Social Network Analysis and Mining (2022) 12:47
https://doi.org/10.1007/s13278-022-00872-1

ORIGINAL ARTICLE

Understanding social engagements: A comparative analysis of user and text features in Twitter

Cagri Toraman[1] · Furkan Şahinuç[1] · Eyup Halit Yilmaz[1] · Ibrahim Batuhan Akkaya[1]

Received: 23 August 2021 / Revised: 23 February 2022 / Accepted: 24 February 2022 / Published online: 31 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Information is spread as individuals engage with other users in the underlying social network. Analysis of social engagements can therefore provide insights to understand the motivation behind how and why users engage with others in different activities. In this study, we aim to understand the driving factors behind four engagement types in Twitter, namely like, reply, retweet, and quote. We extensively analyze a diverse set of features that reflect user behaviors, as well as tweet attributes and semantics by natural language processing, including a deep learning language model, BERT. The performance of these features is assessed in a supervised task of engagement prediction by learning social engagements from over 14 million multilingual tweets. In the light of our experimental results, we find that users would engage with tweets based on text semantics and contents regardless of tweet author, yet popular and trusted authors could be important for reply and quote. Users who actively liked and retweeted in the past are likely to maintain this type of behavior in the future, while this trend is not seen in more complex types of engagements, reply, and quote. Moreover, users do not necessarily follow the behavior of other users with whom they have previously engaged. We further discuss the social insights obtained from the experimental results to understand better user behavior and social engagements in online social networks.

Keywords Natural language processing · Online social network · Social engagement · Text features · Tweet · User features

1 Introduction

With the rapid growth of online social networks, people can connect with each other and reach information easily. This phenomenon creates new challenges for social computing research to understand and analyze the role of social engagements in information spread, such as DARPA's SocialSim Project (Kettler 2018) and the RecSys 2020 Challenge (Anelli et al. 2020).

Information is spread as individuals engage with other users in the underlying social network. The volume and speed at which information spreads can be significant with so many daily activities. Analysis of social engagements can provide insights into the motivations that drive how and why individuals engage with others through various engagements. For instance, gamification approaches increase user engagement with social networks in the long run (Hajarian et al. 2019). Social engagements have important implications in various domains in social media, e.g., preventing hate speech and misinformation (Fortuna and Nunes 2018), promoting public health awareness (Vraga et al. 2018), simulating online social networks (Chung et al. 2019), and developing marketing strategies (Rui et al. 2013).

Social engagements have particular features that reflect the main characteristics of engagements. Existing studies mostly neglect feature comparison (Lee et al. 2014; Zamani et al. 2014) or focus on getting optimal model performance for predicting social engagements (Anelli et al. 2020, 2021; Schifferer et al. 2020). There is a research gap to better understand the main characteristics of social engagements by analyzing important and failed features for various types of social engagements on Twitter. For instance, users' previous activities or tweets' contents can indicate why they

✉ Cagri Toraman
ctoraman@aselsan.com.tr

Furkan Şahinuç
fsahinuc@aselsan.com.tr

Eyup Halit Yilmaz
ehyilmaz@aselsan.com.tr

Ibrahim Batuhan Akkaya
ibakkaya@aselsan.com.tr

[1] Aselsan Research Center, 06200 Yenimahalle, Ankara, Turkey

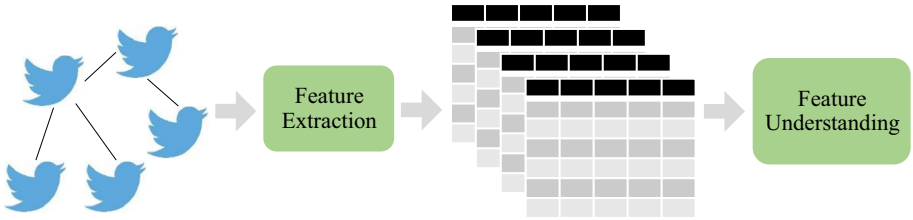

Fig. 1 An illustration of the phases that we follow to understand social engagements

would engage with others. However, not all the features that sound effective might be useful as expected. Our motivation in this study is not to find the most effective or efficient prediction models but to understand the driving factors behind different engagement types in Twitter.

The main objectives of this study are to (i) provide an analysis of the effectiveness of the individual features and feature groups concerning different engagement types, (ii) inspect the effect of the textual features, and (iii) understand which features to prefer and which features to avoid. We thereby extract a diverse set of features that reflect user behaviors, as well as tweet content and semantics through natural language processing. We compare the performance of these features in a supervised task of engagement prediction, using over 14 million multilingual tweets.

In this study, we focus on Twitter, a microblogging platform where users can post tweets with a limited length. Users can interact with other users by four types of engagements in Twitter, namely like (promoting a tweet), retweet (sharing a tweet with the followers), reply (answering to a tweet), and quote (commenting to a tweet while sharing with the followers).

We illustrate the phases that we follow to understand social engagements in Fig. 1. The input is a collection of social engagements that includes Twitter's like, retweet, reply, and quote. We extract substantial features based on user and tweet for each type of engagement. The extracted features are then learned and evaluated by prediction models in a supervised way so that important and failed features can be analyzed to provide insights on social engagements.

The main contributions of this study are as follows:

1. We examine a comprehensive list of features derived from various signals, including user's account and behavior, tweet's meta attributes, and text content. To the best of our knowledge, our study is the first to provide a comprehensive analysis of individual features and feature groups in social engagements with respect to four engagement types, i.e., like, retweet, reply, and quote.
2. We provide a focused analysis on textual features and reveal that users would engage with tweets based on text semantics and contents. We examine both traditional bag-of-words models (Aggarwal and Zhai 2012) and recent text sequence embeddings provided by a deep learning language model, BERT (Devlin et al. 2019).
3. We associate our experimental results with insights on user behavior and provide a better understanding of user behavior and social engagements in online social networks.

The rest of the study is organized as follows: In the following section, we give a summary of related work. Next, we explain our approach for understanding social engagements in Sect. 3. We then present the experimental details and results in Sect. 4, and provide a brief discussion on social insights and limitations of our study in Sect. 5. Finally, we conclude the study in the last section.

2 Related work

In this section, we give a summary of the related work using three sub-topics: (i) Feature extraction in online social networks, (ii) social engagements, and (iii) implications of social engagements.

2.1 Feature extraction in online social networks

User behavior in online social networks can be represented by the features extracted from the user's previous activity, i.e., the tweets the user engaged with in the past. Textual features are exploited to understand the context of target tweets (Savargiv and Bastanfard 2013). Such features are modeled using traditional machine learning and recent deep learning algorithms to accomplish predefined tasks (Zamani et al. 2014; Schifferer et al. 2020). User features (e.g., the number of followers and followees) and textual features (e.g., hashtags, URLs, and mentions) are extracted to predict whether tweets will be retweeted by other users (Chen and Pirolli 2012).

Some of the features that we examine in this study are also used in prediction models in the literature, but not for understanding user behaviors. The number of followers and followees, user's verification status, hashtag, and media existence, tweet type, user's previous activity (e.g., the frequency of liking tweets by the user per day) are examined in (Zamani et al. 2014; Lee et al. 2014; Volkovs et al. 2020).

The idea of using conditional probabilities of interactions concerning categorical features (Schifferer et al. 2020), and extracting textual features (e.g., BERT embeddings and word counts) are studied in (Schifferer et al. 2020; Volkovs et al. 2020), and FastText embeddings in (Silva et al. 2019).

In addition to feature extraction, feature selection also has an important role in analyzing social engagements (Vora et al. 2019). Using parameter tuning and model selection to choose important features from a pool of over 200 features results in high performance in engagement prediction (Schifferer et al. 2020). In addition to the hand-crafted categorical and numerical features, text sequence embeddings obtained from deep language models, such as BERT (Devlin et al. 2019), are utilized with the Attention algorithm (Vaswani et al. 2017) to calculate the similarity between target tweet and user's previous tweets (Volkovs et al. 2020). However, there is still a lack of feature understanding for different social engagements in online social networks. We resolve this by providing a comprehensive feature analysis based on users and tweets.

2.2 Social engagements

Social engagements can be considered in terms of online social networks' structural and non-structural properties. Structural properties are related to network topology. A common problem in this area is link prediction that handles missing connections by predicting future relationships in complex networks (Martínez et al. 2016). Traditional methods propose to utilize neighbors of nodes to find similarities (Martínez et al. 2016), while embedding-based methods employ network embeddings that encode nodes into vector representations (Zhao et al. 2021). There are also efforts to simulate online social networks that focus on not only links but also users (Ryczko et al. 2017; Chung et al. 2019). Non-structural properties of online social networks are based on user and tweet features. Example applications are user profile with location and demographic information (Bergsma et al. 2013), user influence (Almgren and Lee 2016), and tweet content (Cheng et al. 2010).

