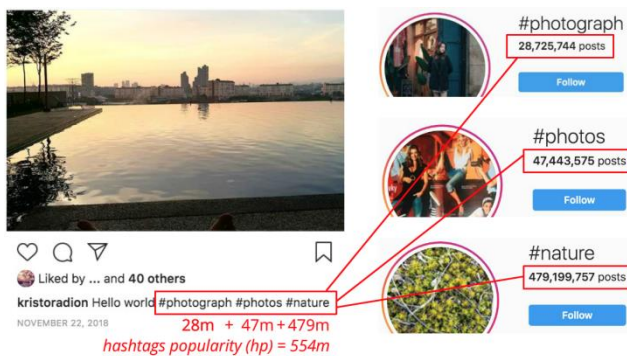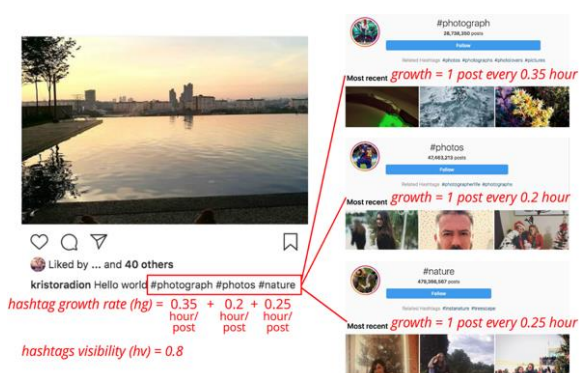

Figure 2. Hashtags popularity (hp) of a post.

Figure 3. Hashtags visibility (hv) of a post.

Choosing hashtags for a post can be tricky since popular hashtags are usually fast-growing. While hp value simulates "audience size," the hv value simulates

"advertisement duration." High hv value indicates a slow-growing hashtag, which is good for a post since the visible duration is longer. Another feature, the hr (hashtag reachability), is the hp multiplied by hv, which is the chance of a post to reach a broader audience.

4.3. Image Assessment Features

There were two image assessment features in this research, i.e., automated assessment and manual (with annotation). The automated assessment was divided into image aesthetic (ia feature) and image technical quality (it feature), which was assessed using the source code available in GitHub [13]. The source code was the implementation of Google's Neural Image Assessment (NIMA) [21]. The technical quality assesses details such as noise, blur, artifacts, and so on, whereas the aesthetic assesses semantic characteristics.

It was suggested that an image's beauty, artistic, emotion, and unique values highly rely on high-level semantic [11, 25]. This creates many challenges for automatic assessment [11]. Thus, a manual assessment was added. To reduce subjectivity, three human judges (photography experts) were assigned. Each image has to be given a discrete score (from 0 to 3) of beauty, artistic, emotion, and unique values. The values from all judges were averaged and converted to a range of 0.00 to 1.00.

4.4. User History Features

For each user, who is the owner of each post, 15 latest posts (with the age of at least 30 days) were captured. To make a fair prediction, the posts that were available inside the main dataset were excluded from the user history. This is to ensure that the created prediction model would be able to forecast future posts of a user.

Besides the post data itself, the likers and the followers of each user were captured. By using these data, the unique likers can be captured. A unique liker is either a follower, which is called as *engaging followers* (*ef*); or a non-follower, which is called as *engaging outsiders* (*eo*). Both *ef* and *eo* were used to estimate the active audience size of each user in order to improve prediction accuracy.

4.5. Prediction Output

The output of the prediction is the ER within one month from the upload date, which is the number of (likes+comments) divided by followers. The range of the ER was 1.1% to 19.9%.

5. Popularity Analysis

In this section, features are plotted to examine the contribution of each factor to the ER. In Instagram market studies, it was proven that the number of followers highly contributes to the ER. For example,

users with <2,000 followers have an average ER of 10.7%, all the way up to users with >1 million followers have 1.5% average ER [24]. The linear trend line in Figure 4 shows that the ER is decreasing by 0.026% for every 10,000 followers increase, according to our data.
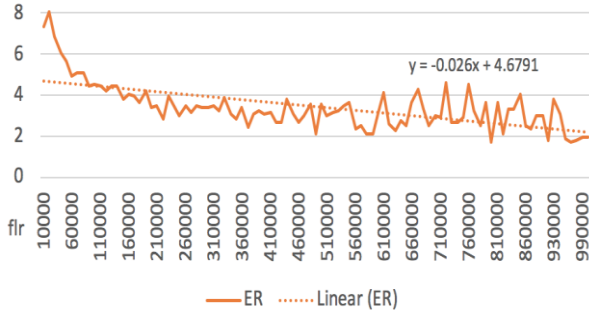

Figure 4. Number of followers vs. ER.