There is a lack of studies to examine structural and non-structural features to understand the driving factors behind social engagements. Our study can be placed in between structural and non-structural approaches. The links in network topology infer social engagements, while user and text features provide additional information for why users engage with others.

2.3 Implications of social engagements

Understanding social engagements can provide new opportunities for a better recommendation of the content that might be interesting to the users. Collaborative filtering, content-based recommendation, and hybrid approaches are the main solutions in tweet recommendation (Kywe et al. 2012). Another implication is the ability of social engagements to have positive or harmful consequences for society. For example, analysis of online identities and features of Twitter help uncover the role of self-identified activists in social justice movements (Choi et al. 2020). On the contrary, social network analysis and topic modeling reveal the role of social media interactions, specifically retweets, during the protests of COVID-19 measures that would threaten public health and safety (Haupt et al. 2021). Unlike other studies, we follow a systematic approach to shed light on the main characteristics of social engagements in a more general and domain-independent setting, including over 14 million multilingual tweets.

3 Understanding social engagements

We illustrate the details of understanding social engagements in Fig. 2. We analyze social engagements in terms of two feature categories: user and tweet. User features are based on the author of a targeted tweet (i.e., engagee) and the candidate user to engage with the target tweet (i.e., engager). Tweet features are based on a targeted tweet's meta attributes and contents. In order to analyze feature performances, we employ a supervised learning task of engagement prediction for four engagement types (i.e., like, retweet, reply, and quote). In this section, we first explain the process of extracting the features. We then explain the traditional and neural models that we employ to learn prediction models.

3.1 Feature extraction

We categorize the features into user and tweet features, listed in Table 1. Each group is divided into subgroups to reflect a particular aspect of that group. Before the training process, all features are normalized according to their respective values.

3.1.1 User features

User features represent information related to the author of a tweet, called *engagee*, and the user who would interact with this tweet, called *engager*. We extract user's meta attributes, e.g., being influential and verified user, since one could interact with a tweet based on the popularity of its author. In order to reflect the activity, social, and time patterns of engager behavior, we extract prior and conditional probabilistic features based on the previous engagements using Bayesian modeling (Manning et al. 2008). We list the details of user features as follows:

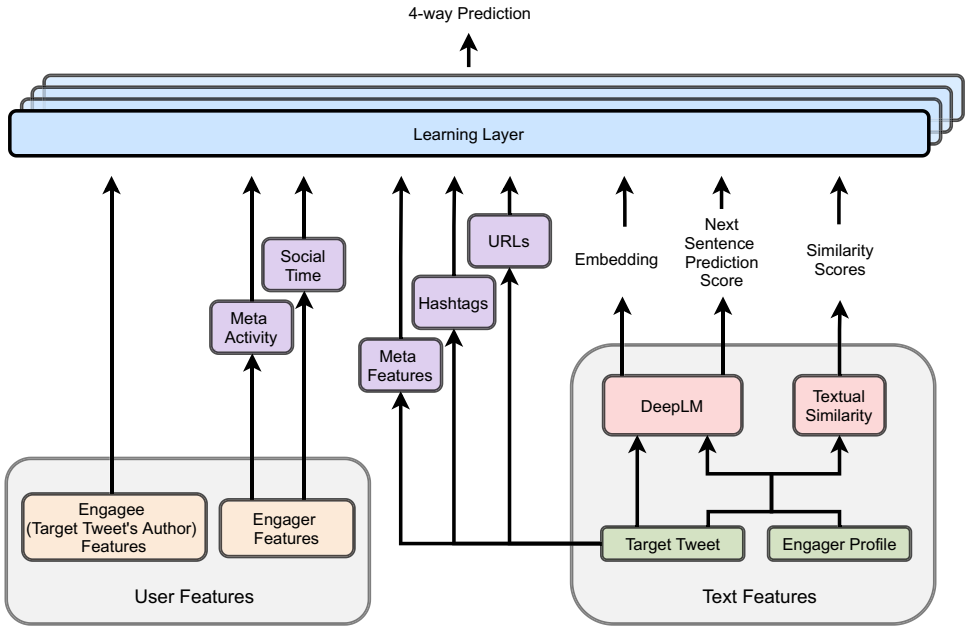

Fig. 2 An illustration of user and tweet features for supervised social engagement prediction

Table 1 The list of the features that we extract and analyze in this study

Group			Feature name	Short description	Range
User	Engagee	Meta	Account Age	How long ago engagee created the account	[0, ∞)
			Influential	The ratio between followees and followers of engagee	(-∞, ∞)
			Verified	Verification status of tweet author.	True of false
	Engager	Meta	Influential	The ratio between followees and followers of engager	(-∞, ∞)
			Verified	Verification status of engager	True or false
		Activity	Conditional Activity	Prob. of engager being interacted with tweet author	[0, 1]
			Prior Activity	Prob. of engager having any engagement	[0, 1]
		Social	Conditional Social	Prob. of engagement between the given engager and previous engagers of the given tweet	[0, 1]
		Time	Conditional Time	Prob. of engager having any engagement in time slot	[0, 1]
Tweet	Meta	Meta	Media	Media type if tweet contains	Photo, video, GIF, or none
			Language	Language of tweet	{1,..,66}
			Type	Type of tweet (not true engagement label)	RT, quote, reply, or top
	Content	Hashtag	Hashtag Existence	Indication of existence of any hashtag in tweet.	True or false
			Conditional Hashtag	Prob. of engager having engagements with hashtag	[0, 1]
			Prior Hashtag	Prob. of hashtag being observed	[0, 1]
		URL	URL Existence	Indication of existence of any URL in tweet	True of false.
			Conditional URL	Prob. of engager having engagements with URL	[0, 1]
			Prior URL	Prob. of URL being observed.	[0, 1]
		Text	Length	The total number of BERT tokens in tweet	{0,..,512}
			Embeddings	Pre-trained BERT sequence embeddings	768-dim.
			Sim. (Cos/BOW)	Cosine sim. between tweet and engager profile with BOW	[0, 1]
			Sim. (Cos/TOK)	Cosine sim. between tweet and engager profile with TOK	[0, 1]
			Sim. (Dice/BOW)	Dice sim. between tweet and engager profile with BOW	[0, 1]
			Sim. (Dice/TOK)	Dice sim. between tweet and engager profile with TOK	[0, 1]
			Sim. (BERT NSP)	NSP task between tweet and engager profile	[0, 1]

– *Account Age (Engagee):* Fresh accounts are likely to have spamming behavior (Farooqi and Shafiq 2019). The account age of the tweet author, *accountAge*(*engagee*), is the difference between the timestamp that the tweet is shared, t_{tweet}, and the timestamp that the account of the tweet author is created, $t_{engagee}$, as follows:

$$\text{accountAge}(\text{engagee}) = t_{\text{tweet}} - t_{\text{engagee}} \tag{1}$$

– *Influential (Engagee):* If the tweet author is an influential user, the probability of interactions with this tweet is expected to increase since the tweet would reach a wide audience. The influential ratio of the tweet author, *influential*(*engagee*), is given as follows (Laplace smoothing is applied in case of no followee or follower count):

$$\text{influential}(\text{engagee}) = \log \frac{n_{\text{followee}}}{n_{\text{follower}}} \tag{2}$$

– *Verified (Engagee):* Some user accounts on Twitter have a blue verified badge to represent that the user is authentic, notable, and active. If the tweet author is a verified user, the probability of interactions is expected to increase. This feature has a Boolean value that indicates whether the tweet author has a verified account.
– *Influential (Engager):* If the engager is an influential user, the probability of interactions is expected to increase since the audience of the engager can also be aware of the engaged tweets. Influential ratio of the engager, *influential*(*engager*), is calculated similarly as in Eq. 2.
– *Verified (Engager):* This feature has a Boolean value that indicates whether the engager has a verified account.
– *Conditional Activity (Engager):* If the engager frequently interacts with a specific tweet author, the likelihood of engaging with the same author increases. The conditional activity of the engager is the probability of the interaction between the engager and the tweet author. The conditional activity is given in Eq. 3, where $n_{engager,engagee}$ is the number of times the engager interacted with the tweet author in the past, and $n_{engager}$ is the total number of interactions by the engager.

$$P(\text{engagee}|\text{engager}) = \frac{n_{\text{engager,engagee}}}{n_{\text{engager}}} \tag{3}$$

– *Prior Activity (Engager):* The prior activity of the engager is the probability of the engager having a possible interaction. The prior activity is given in Eq. 4, where $n_{engager}$ is the total number of interactions by the engager, and $n_{engager_i}$ is the number of interactions by $engager_i$ ($i \in M$ where M is the set of all users).

$$P(\text{engager}) = \frac{n_{\text{engager}}}{\sum_i n_{\text{engager}_i}} \tag{4}$$

– *Conditional Social (Engager):* If there is a connection between the engager and the previous engagers of the tweet, the likelihood of the engager interacting with the tweet increases. This feature reflects the engager's group patterns to an extent, defined as the probability of the engagement between the engager and previous engagers of the tweet, given in Eq. 5. $P(engager_i|engager)$ is the conditional probability of interaction between the engager and a previous engager, $engager_i$ ($i \in K$ where K is the set of previous engagers of the tweet), calculated as in Eq. 3.

$$P(\text{engagers}|\text{engager}) = \prod_i P(\text{engager}_i|\text{engager}) \tag{5}$$

– *Conditional Time (Engager):* Users can be more active in online social networks during specific times of the day. The likelihood of interaction increases if the tweet is shared during a time when the engager is more active. We divide a day into six time periods starting from the beginning of the day, each having four hours. The conditional time is the probability of the engager interacting in a given period, given in Eq. 6, where $n_{engager,time}$ is the number of interactions by the engager in the period when the tweet is shared, and $n_{engager}$ is the total number of interactions by the engager.

$$P(\text{time}|\text{engager}) = \frac{n_{\text{engager,time}}}{n_{\text{engager}}} \tag{6}$$

3.1.2 Tweet features

We divide tweet features into meta and content features. The meta attributes of a tweet are descriptive features to assess the importance of that tweet. We consider the following meta attributes.

– *Media:* Tweets containing interesting and popular media elements can attract user interest. This feature captures the information of a media element in the tweet; in terms of image, video, GIF, or no media content at all. We map each media type to a discrete value.
– *Language:* Although users are more likely to interact with tweets in their own language, there may also be interactions in a foreign language. This feature indicates the language of the tweet. In our dataset, tweets are written in 66 different languages. We map each language to a discrete value.
– *Type:* This feature indicates the type of tweet in terms of top-level, retweet, reply, or quote. Twitter conversations

can have cascades (Cheng et al. 2014), i.e., users make decisions sequentially, and discussions can have lots of layers starting from the top-level tweet. The type of tweet is thereby an important signal for capturing information spread. We map each type to a discrete value.

Users are more inclined to engage with content that they are most interested in; thus, tweet content is just as crucial as user behavior. We examine three content features: Hashtags, URLs, and tweet text.

- *Hashtag Existence:* Users are most likely to interact with tweets that contain popular hashtags (i.e., trend topics). This Boolean feature implies the existence of any hashtags in the tweet.
- *Conditional Hashtag:* Users are more likely to engage with tweets with hashtags that they are interested in. The conditional hashtag is the probability of whether the engager interacts with the tweet containing a particular hashtag, considering engager's previous interactions with that hashtag, given in Eq. 7, where $n_{engager,hashtag}$ is the number of times the engager interacted with a given hashtag, and $n_{engager}$ is the total number of interactions by the engager.

$$P(\text{hashtag}|\text{engager}) = \frac{n_{\text{engager,hashtag}}}{n_{\text{engager}}} \tag{7}$$

- *Prior Hashtag:* If a hashtag is unpopular and rarely observed in tweets, the probability of interactions with this hashtag is likely to decrease. The prior hashtag reflects the popularity of a hashtag, given in Eq. 8, where $n_{hashtag}$ is the total number of times the hashtag is shared, and $n_{hashtag_i}$ is the total number of times $hashtag_i$ is shared ($i \in H$ where H is the set of all hashtags).

$$P(\text{hashtag}) = \frac{n_{\text{hashtag}}}{\sum_i n_{\text{hashtag}_i}} \tag{8}$$

- *URL Existence:* Users are more likely to interact with tweets with interesting or popular URLs (links to external web pages). This Boolean feature implies the existence of any URLs in the tweet.
- *Conditional URL:* The conditional URL is the probability of whether the engager interacts with the tweet containing a particular URL, considering engager's previous interactions with that URL, given in Eq. 9, where $n_{engager,url}$ is the number of times the engager interacted with a given URL, and $n_{engager}$ is the total number of interactions by the engager.

$$P(\text{URL}|\text{engager}) = \frac{n_{\text{engager,url}}}{n_{\text{engager}}} \tag{9}$$

- *Prior URL:* The prior URL reflects the popularity of a URL, given in Eq. 10, where n_{url} is the total number of times the URL is shared, and n_{url_i} is the total number of times url_i is shared ($i \in L$ where L is the set of all URLs).

$$P(\text{URL}) = \frac{n_{\text{url}}}{\sum_i n_{\text{url}_i}} \tag{10}$$

In addition to hashtags, URLs, and media content, we utilize tweet text to understand tweet semantics. Users are more likely to engage with semantically rich tweets in context, e.g., useful information or breaking news. We consider the following content features to represent tweet semantics.

- *Embeddings (BERT):* We extract tweet semantics by encoding tweet text with the pre-trained multilingual sequence embeddings (cased mBERT) (Devlin et al. 2019). BERT is a state-of-the-art deep learning language model built on bidirectional contextual representations of words. BERT's text sequence embeddings perform higher effectiveness than traditional methods in several natural language processing tasks (Devlin et al. 2019).
- *Tweet Length:* Compared to shorter tweets, longer tweets are more likely to include more information. The tweet length is the total number of BERT tokens. Tokenization divides a text into meaningful pieces, such as words or subwords (Devlin et al. 2019).

When tweets are relevant to a user's interests, they can be appealing to a user. We refer to the user's interests as the user profile. We measure the similarity between tweet content and engager's profile to estimate the level of engager's interest, with the following features.

- *Similarity (Cos/BOW):* We encode tweets in the bag-of-words model (BOW) and then calculate the Cosine similarity (Cos) measurement (Manning et al. 2008) between the tweet and user profile. Bag-of-words is a document encoding model in which a fixed-length vector represents each document. We use unigram vector representations that consider each word or term independently. If the user profile does not exist, we set the similarity score to zero. The Cosine similarity measurement is given in Eq. 11, where N is the dictionary size, $tweet \in \mathbb{R}^N$ is target tweet vector, $profile \in \mathbb{R}^N$ is engager's profile vector, using TF-IDF (Term Frequency - Inverse Document Frequency) term weightings (Salton and Buckley 1988).

$$\cos(\text{tweet}, \text{profile}) = \frac{\sum_{i=1}^{N} \text{tweet}_i \times \text{profile}_i}{(\sqrt{\sum_{i=1}^{N} (\text{tweet}_i)^2} \times \sqrt{\sum_{i=1}^{N} (\text{profile}_i)^2})} \tag{11}$$

- *Similarity (Cos/TOK):* Instead of using unigrams, as a novel approach, we propose to convert text to the BERT

tokens and then apply the Cosine similarity measurement. BERT tokens are not word embeddings but represent the subwords in a sentence obtained by tokenization. For instance, two contextually similar words in different languages can be matched by BERT tokens, e.g., the tokens of *coffee* in English can overlap with the tokens of *Kaffee* in German.