The ER comparison is not fair across users with a different number of followers, since higher tier ER values are produced from users with lesser followers. This is consistent with a report [24]. Thus, for a fairer comparison, the EG metric was defined. The ER chart [24] was divided equally into 12 regions, with 4 below average regions and 8 above average regions, as can be seen in Figure 5.
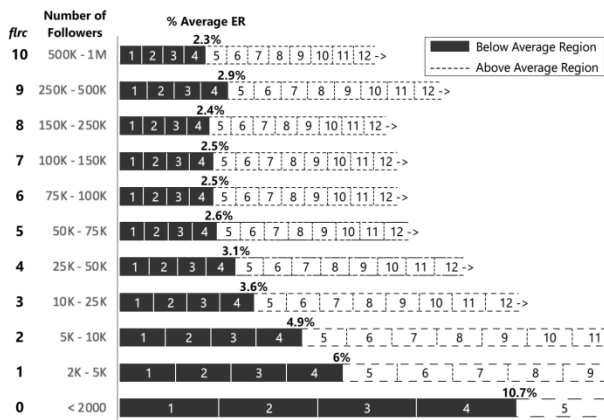

Figure 5. Follower category (*flrc*) and EG regions.

EG is ranging from 1 to 12. For example, for a user with 260K followers, and ER 3.0%, the EG is 5 since it is inside region 5. The EG metric sets a higher standard for users with lesser followers, and vice versa.

5.1. Contribution of Users Features on EG

The following Table 1 shows the EG and user metadata features comparison. The table shows that the number of followers and following are equally distributed, which shows a fairer comparison, unlike Figure 4. The features *bl* and *lin* didn't show a significant contribution to EG. Interestingly, the *flg* and *pos* decrease as the EG increase, as seen in Figure 6. This means that better EGs were coming from less spammy users, as also suggested by [17].

Table 1. EG and user features comparison.

| EG	Data Count	flg	flr	pos	bl	lin
1	675	1361.0	1412.0	2297.8	136.3	0.60
2	2,120	1137.7	30833.3	1600.8	135.2	0.56
3	2,786	1099.7	94837.7	1952.1	142.9	0.56
4	2,358	1021.2	86672.5	1746.5	146.1	0.58
5	2,146	1007.1	80331.6	1535.3	144.3	0.58
6	1,745	1005.4	72425.7	1365.8	143.5	0.54
7	1,495	1174.3	74240.5	1267.8	152.0	0.58
8	1,101	1094.2	69767.0	1312.9	150.9	0.61
9	994	1133.0	67713.8	1206.8	150.0	0.58
10	802	1120.1	67714.2	1015.5	143.5	0.56
11	662	1102.9	61316.8	898.9	148.6	0.55
12	2,440	881.6	69483.7	740.7	139.5	0.53
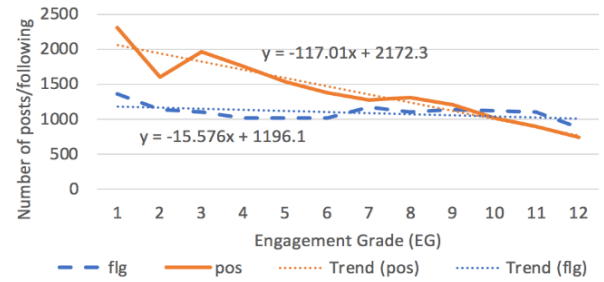
Figure 6. Higher number of following and posts lead to lesser EG.

5.2. Contribution of Post and Hashtag Features on EG

The next comparison is EG with post features, shown in Table 2. One important aspect from this table is that posts with higher EG used more user tags, with an average of 2 user tags on the highest EG. The features *hc*, *cl*, *lt*, didn't show a significant contribution to EG.

Table 2. EG and post features comparison.