- *Similarity (Dice/BOW):* We encode tweets in the bag-of-words model and then calculate the Dice similarity measurement (Manning et al. 2008) that considers overlapping terms between an engager profile, $profile \in \mathbb{R}^N$, and a targeted tweet, $tweet \in \mathbb{R}^N$, given in Eq. 12, where n_{common} is the number of common or overlapping terms, n_{tweet} is the number of terms in the target tweet, and $n_{profile}$ is the number of terms in engager's profile. We use TF (Term Frequency) term weighting for Dice/BOW.

$$dice(\text{tweet}, \text{profile}) = \frac{2 \times n_{\text{common}}}{(n_{\text{tweet}} + n_{\text{profile}})} \tag{12}$$

- *Similarity (Dice/TOK):* We apply the Dice similarity measurement between the tweet and user profile based on BERT tokens instead of unigrams.
- *Similarity (BERT NSP):* We do not use the similarity between two sequence embeddings since similarity calculations with BERT embeddings are reported to yield poor results (Reimers and Gurevych 2019). Instead, we adapt the next-sequence prediction task of BERT (Devlin et al. 2019) to measure the similarity between two text sequences. BERT takes two input text sequences and outputs a probability score that estimates whether the latter follows the former in natural text. Given an engager profile as the first sequence, *profile*, and a targeted tweet as the second sequence, *tweet*; the next-sequence probability, *NSP*(*tweet*, *profile*), is given in Eq. 13, where W is a weight matrix, C is [CLS] token embedding provided by BERT that encodes the classification task, and b is a bias vector of the fully connected layer following C. Input is encoded in the format of [CLS] *profile* [SEP] *tweet*, where [SEP] denotes a separator between two sequences, so that the target tweet is a possible successive sequence of the user profile. We employ the pre-trained multilingual BERT model (cased mBERT).

$$NSP(\text{tweet}, \text{profile}) = \text{softmax}(W \times C + b) \tag{13}$$

We assume that users' previously engaged tweets can construct user profiles. Since the number of engagements per user is approximately 0.72 in the dataset that we use in the experiments (there are approximately 7.18m engaged tweets and 9.99m engagers), we take the user's most recent engaged tweet as the profile. Besides, in the preliminary experiments, we observe deterioration in the prediction performance when more than one engagement is considered, possibly due to the noise created by multiple tweets.

3.2 Learning social engagement models

Assume that $x \in \mathbb{R}^N$ is a data instance, where N is the length of the feature set, and $y \in \{0, 1\}$ the corresponding true label, representing if a particular engagement type occurs. The task is to predict the probability of a possible social engagement, $P(y^k|x^k)$, for each engagement type $k \in \{like, retweet, reply, quote\}$, where $P \in [0, 1]$.

Considering that different machine learning algorithms can perform differently when exploiting information from features, we utilize two predictors, Light Gradient Boosting Machine (LightGBM) (Ke et al. 2017) and Multilayer Perceptron (MLP). While LightGBM uses boosting, MLP utilizes regularization to achieve better generalization. We have a separate predictor model for each engagement type, trained in a supervised way using the same dataset, but target labels change according to engagement type.

3.2.1 Light gradient boosting machine

Decision tree-based learning algorithms are fast and effective for prediction tasks, specifically gradient boosting machine (GBM) (Friedman 2001). GBM uses the boosting method of ensemble learning. LightGBM is proposed to accelerate the training process of GBM when data dimensionality is high in terms of features and instances (Ke et al. 2017).

LightGBM employs a loss function in training to estimate the quality of predictions in each learning step. We use the mean squared error loss function, $\mathcal{L}_{LightGBM}$, that measures the distance between our predictions and true labels, given as follows:

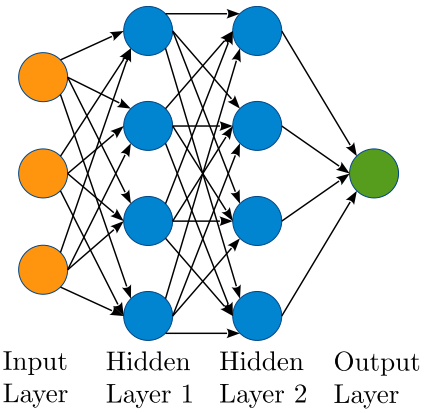

Fig. 3 An illustration for a two layer perceptron

Table 2 The distribution of tweets, engagers, and engagees to the engagement types (like, retweet, reply, and quote). Negative means items having no engagement at all. Positive means items having at least one engagement

Item	Total	Negative	Positive	Like	Retweet	Reply	Quote
Tweet	14,115,364	6,936,057	7,179,307	6,183,928	1,581,922	379,874	108,216
Engager	9,989,359	4,738,788	5,250,571	4,550,366	1,250,552	363,948	105,175
Engagee	4,077,535	1,655,972	2,421,563	2,118,314	722,879	277,995	82,433

$$\mathcal{L}_{\text{LightGBM}} = \frac{1}{M}\sum_{i=1}^{M}(y_i - F(x_i))^2 \tag{14}$$

where x_i is a data instance, y_i is the true engagement label for prediction, M is the number of instances, and $F(x_i)$ is the probability that our model assigns for that instance, $P(y_i|x_i)$. The model finds an optimal regressor estimator, $\hat{F}$, for a given training set, as follows:

$$\hat{F} = \arg\min_{F} \mathbb{E}_{y,x}\mathcal{L}_{\text{LightGBM}} \tag{15}$$

Employing LightGBM is non-trivial as the number of tweets and features increases. Note that we have over 14 million tweets. Since LightGBM gives more importance to the instances with larger gradients and bundles mutually exclusive features, training can be done efficiently.

3.2.2 Multilayer perceptron

Feed-forward neural networks with hidden layers are universal function approximators, making them suitable for engagement prediction. Multilayer perceptron (MLP) is a special type of neural network consisting of input, output, and usually multiple hidden layers. Each neuron of each layer is connected to all neurons of the following layer, followed by a nonlinear activation function that allows complex mapping between inputs and outputs. A softmax layer follows the last layer of MLP to obtain a probability distribution for target classes. We provide a diagram depicting a network architecture with an arbitrary number of parameters in Figure 3. We use the binary cross-entropy loss function, $\mathcal{L}_{MLP}$, to estimate the quality of predictions in each learning step, given as follows:

$$\mathcal{L}_{\text{MLP}} = -\frac{1}{M}\sum_{i=1}^{M}\sum_{c} y_i^c \log P(y_i^c|x_i) \tag{16}$$

where c is a target class (engagement type), x_i is a data instance, y_i^c is the true label of the engagement type c, $P(y_i^c|x_i)$ is the probability that our model assigns for the instance x_i with respect to the engagement type c, and M is the total number of instances. The predictor model's training objective for a parameter set, θ, is given as follows:

$$\theta^* = \arg\min_{\theta} \mathcal{L}_{\text{MLP}} \tag{17}$$

where θ^* represents the optimal parameter set that minimizes the loss function. MLP is a proper choice to exploit a high number of data instances as in our case.

4 Experiments

In this section, we explain the dataset, evaluation metrics, experimental setup, and experimental design. We then report our experimental results.

4.1 Dataset

We use a modified version[1] of the dataset provided by (Belli et al. 2020). Each data instance has a set of properties; including tweet's text and author, candidate engager, and true engagement label. We sample approximately 14 millions of tweets randomly with uniform distribution. The number of unique tweets, engagers, and engagees are given in Table 2 with respect to four engagement types (like, retweet, reply, and quote). Note that a tweet can have more than one engagement type, e.g. both like and retweet at the same time. To like or retweet a tweet, users simply click a button, whereas interactions with reply and quote require writing a text sequence. The large frequency of like and retweet instances compared to reply and quote is probably due to the simplicity of engagement.

In our experiments, we sort tweets by time and allocate the first 80% of the total instances for training and the remaining 20% for evaluation or test. Table 3 gives the distribution of three entities (i.e., engagers, hashtags, and URLs) to the training and test sets. We report the number of overlapping entities between the train and test sets since conditional features match the same entity in both train and test.

[1] The dataset includes publicly available tweet IDs, in compliance with Twitter's Terms and Conditions. The dataset can be accessed from https://github.com/avaapm/Understanding-Social-Engagements.

Table 3 The distribution of engagers, hashtags, and URLs to the training and test sets

Item	All dataset	Only test	Only train	Both train and test
Engager	9,989,359	1,599,330	7,335,108	1,054,921
Hashtag	747,049	120,877	507,120	119,052
URL	952,381	209,029	725,895	17,457

Table 4 The RCE scores obtained by the random and replicate predictors for each engagement type

Method	Like	Retweet	Reply	Quote
Random prediction	−2410.713	−1923.380	−1350.714	−1061.782
Replicate prediction	−0.569	−0.242	−0.004	−0.004

4.2 Evaluation metrics

We evaluate prediction effectiveness by the Relative Cross-Entropy (RCE) score. RCE measures the percentage improvement over the naive or trivial prediction that always predicts the same prediction score based on the mean of the distribution of true labels in the training set. Assume that the cross-entropy between true labels and naive prediction is $\mathrm{CE}_{\mathrm{naive}}$, and the cross-entropy between true labels and our prediction is $\mathrm{CE}_{\mathrm{pred}}$; then, the RCE score is calculated as follows:

$$\mathrm{RCE} = \frac{\mathrm{CE}_{\mathrm{naive}} - \mathrm{CE}_{\mathrm{pred}}}{\mathrm{CE}_{\mathrm{naive}}} \times 100 \tag{18}$$

The higher the RCE score, the better the prediction performance. By reporting the RCE scores, we aim to find the discriminative features for social engagements that perform better than a trivial prediction with a confidence estimate.

The Precision-Recall Area Under Curve (PR-AUC) score is another evaluation metric for prediction performance. However, PR-AUC ignores the potential differences in the confidence scores of predictions. Other studies also note PR-AUC's incompetence (Zhao et al. 2021; Volkovs et al. 2020; Schifferer et al. 2020), and the RecSys 2021 Challenge does not use PR-AUC as an evaluation metric while keeping RCE (Anelli et al. 2021). We thereby analyze important and failed features based on the RCE scores. We provide the PR-AUC scores in Supplementary Material for further investigations.

In addition to RCE and PR-AUC, we evaluate sparsity and cold-start in order not to attribute any data-dependent failure to the features. Sparse data typically have lots of empty elements or zero values. The sparsity ratio is given in Eq. 19; where n_{zero} is the number of instances with the corresponding feature having zero value, and n is the total number of instances in the dataset.

$$\mathrm{sparsity} = \frac{n_{\mathrm{zero}}}{n} \tag{19}$$

The cold-start problem occurs when the model cannot provide any inferences for new users or items (Schein et al. 2002). In our case, the cold-start problem occurs when engagers, hashtags, or URLs emerge in the test set for the first time. The cold-start ratio is given in Eq. 20 where n_{test} is the number of entities observed only in the test set (see Table 3), and n_{both} is the number of entities that are present both in the train and test sets.

$$\mathrm{coldstart} = \frac{n_{\mathrm{test}}}{n_{\mathrm{test}} + n_{\mathrm{both}}} \tag{20}$$

The cold-start problem usually causes sparsity. However, the reason for sparsity is not always the cold-start issue. Sparse data can still be observed due to other factors, such as an inadequate number of hashtags or URLs.

4.3 Experimental setup

We employ the Huggingface (Wolf et al. 2020) implementation to extract BERT-based features, the scikit-learn library (Pedregosa et al. 2011) for evaluation, the Pytorch framework (Paszke et al. 2019) for MLP, and Microsoft's implementation (Microsoft 2020) for LightGBM.

The experiments for MLP are conducted on an RTX2080 Ti GPU with an 11-GB memory. We utilize the Adam optimizer (Kingma and Ba 2015) with a learning rate of 0.01 and apply dropout after each layer to regularize the network to prevent overfitting. We utilize a two-layer perceptron as neural network architecture. The number of neurons in each hidden layer and dropout probabilities is determined by a hyperparameter optimization framework, called Optuna (Akiba et al. 2019). The search interval for a number of neurons in hidden layers is [400, 12800] if BERT features are used as input, [4, 128] otherwise. Dropout ratios are searched between 0.2 and 0.7.

The experiments for LightGBM are conducted on a CPU with 48 threads. We employ 50 estimators and select the feature fraction as 0.06, bagging fraction as 0.67, and bagging frequency as 1.00.

4.4 Experimental design

We design four experiments to analyze different aspects of social engagements.

i. The first experiment aims to analyze the effectiveness of the individual features that we list in Sect. 3 with respect to different engagement types (i.e., like,

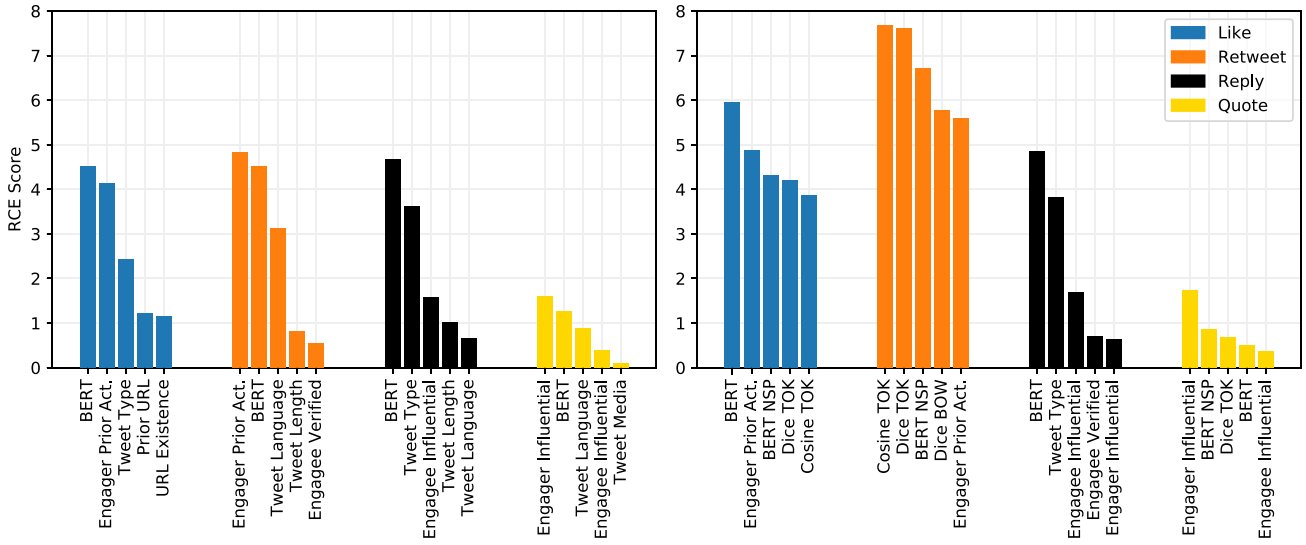

Fig. 4 The top-5 high performing features in terms of RCE for each engagement type using LightGBM (left) and MLP (right)

retweet, reply, and quote). Individual features are the fundamental building blocks for understanding social engagements.

ii. Feature groups containing a set of individual features can expose the main characteristics of social engagements. We compare the effectiveness of feature groups for different engagement types.

iii. We provide a detailed analysis of the textual features. We examine various strategies for measuring the text-similarity between target tweet and engager's profile. We also investigate whether the semantics of the target tweet can be represented by the BERT sequence embeddings (Devlin et al. 2019).

iv. Based on the overall results, we present the ingredients of social engagements; an analogy that describes a recipe for social engagements in terms of which features to prefer and which to avoid. That is, we list important feature sets to understand the main characteristics of social engagements and the failed ones to gain insights into why such features fail in social engagements.

4.5 Experimental results

To provide a starting point for the effectiveness of the features, we first give the prediction results obtained by the random and replicate predictors in Table 4, which are used in similar engagement studies (Chung et al. 2019; Bollenbacher et al. 2021; Kettler 2018). The *random* predictor assigns 25% of instances randomly for each engagement type. Poor results of the random predictor verify that social engagements do not follow a uniform random distribution; in other words, users do not engage with others by chance. The *replicate* predictor assigns the mean of the distribution of engagement types as in the training set. The motivation of the replicate method is that users are expected to follow similar patterns in their social behavior. Note that the replicate method is similar to the naive predictor used to calculate the RCE score. The RCE scores of the replicate method are therefore close to zero.

4.5.1 Individual feature comparison by engagement types

We examine the performances of the extracted features individually. In Fig. 4, the top-5 highest performing features[2] are given for each engagement type using LightGBM and MLP. The effectiveness score is calculated solely based on the reported feature.

Engager's prior activity is consistently observed in the top-5 features for like and retweet, showing that engager's behavior is important in social engagements. Engager's influential feature is, interestingly, the most successful feature for the quote. Engagee's influential feature and tweet type are promising features for the reply. Not only user behavior but also user popularity can be a strong indicator for social engagements.

Textual features, specifically BERT embeddings (Devlin et al. 2019), are useful features across different engagement types and prediction models. Without relying on the user's

[2] In Supplementary Material, we provide the results of all individual features, as well as feature groups, for further investigations.