EG	hc	lt	cl	ut	hp	hv	hr
1	17.29	0.40	328.6	0.32	0.20	0.16	0.08
2	17.52	0.45	347.5	0.63	0.20	0.18	0.09
3	15.57	0.44	367.2	0.83	0.17	0.14	0.07
4	15.24	0.45	363.0	1.08	0.17	0.14	0.08
5	15.17	0.48	376.7	1.22	0.15	0.15	0.07
6	15.16	0.49	375.0	1.45	0.15	0.15	0.07
7	15.55	0.49	395.3	1.64	0.15	0.14	0.06
8	15.15	0.48	394.4	1.80	0.13	0.13	0.05
9	15.40	0.48	372.1	1.56	0.14	0.12	0.05
10	15.84	0.49	386.1	2.09	0.14	0.12	0.05
11	15.98	0.46	390.8	2.23	0.15	0.14	0.06
12	16.03	0.38	364.3	2.00	0.12	0.12	0.06

All hashtag features (*hp, hv, hr*) contribute to slightly lowering EG value, as can be seen in Figure 7. Higher hashtag count (*hc*), on the other hand, can slightly increase EG, as can be seen in Figure 8. The linear trend shows that every 1 hashtag increase will increase EG by +0.0134. In addition, the plot also shows that the best number of hashtags to get the highest EG is 20. All these data show that hashtag tricks do not help a lot in increasing engagement, but post quality does.
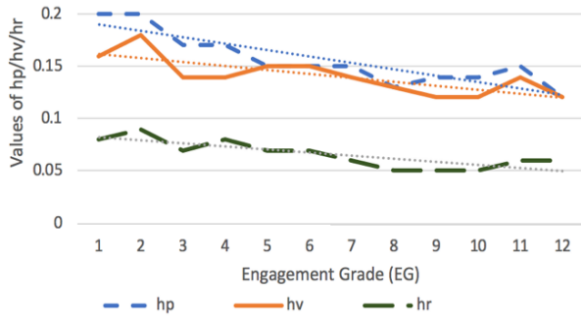
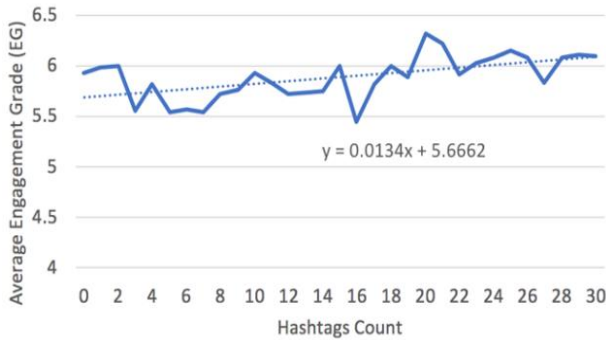Figure 7. EG Comparison with hashtag features.

Figure 8. Higher hashtags count slightly increases EG.

There were two categorical post features, i.e., day and time. The plot of day and time against average EG can be seen in Figures 9 and 10. It can be concluded, according to our data, that the best day for uploading posts is weekdays, especially Tuesday and Wednesday. The best upload time is 00am to 06am.
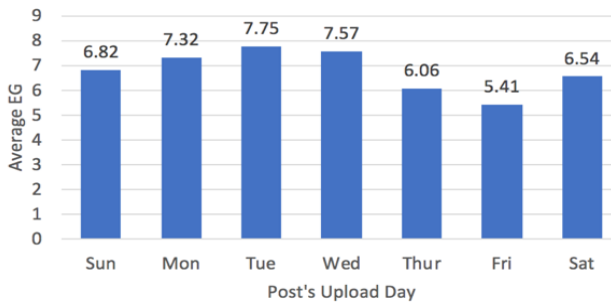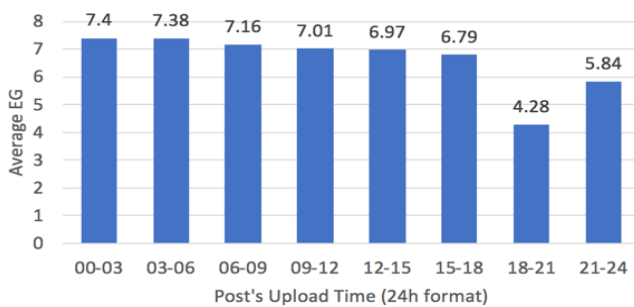

Figure 9. Upload day vs. average EG.