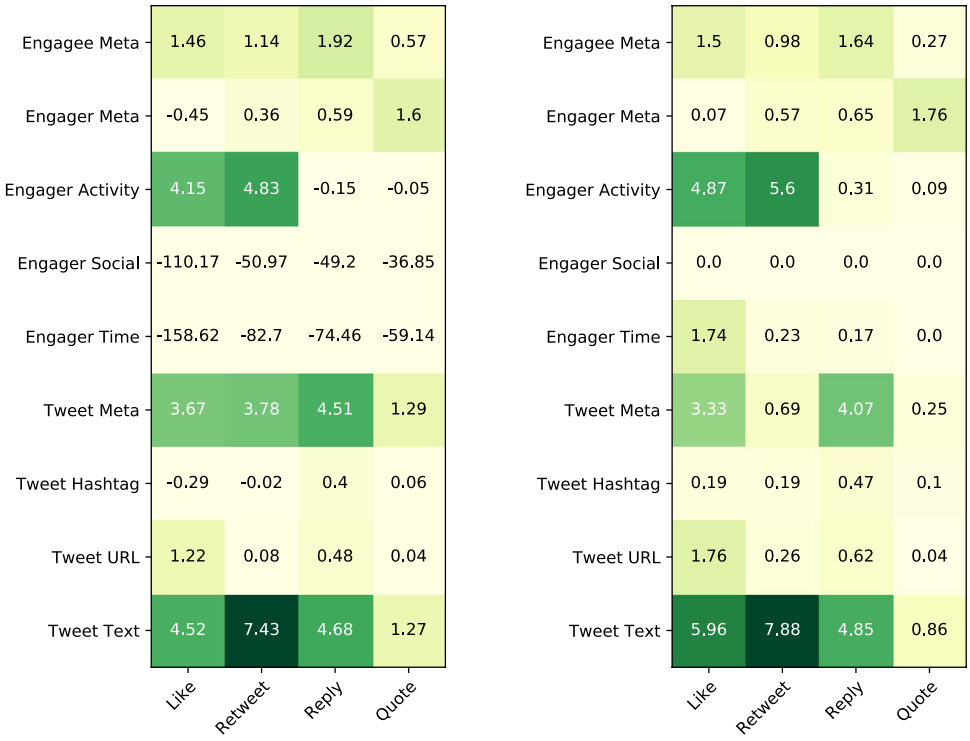

Fig. 5 RCE heatmap of feature groups for each engagement type using LightGBM (left) and MLP (right). Darker color means better performance

activity history, the semantics provided by BERT embeddings can reflect a potential engagement. When MLP is used in Fig. 4, the similarity-based features are also observed in the top features. Since the similarity-based features are based on the user's profile, we argue that the content favored by users is contextually coherent with their previous engagements.

Although our research aims to better understand feature performances in social engagements, we would like to point out the variations in the performances of learning methods. Despite the imbalance between like and other engagement types in the training set, as observed in Table 2, LightGBM can achieve similar RCE scores for like, retweet, and reply. However, the lack of positive instances in the dataset deteriorates the performance of quote predictions. Using parameter optimization, the MLP classifier can improve the performances of like and retweet predictions. We argue that LightGBM does not completely exploit the potential of textual features since textual features are observed more frequently in MLP's top features than in LightGBM's.

4.5.2 Feature group comparison by engagement types

Feature groups are sets of features with common attributes. The comparison of feature groups is given as an RCE heatmap in Fig. 5. We report the highest score that we can achieve in the corresponding group.

Tweet's text has the highest performance for all cases except the quote. Tweet's meta attributes are also useful for all engagement types. Engager's meta attributes, such as being an influential user, have the highest RCE score for the quote. Engager's activity feature plays a significant role in like and retweet engagements, as previously observed in Fig. 4.

Not all features in a group are always useful in understanding social engagements (see Supplementary Material for details). For instance, a tweet's meta attributes (language, media, and type) significantly perform as a group. In contrast, the performance of a tweet's text features mostly relies on BERT sequence embeddings.

4.5.3 A focused analysis: Textual features

The previous experiments reveal that tweet semantics based on textual features are a key factor for understanding social engagements. We thereby focus on textual features in Table 5. The main objective of this experiment is to examine the role of different combinations of textual features in social engagements.

We observe that BERT's sequence embeddings benefit from text length and token similarity (both Cosine and Dice) for like and retweet engagements. The similarity-based features consistently improve the performance of retweet engagements, implying that the user's latest retweet can be representative content for the engager's profile. The quote scores are slightly improved by using BERT with the Dice similarity measurement, whereas the performance of reply engagements is not improved.

Table 5 Performance of different combinations of textual features in terms of the RCE score. Improvements over BERT text sequence embeddings are given in bold

Textual features	Like	Retweet	Reply	Quote
Text Sequence Embeddings (BERT)	5.96	5.22	**4.85**	1.27
BERT + Length	5.83	5.04	4.74	1.23
BERT + Dice BOW	5.68	**5.92**	4.17	**1.32**
BERT + Dice TOK	**6.43**	**6.37**	4.10	**1.44**
BERT + Dice BOW + Cos BOW	3.52	**6.38**	0.00	0.00
BERT + Dice BOW + Cos BOW + BERT NSP	2.77	**7.30**	−0.43	0.02
BERT + Dice BOW + Cos BOW + Length	5.60	**7.50**	3.28	0.37
BERT + Dice BOW + Cos BOW + BERT NSP + Length	3.48	**7.40**	0.78	0.02
BERT + Dice TOK + Cos TOK	5.18	**7.70**	0.00	0.00
BERT + Dice TOK + Cos TOK + BERT NSP	3.26	**7.11**	0.17	0.00
BERT + Dice TOK + Cos TOK + Length	**7.11**	**8.65**	3.49	0.51
BERT + Dice TOK + Cos TOK + BERT NSP + Length	4.57	**7.88**	0.00	-0.01

Table 6 Important features for each engagement type in terms of RCE. Darker color means better performance. Not only individual features but also feature groups are considered. ER stands for engager, and EE for engagee

Like	Retweet	Reply	Quote
5.96 Tweet Text/BERT	7.88 Tweet Text/BERT	4.85 Tweet Text/BERT	1.76 ER Meta/Influ.
	Tweet Text/Dice TOK		ER Meta/Verified
	Tweet Text/Cos TOK		
	Tweet Text/BERT NSP		
	Tweet Text/Length		
4.87 ER Activity/Prior	5.60 ER Activity/Prior	4.51 Tweet Meta/Media	1.29 Tweet Meta/Media
		Tweet Meta/Lang.	Tweet Meta/Lang.
		Tweet Meta/Type	Tweet Meta/Type
3.67 Tweet Meta/Media	3.78 Tweet Meta/Media	1.92 EE Meta/Acc. Age	1.27 Tweet Text/BERT
Tweet Meta/Lang.	Tweet Meta/Lang.	EE Meta/Influ.	
Tweet Meta/Type	Tweet Meta/Type	EE Meta/Verified	
1.76 Tweet URL/Exist.	1.14 EE Meta/Acc. Age	0.65 ER Meta/Influ.	0.57 EE Meta/Acc. Age
Tweet URL/Cond.	EE Meta/Influ.		EE Meta/Influ.
Tweet URL/Prior.	EE Meta/Verified		EE Meta/Verified
1.74 ER Time/Cond.	0.57 ER Meta/Influ.	0.62 Tweet URL/Exist.	– –
1.50 EE Meta/Acc. Age	– –	– –	– –

We implement two encoding types for text similarity, using unigrams of bag-of-words (BOW) and BERT tokens (TOK). The results show that using BERT tokens rather than unigrams is more effective. We attribute the success of tokenization to the fact that the dataset includes multilingual tweets. In addition, tokenization can also solve issues caused by short and noisy text.

4.5.4 Ingredients of social engagements: Important and failed features

We call the list of important and failed features as *the ingredients of social engagements*, referring to an analogy that describes a recipe for social engagements in terms of which features to prefer and which features to avoid.

The following are our criteria for identifying important and failed features. First, we select important features resulting in the highest RCE score in a feature group using either learning algorithm. The features are then ranked according to the RCE scores for each engagement type, and those with less than 0.5 are discarded since their contribution to prediction is less than 0.5%, compared to the naive prediction. Next, we determine the failed features by selecting RCE scores less than 0.1 using both learning algorithms. We empirically set 0.1 as a safety margin.

We list the important features in Table 6. The findings highlight the importance of textual features in all types of engagements, though quote engagements show a lower performance gain. The combination of different textual features is helpful for retweet engagements. Other ingredients include tweet's meta attributes for all engagement types, engager's prior activity for like and retweet, and engager's meta attributes for quote.

Table 7 Failed features for each engagement type along with their sparsity and cold-start ratio. ER stands for engager, and EE for engagee

Features	Sparsity	Cold-start
Like		
Tweet Text / Text Length	0.000	–
ER Social / Conditional	0.536	0.603
Tweet Hashtag / Prior	0.801	0.504
ER Meta / Influential	0.845	0.603
ER Meta / Verified	0.998	–
Tweet Hashtag / Conditional	0.999	0.504
Tweet URL / Conditional	0.999	0.923
Retweet		
ER Social / Conditional	0.536	0.603
Tweet Hashtag / Existence	0.801	–
Tweet Hashtag / Prior	0.801	0.504
Tweet URL / Prior	0.862	0.923
ER Meta / Verified	0.998	–
Tweet URL / Conditional	0.999	0.923
Reply		
ER Social / Conditional	0.536	0.603
Tweet URL / Prior	0.862	0.923
Tweet Text / Cos BOW	0.964	0.603
ER Activity / Conditional	0.994	0.603
ER Meta / Verified	0.998	–
Tweet Hashtag / Conditional	0.999	0.504
Quote		
Tweet Text / Text Length	0.000	–
EE Meta / Age	0.000	–
Tweet Meta / Type	0.000	–
ER Social / Conditional	0.536	0.603
Tweet Hashtag / Exist	0.801	–
Tweet Hashtag / Prior	0.801	0.504
Tweet URL / Existence	0.862	–
Tweet URL / Prior	0.862	0.923
Tweet Text / Dice BOW	0.989	0.623
ER Time / Conditional	0.993	0.603
ER Activity / Conditional	0.994	0.603
Tweet Hashtag / Conditional	0.999	0.504
Tweet URL / Conditional	0.999	0.923
ER Meta / Verified	0.998	–
ER Activity / Prior	1.000	0.603

We list the failed features in Table 7, along with their sparsity and cold-start ratio. The RCE scores are not reported since our concern is not the degree of failure. Rather than failing to model social engagements accurately, the features with high sparsity and cold-start ratio values may have failed due to data-dependent issues. The conditional and prior features of hashtags and URLs are examples of failed features. The results call for a dedicated solution to mitigate such data-dependent issues for further datasets.

The features that are not suffering from sparsity and cold-start can provide useful information to understand social engagements. We argue that the length of a tweet is a trivial feature for like and quote engagements. Tweet type and engagee's account age are not useful features for quote engagements, meaning that the target tweet can be of any type, and its author's account can be of any age in quote engagements. Lastly, we observe no substantial relation between the engager's behavior and the tweet's previous engagers.

5 Discussion

In this section, we provide the main outcomes and social insights, along with the limitations of our research.

5.1 Main outcomes and social insights

We list the following findings obtained from the experimental results.

- The group of textual features is an important component of social engagements, as observed in Fig. 5 and Table 6. Users are likely to engage with tweets based on text semantics regardless of tweet author. This observation is consistent with previous research, e.g., tweet content is a key factor for tweet popularity (Silva et al. 2019) and engagement prediction (Volkovs et al. 2020). Furthermore, social engagements, especially retweets, become more predictable when text semantics are combined with similarity-based features that capture user interests, as observed in Table 5, supporting the importance of user preferences in social engagements (Majmundar et al. 2018).
- Engager's previous activity is strongly correlated with like and retweet engagements, as seen in Fig. 4 and Table 6. Figure 5 also supports that the group of engager's activity features is useful for like and retweet engagements. We thereby argue that users who have actively liked and retweeted information in the past will continue to do so in the future. Since retweeting is a critical tool for information spread in Twitter (Lee et al. 2014), users with a long history of retweeting are more likely to receive and spread false information. On the other hand, the trend of like and retweet engagements is not observed in more complex types of engagements, reply, and quote, probably due to requesting additional content.
- Conditional social feature fails in all engagement types, as observed in Table 7, supporting that the user does not necessarily follow the behavior of other users with whom

she has previously engaged. This observation conflicts with collaborative filtering in recommendation systems (Kywe et al. 2012). Collaborative filtering assumes that users are likely to interact with items (tweets in our case) that similar users have previously interacted with. Our social feature assumes that similar users are found by overlapping engagement history. We argue that overlapping history among users is insufficient to model social engagements due to a long and diverse list of engagement history, yet other methods to find similar users in collaborative filtering might still work.

- Tweet's meta attributes (tweet's language, type, and media contents) are important components of possible engagements, as seen in Fig. 5 and Table 6. Users are likely to engage with tweets in their own language. Tweet type implies user behavior to some extent. One can engage mostly with top-level tweets, while another user can participate in discussion cascades with many replies and quotes. This observation is supported by the results of reply engagements in Fig. 4. Since users are limited to 280 characters, media contents (image, video, or GIF) are important tools for content enrichment to attract social engagements.
- Tweet author's meta attributes are important signals for all engagement types, as observed in Fig. 5 and Table 6. This observation supports that users are likely to engage with popular and trusted tweet authors. Compared to other engagement types, tweet author's meta attributes are more important for reply engagements, and engager's meta attributes for quote engagements. Influential and trusted users can be modeled as the source of information since users participate in discussions by replying to tweet authors. On the other hand, quote can be considered a means of distributing information and triggering discussion by engagers.
- Hashtags are special social media components that enable users to participate in particular topics. Tweet including a hashtag does not strongly relate to getting likes and replies, possibly due to unpopular hashtags and hashtag hijacking (Sedhai and Sun 2015). One can further weigh hashtags to understand their role in engagements according to hashtag importance. On the other hand, we observe that URLs are important for like and reply engagements in Fig. 5 and Table 6. We argue that external content in tweets attracts users for possible like or reply engagement.

5.2 Limitations

We acknowledge a set of limitations to our study. (i) The analysis of features in this study depends on the computational process of feature extraction. One can propose different features to analyze different aspects of social engagements. (ii) The learning of features in this study depends on two state-of-the-art learning algorithms, LightGBM and MLP. One can exploit the features with a different learning algorithm to improve effectiveness. However, our main concern in this study is not finding an optimal prediction model but understanding the dynamics behind social engagements. (iii) Our results are obtained on a Twitter dataset. Therefore, one can validate the generalization of our results to other social media platforms. (iv) We assume that their latest engaged tweet represents the user profile. Understanding the motivation for user profiles participating in different topics is still open research. For instance, user engagements can be examined in the scope of the *fuzzy like* human behavior (Hajarian et al. 2017). (v) We observe the sparsity and cold-start issues, showing an accurate reflection of the unavailability of user history in real-life settings.

6 Conclusion

We extract and analyze substantial features based on users and tweets to understand the main characteristics of social engagements in Twitter. We compare the effectiveness of individual features and feature groups for each engagement type in Twitter, namely like, retweet, reply, and quote. We provide a focused analysis on textual features including BERT's text sequence embeddings and similarity-based features. We lastly list a set of important and failed features in social engagements and discuss the social insights obtained from the experimental results.

The future work can provide more insights on social engagements in different data collections and social media platforms. The features that perform high effectiveness can be utilized in related tasks, such as detecting misinformation and hate speech in online social networks. We also plan to focus on cross-lingual engagements in which user interacts with tweets in a foreign language.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13278-022-00872-1.

Funding No funding was received for conducting this study.

Data availability All data analyzed during this study are included in this published article (Belli et al. 2020). We use a subset of this dataset, which is available at https://github.com/avaapm/Understanding-Social-Engagements.

Code availability The source code used for analysis during the current study is available from the corresponding author on reasonable request.

Declaration

Conflict of interest The authors declare that they have no competing interests.