Figure 10. Upload time vs. average EG.

5.3. Contribution of Image Content on EG

To compare between different image content categories, another plot of image contents vs. average EG is shown in Figure 11. According to our data,

pictures containing food are the most interesting to users, followed by text images (usually memes or motivational texts) and activity (usually photograph of person/people). The least interesting is talent images, such as people dancing or playing music. Figure 12 shows some examples of each kind of picture. Note that one image can contain some image contents.
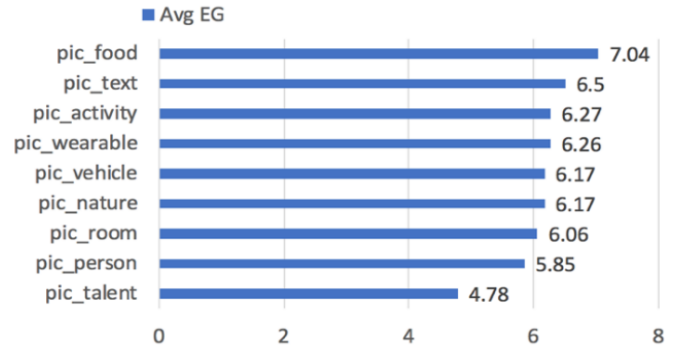

Figure 11. Comparison between image contents and average EG value.

Figure 12. Examples of image contents.

5.4. Contribution of Image Assessment on EG

The EG comparison with image assessment features is shown in Table 3. It is shown consistently that higher EG values have higher quality images. In terms of automatic assessment, tech outperforms aest with an average increase of 0.031 between two adjacent levels. In terms of manual assessment, beauty and artistic are the highest performers. Every improvement of image aesthetic by 0.0019, or technical quality by 0.0031, or beauty value by 0.0089, or artistic value by 0.0092, can lead to an increase of EG by 1 level.

Table 3. EG and image assessment features.

EG	aest	tech	beauty	artistic	emotion	unique
1	0.453	0.502	0.75	0.71	0.68	0.33
2	0.457	0.511	0.79	0.75	0.70	0.31
3	0.469	0.515	0.82	0.78	0.70	0.34
4	0.471	0.516	0.84	0.79	0.70	0.33
5	0.476	0.522	0.84	0.80	0.70	0.34
6	0.475	0.525	0.85	0.80	0.70	0.34
7	0.477	0.525	0.84	0.81	0.70	0.35
8	0.480	0.528	0.84	0.80	0.69	0.34
9	0.477	0.528	0.84	0.81	0.69	0.34
10	0.478	0.529	0.84	0.81	0.68	0.34
11	0.479	0.535	0.85	0.81	0.68	0.33
12	0.475	0.536	0.84	0.81	0.68	0.32
Avg Increase	0.0019	0.0031	0.0089	0.0092	0.0006	-0.0009

5.5. Contribution of User History on EG

Moving on to the user audience features, Figure 13 shows EG as the X-axis and values of *ef* and *eo* as Y-axis. It is shown that higher EG values are always caused by a more active audience. Creating more

interesting content, such as improving image quality, can be a way to increase audience activeness.
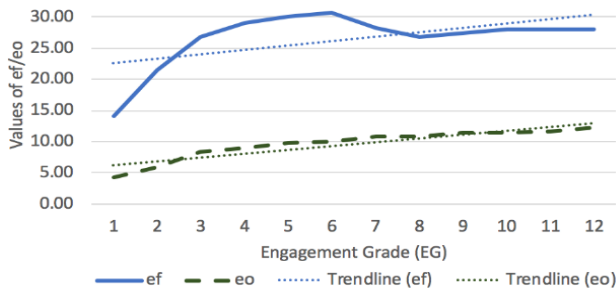

Figure 13. EG Comparison with engaging followers (ef) and engaging outsiders (eo).

Statistics-based history features (average, standard deviation) were used to help in the popularity forecast of each user. These features helped in rectifying the huge deviation of ER between users, even from the same follower category (*flrc*). The average of *av-er,* as well as the deviation, can be seen in Figure 14.
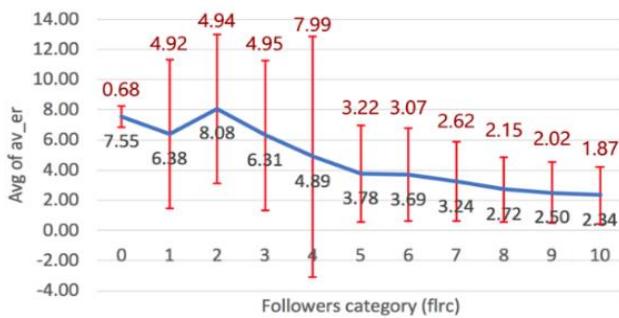

Figure 14. Followers category vs. average and standard deviation of *av-er* (Global dataset).

Additionally, in order to prove H1, a local dataset was collected from various Instagram analytics websites. It consisted of 1,152 Indonesian influencers, with a minimum number of followers of at least 2,000. The ER were calculated from influencers that, on average, has 15 latest posts each. The plot of average and standard deviation of *av-er* (Figure 15) of them show a significant difference with the previous plot. This proves that the local dataset (influencers) has lesser ER variance compared to the global dataset.
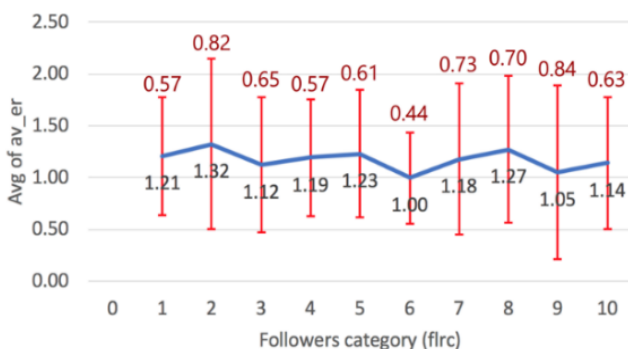

Figure 15. Followers category vs. average and standard deviation of *av-er* (local dataset).

6. Popularity Prediction

In this section, machine learning regression methods are compared to predict (ER, not EG), as a floating number. Compared to EG, ER is a more usable number since it directly shows the rate of likes. Along with the prediction, features analysis was carried out and discussed in this section. The tested methods are Linear Regression (LR), Random Forest regressor (RF), and Support Vector Regression (SVR).

6.1. Prediction Results

There were challenges in getting some features, i.e., costly human annotation, and extensive API usage to get user history. Thus, in the regression implementation, several combinations of features group were tested and presented to give ideas for future studies.

As the main accuracy measure, R^2 (R-squared) was used, along with the error measures, i.e., Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Relative Absolute Error (RAE), and Root Relative Squared Error (RRSE). The methods used were LR, regressor RF, SVR with a radial kernel. All prediction results are shown in Table 4.

Table 4. Regression prediction results (using 10-folds cross validation).

Features Group	Method	R^2	MAE	RMSE	RAE	RRSE	Best R^2
UPH	LR	0.329	2.53	3.131	0.806	0.819	0.426 (RF)
	RF	0.426	2.28	2.901	0.726	0.759	
	SVR	0.396	2.294	3.003	0.731	0.786	
UPH, Aa	LR	0.331	2.528	3.127	0.805	0.818	0.427 (RF)
	RF	0.427	2.284	2.9	0.727	0.759	
	SVR	0.397	2.294	2.999	0.731	0.785	
UPH, Aa, UH	LR	0.516	1.913	2.676	0.609	0.7	0.648 (RF)
	RF	0.648	1.697	2.269	0.54	0.594	
	SVR	0.641	1.671	2.31	0.532	0.604	
UPH, Aa, UH, Am	LR	0.579	1.809	2.492	0.576	0.652	0.731 (SVR)
	RF	0.697	1.58	2.112	0.503	0.553	
	SVR	0.731	1.464	1.995	0.466	0.522	

Prediction using only UPH features produced R^2 of 0.426, which is simply not enough. The addition of automatic image assessment (UPH, Aa) did not really help either. By adding the user history features (UPH, Aa, UH), the best R^2 was increased to 0.648. This group (UPH, Aa, UH) used all readily available features, without the addition of the subjectivity of manual assessment. By adding the manual assessment, the best R^2 was increased to 0.731. This shows the rate of likes involves the general characteristics of an image.

6.2. Features Importance

Although the best R^2 was acquired from SVR, there is no native way to get features importance for SVR. Thus, another way to achieve this is by calculating the Pearson's correlation of each feature with ER, as

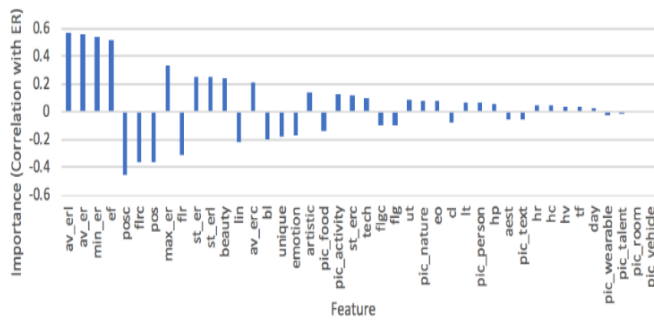shown in Figure16, or from RF attribute importance calculation, as shown in Figure 17.
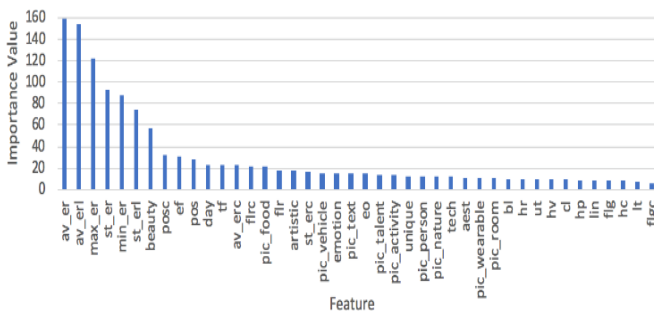

Figure16. Features importance based on correlation.

Figure 17. Features importance from random forest.

7. Determinants of Popularity

The correlation (Figure 16) is quite consistent with the RF's features importance (Figure 17). Based on the features importance (section 6) with popularity trend (section 5), it can be concluded that:

- User history features were the most important predictors. These features were used to get the followers demography, particularly how many are active followers, to mitigate the huge deviation of ER, as presented in Figure 14. Thus, history is useful for a future forecast of posts from a particular user.
- Increasing image quality, especially *beauty* and *artistic*, can improve ER. For the aesthetic and technical quality (auto values), even though they have a low rank as predictors, the popularity trend shows that improving them still helps in increasing ER.
- There were five user features. Biography length (*bl*) and link availability (*li*) did not show significance both in features importance and popularity trend. The number of followers (*flr, flrc*), posts (*pos, posc*) and following (*flg, flgc*) have a quite high rank on predicting ER. Consistent with earlier discussions (Figure 4), that bigger size users (high followers and posts) have indeed lesser ER. Thus, the conclusion cannot be drawn from ER, but from EG. Earlier data in Figure 6 showed that less spammy users (less following and posts) gain higher EG.
- Day (day) and time of post (*tf*) are significant, as can be seen in RF features importance. Since these

features are categorical, the correlation rank (Figure 16) didn't give a suitable result as they were being treated as numerical. Both feature importance and popularity trend show that it is important to pick the best day (Tuesday and Wednesday) and time (00am to 06am) to upload a post in order to raise EG.

- Two post features, i.e., location (*lt*) and caption length (*cl*) didn't show significance in both importance and trend. The user tags (*ut*), on the other hand, can help in increasing EG, as seen in Table 2.
- Among the image content features, pic-food has a higher rank as a predictor. Earlier data in Figure 11 also showed that food pictures get higher EG. Both these data amplify the fact that food pictures are the best pictures to post.
- Increasing hashtags count, as shown in Figure 8, can help in very slightly increasing EG. However, these hashtag features (*hp, hv, hr, hc*) have a low rank as predictors. This means that engagement is contributed mostly by other factors, not hashtags.

8. Conclusions

In this study, we have done an analysis of popularity trend and prediction on Instagram, using a set of features acquired from user metadata, post, hashtag, image assessment, and user history. In the analysis of popularity trend, EG is used in comparison to respect the lower ER of users with higher followers. In the prediction, ER was used as the output since it is more readable.

It was found that the most important factors in raising EG were image quality, day and time of post, user tags, and type of image. In terms of predicting (or forecasting) ER, the user's history data is very important. The history was used to mitigate the high variability of ERs between users in the global dataset. Compared to the local dataset, the global dataset was proven to have a much higher deviation of ER. Other important features that contributed to ER were image quality, upload day, and time of the post.

Prediction accuracy, measured with R^2, can reach up to of 73.1% with all features, and 64.8% without manual image assessment. This accuracy is enough for practical use and has a significance compared to previous studies. Regular users can take advantage of the popularity trend results in determining how to increase likes. Business users can also be benefited in terms of finding influencers for brand marketing.

In future research, the manual assessment values in this study can be changed to similar automated values in order to reduce subjectivity. Other features, such as user history, can still be tuned to get better results. Text analysis features, such as sentiment analysis and concept semantic similarity [26], can also be added to distinguish between popular or less popular posts.

References

[1] Almgren K., Lee J., and Kim M., "Predicting the Future Popularity of Images on Social Networks," *in Proceedings of The 3rd Multidisciplinary International Social Networks Conferenc*e, Union, pp. 1-6, 2016.

[2] Bae Y. and Lee H., "Sentiment Analysis of Twitter Audiences: Measuring the Positive or Negative Influence of Popular Twitterers," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 12, pp. 2521-2535, 2012.

[3] Berger J. and Milkman K., "What Makes Online Content Viral," *Journal of Marketing Research*, vol. 49, no. 2, pp. 192-205, 2012.

[4] Cakmak K., Cikrikcioglu I., and Demiralp O., "The Causal Determinants of Popularity in Instagram," Technical Report, 2017.

[5] Code M., Instagram, Social Media, and The "Like": Exploring Virtual Identity's Role in 21st Century Students' New Socialization Experience, M.S Theses, Brock University, 2015.

[6] De S., Maity A., Shitole S., Goel V., and Bhattacharya A., "Predicting the Popularity of Instagram Posts for a Lifestyle Magazine Using Deep Learning," *in Proceedings of 2nd IEEE International Conference on Communication Systems, Computing and Information Technology Applications*, Mumbai, pp. 174-177, 2017.

[7] Ding K., Ma K., and Wang S., "Intrinsic Image Popularity Assessment," *in Proceedings of the 27th ACM International Conference on Multimedia*, Nice, pp. 1979-1987, 2019.

[8] Feehan B., "Social Media Industry Benchmark Report." https://www.rivaliq.com/blog/2019-social-media-benchmark-report/, Last Visited, 2020.

[9] Gelli F., Uricchio T., Bertini M., Del Bimbo A., and Fu Chang S., "Image Popularity Prediction in Social Media Using Sentiment and Context Features," *in Proceedings of the 23rd ACM International Conference on Multimedia*, New York, pp. 907-910, 2015.

[10] Iqbal M., "Instagram Revenue and Usage Statistics." https://www.businessofapps.com/data/instagram-statistics/, Last Visited, 2020.

[11] Joshi D., Datta D., Fedorovskaya E., Lu X., Luong Q., Wang J., Li J., and Luo J., *On Aesthetics and Emotions in Scene Images: A Computational Perspective*, MIT Press Scholarship Online, 2014.

[12] Khalid N., Jayasainan S., and Hassim N., "Social Media Influencers-Shaping Consumption Culture Among Malaysian Youth," *International Conference on Humanities and Social Sciences*, Kuala Lumpur, pp. 1-12, 2018.

[13] Lennan C., Nguyen H., and Tran D., Image Quality Assessment, https://github.com/idealo/image-quality Assessment, Last Visited, 2019.

[14] Mazloom M., Rietveld R., Rudinac S., Worring M., and Dolen W., "Multimodal Popularity Prediction of Brand-Related Social Media Posts," *in Proceedings of the 24th ACM international Conference on Multimedia*, Amsterdam, pp. 197-201, 2016.

[15] Nandagiri V. and Philip L., "Impact of Influencers from Instagram and Youtube on Their Followers," *International Journal of Multidisciplinary Research and Modern Education*, vol. 4, no. 1, pp. 2454-6119, 2018.

[16] Paul K., "Instagram tests hiding how many people like a post. That has influencers worried." https://www.theguardian.com/technology/2019/nov/15/instagram-likes-influencers-social-media, Last Visited, 2019.

[17] Qian C., Tang, J., Penza M., and Ferri C., "Instagram Popularity Prediction via Neural Networks and Regression Analysis," pp. 2561-2570, 2017.

[18] Saxton G. and Waters R., "What do Stakeholders Like on Facebook? Examining Public Reactions to Nonprofit Organizations' Informational, Promotional, and Community-Building Messages," *Journal of Public Relations Research*, vol. 26, no. 3, pp. 280-299, 2014.

[19] Southern M., "Instagram Has 1 Billion Monthly Users, Now the Fastest Growing Social Network." https://www.searchenginejournal.com/instagram-1-billion-monthly-users-now-fastest-growing-social-network/258127/, Last Visited, 2019.

[20] Statista., "Most Famous Social Network Sites Worldwide as of, Ranked By Number Of Active Users (in millions)," https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/, Last Visited, 2018.

[21] Talebi H. and Milanfar P., "Nima: Neural Image Assessment," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3998-4011, 2018.

[22] Top-Hashtags, Top 100 HashTags on Instagram. https://top-hashtags.com/instagram/, Last Visited, 2020.

[23] Trzcinski T. and Rokita P., "Predicting popularity of Online Videos Using Support Vector Regression," *IEEE Transactions on Multimedia*, vol. 19, no. 11, pp. 2561-2570, 2017.

[24] Vorhaus J., "Instagram Influencer Rates," http://blog.influence.co/instagram-influencer-rates/, Last Visited, 2019.

[25] Windt G., Artistic Creativity: Transforming Sorrow Into Beauty, Truth and Art, M.S Theses,

Simon Fraser University, 2004.

[26] Zhang X., Sun S., and Zhang K., "A New Hybrid Improved Method for Measuring Concept Semantic Similarity in WordNet," *The International Arab Journal of Information Technology*, vol. 17, no. 4, pp. 433-439, 2020.

[27] Zhang Z., Zhou Z., Li J., Chen T., and Luo J., "How to Become Instagram Famous: Post Popularity Prediction with Dual-Attention," *in Proceedings of IEEE International Conference on Big Data* , Seattle, pp. 2383-2392, 2018.

[28] Zohourian A., Sajedi H., and Yavary A., "Popularity Prediction of Images and Videos on Instagram," *in Proceedings of 4th International Conference on Web Research*, Tehran, pp. 111-117, 2018.

Kristo Radion Purba is currently a computer science PhD student at Taylor's University Malaysia, starting from 2018. His research interests are in artificial intelligence, machine learning, and social network influence maximization. Prior to joining Taylor's, he was an informatics lecturer at Petra Christian University, Indonesia for 4 years (2014-2018), and also a contracted programmer at EHS (Environment, Health and Safety) department at PT. HM. Sampoerna, Tbk, Indonesia (2013-2017). He is also an active mobile apps, games, websites developer since 2008 until now.

David Asirvatham Dr. David Asirvatham is currently the Head for the School of Computing and IT, Taylor's University. Prior to this, he was the Director for the Centre of Information Technology at University of Malaya. He has held numerous posts such the Associate Dean for Faculty of Information Technology (Multimedia University), Project Manager for the Multimedia and IT Infrastructure Development for a university campus (US$14 million), Finance Committee for Multimedia University, SAP Advisory Council, Consultant for e-University Project and many more. Dr. David completed his Ph.D. from Multimedia University, M.Sc. (Digital System) from Brunel University (U.K.), and B.Sc. (Hons) Ed. and Post-Graduate Diploma in Computer Science from University of Malaya. He has been lecturing as well as managing ICT projects for the past 25 years. His area of expertise will include Neural Network, E-Learning, ICT Project Management, Multimedia Content Development and recently he has done some work on Big Data analytics.

Raja Kumar Murugesan Dr Raja Kumar Murugesan is an Associate Professor of Computer Science, and Head of Research for the Faculty of Innovation and Technology at Taylor's University, Malaysia. He has a PhD in Advanced Computer Networks from the Universiti Sains Malaysia, and has over 28 years' experience as an educator. His research interests include IPv6, and Future Internet, Internet Governance, Computer Networks, Network Security, IoT, Blockchain, Machine Learning, and Affective Computing. He is a member of the IEEE and IEEE Communications Society, Internet Society (ISOC), and associated with the IPv6 Forum, Asia Pacific Advanced Network Group (APAN), Internet2, and Malaysia Network Operator Group (MyNOG) member's community.