References

Aggarwal CC, Zhai C (2012) Mining text data. Springer, Boston, MA, USA. https://doi.org/10.1007/978-1-4614-3223-4

Akiba T, Sano S, Yanase T, Ohta T, Koyama M (2019) Optuna: a next-generation hyperparameter optimization framework. In: proceedings of the 25th ACM SIGKDD International conference on knowledge discovery & data mining-KDD, ACM, New York, NY, USA, pp 2623–2631, https://doi.org/10.1145/3292500.3330701

Almgren K, Lee J (2016) An empirical comparison of influence measurements for social network analysis. Soc Netw Anal Min 6(1):1–18. https://doi.org/10.1007/s13278-016-0360-y

Anelli VW, et al. (2020) RecSys 2020 challenge workshop: engagement prediction on Twitter's home timeline. In: Fourteenth ACM conference on recommender systems, ACM, New York, NY, USA, pp 623–627, https://doi.org/10.1145/3383313.3411532

Anelli VW, et al. (2021) RecSys 2021 challenge workshop: Fairness-aware engagement prediction at scale on Twitter's home timeline. In: RecSys '21: fifteenth ACM conference on recommender systems, ACM, pp 819–824, https://doi.org/10.1145/3460231.3478515

Belli L, et al. (2020) Privacy-preserving recommender systems challenge on Twitter's home timeline. arXiv preprint arXiv:2004.13715

Bergsma S, Dredze M, Van Durme B, Wilson T, Yarowsky D (2013) Broadly improving user classification via communication-based name and location clustering on Twitter. Proc. NAACL, ACL, Atlanta, Georgia, pp 1010–1019

Bollenbacher J, Pacheco D, Hui PM, Ahn YY, Flammini A, Menczer F (2021) On the challenges of predicting microscopic dynamics of online conversations. App Net Sci 6(1):1–21. https://doi.org/10.1007/s41109-021-00357-8

Chen J, Pirolli P (2012) Why you are more engaged: Factors influencing Twitter engagement in occupy Wall Street. In: Proceedings of ICWSM, AAAI Press, Dublin, Ireland

Cheng J, Adamic L, Dow PA, Kleinberg JM, Leskovec J (2014) Can cascades be predicted? In: proceedings of WWW, ACM, New York, NY, USA, pp 925–936, https://doi.org/10.1145/2566486.2567997

Cheng Z, Caverlee J, Lee K (2010) You are where you tweet: A content-based approach to geo-locating Twitter users. In: Proceedings of CIKM, ACM, New York, NY, USA, p 759–768, https://doi.org/10.1145/1871437.1871535

Choi JO, Herbsleb J, Hammer J, Forlizzi J (2020) Proc Identity-based roles in rhizomatic social justice movements on Twitter. AAAI Press, ICWSM, California, pp 488–498

Chung W, Toraman C, Huang Y, Vora M, Liu J (2019) A deep learning approach to modeling temporal social networks on Reddit. In: 2019 IEEE International Conference on ISI, pp 68–73, https://doi.org/10.1109/ISI.2019.8823399

Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. In: proceedings of NAACL, ACL, Minneapolis, MN, pp 4171–4186, https://doi.org/10.18653/v1/N19-1423

Farooqi S, Shafiq Z (2019) Measurement and early detection of third-party application abuse on Twitter. In: Proceedings of WWW, ACM, New York, NY, USA, pp 448–458, https://doi.org/10.1145/3308558.3313515

Fortuna P, Nunes S (2018) A survey on automatic detection of hate speech in text. Comput Surv 51(4):1–30. https://doi.org/10.1145/3232676

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232. https://doi.org/10.1214/aos/1013203451

Hajarian M, Bastanfard A, Mohammadzadeh J, Khalilian M (2017) Introducing fuzzy like in social networks and its effects on advertising profits and human behavior. Comput Hum Behav 77(C):282–293. https://doi.org/10.1016/j.chb.2017.08.046

Hajarian M, Bastanfard A, Mohammadzadeh J, Khalilian M (2019) A personalized gamification method for increasing user engagement in social networks. Soc Netw Anal Min 9(1):47:1-47:14. https://doi.org/10.1007/s13278-019-0589-3

Haupt MR, Jinich-Diamant A, Li J, Nali M, Mackey TK (2021) Characterizing Twitter user topics and communication network dynamics of the Liberate movement during COVID-19 using unsupervised machine learning and social network analysis. Online Soc Net Med. https://doi.org/10.1016/j.osnem.2020.100114

Ke G, et al. (2017) LightGBM: a highly efficient gradient boosting decision tree. NIPS, Curran Associates Inc 30:3146–3154

Kettler B (2018) Socialsim project. https://www.darpa.mil/program/computational-simulation-of-online-social-behavior. Accessed 3 January 2022

Kingma DP, Ba J (2015) Adam: A method for stochastic optimization. In: ICLR, San Diego, CA, USA

Kywe SM, Lim EP, Zhu F (2012) A survey of recommender systems in Twitter. In: Soc. Inf., Springer Berlin Heidelberg, Berlin, Heidelberg, pp 420–433, https://doi.org/10.1007/978-3-642-35386-4_31

Lee K, Mahmud J, Chen J, Zhou M, Nichols J (2014) Who will retweet this? Automatically identifying and engaging strangers on Twitter to spread information. In: proceedings of the 19th international conference on Intelligent User Interfaces, ACM, New York, NY, USA, pp 247–256, https://doi.org/10.1145/2557500.2557502

Majmundar A, Allem JP, Boley Cruz T, Unger JB (2018) The why we retweet scale. PLOS ONE 13(10):1–12. https://doi.org/10.1371/journal.pone.0206076

Manning CD, Schütze H, Raghavan P (2008) Introduction to information retrieval. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511809071

Martínez V, Berzal F, Cubero JC (2016) A survey of link prediction in complex networks. Comput Surv. https://doi.org/10.1145/3012704

Microsoft (2020) LightGBM. https://github.com/microsoft/LightGBM. Accessed 3 January 2022

Paszke A, et al. (2019) Pytorch: an imperative style, high-performance deep learning library. NIPS, Curran Associates Inc 32:8026–8037

Pedregosa F, et al. (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12(85):2825–2830

Reimers N, Gurevych I (2019) Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In: Proceedings of 2019 EMNLP-IJCNLP, ACL, Hong Kong, China, pp 3982–3992, https://doi.org/10.18653/v1/D19-1410

Rui H, Liu Y, Whinston A (2013) Whose and what chatter matters? The effect of tweets on movie sales. Decis Support Syst 55(4):863–870. https://doi.org/10.1016/j.dss.2012.12.022

Ryczko K, Domurad A, Buhagiar N, Tamblyn I (2017) Hashkat: large-scale simulations of online social networks. Soc Netw Anal Min 7(1):4

Salton G, Buckley C (1988) Term-weighting approaches in automatic text retrieval. Inf Process Manag 24(5):513–523. https://doi.org/10.1016/0306-4573(88)90021-0

Savargiv M, Bastanfard A (2013) Text material design for fuzzy emotional speech corpus based on Persian semantic and structure. In: International conference on fuzzy theory and its applications (iFUZZY), pp 380–384, https://doi.org/10.1109/iFuzzy.2013.6825469

Schein AI, Popescul A, Ungar LH, Pennock DM (2002) Methods and metrics for cold-start recommendations. In: Proc. 25th Int. ACM SIGIR Conf. Res. Develop. Inf. Ret., ACM, New York, NY, USA, pp 253–260, https://doi.org/10.1145/564376.564421

Schifferer B, et al. (2020) GPU accelerated feature engineering and training for recommender systems. In: Proceedings of the recommender systems challenge 2020, ACM, New York, NY, USA, pp 16–23, https://doi.org/10.1145/3415959.3415996

Sedhai S, Sun A (2015) HSpam14: A collection of 14 million tweets for hashtag-oriented spam research. In: Proc. 38th Int. ACM SIGIR Conf. Res. Develop. Inf. Ret., ACM, New York, NY, USA, p 223–232, https://doi.org/10.1145/2766462.2767701

Silva MG, Domínguez MA, Celayes PG (2019) Analyzing the retweeting behavior of influencers to predict popular tweets, with and without considering their content. In: Inf. Manage. and Big Data, Springer International Publishing, Cham, pp 75–90, https://doi.org/10.1007/978-3-030-11680-4_9

Vaswani A, et al. (2017) Attention is all you need. In: NIPS, Curran Associates, Inc., vol 30

Volkovs M, et al. (2020) Predicting twitter engagement with deep language models. In: Proceedings of the recommender systems challenge 2020, ACM, New York, NY, USA, pp 38–43, https://doi.org/10.1145/3415959.3416000

Vora M, Chung W, Toraman C, Huang Y (2019) SimON-Feedback: An iterative algorithm for performance tuning in online social simulation. In: 2019 IEEE Conf. Intell. and Security Inform., pp 13–17, https://doi.org/10.1109/ISI.2019.8823438

Vraga EK, et al. (2018) Cancer and social media: a comparison of traffic about breast cancer, prostate cancer, and other reproductive cancers on Twitter and Instagram. J Health Comm 23(2):181–189. https://doi.org/10.1080/10810730.2017.1421730

Wolf T, et al. (2020) Transformers: State-of-the-art natural language processing. In: Proceedings of EMNLP, ACL, Online, pp 38–45, https://doi.org/10.18653/v1/2020.emnlp-demos.6

Zamani H, Shakery A, Moradi P (2014) Regression and learning to rank aggregation for user engagement evaluation. In: Proceedings of the 2014 recommender systems challenge, ACM, New York, NY, USA, p 29-34, https://doi.org/10.1145/2668067.2668077

Zhao F, Zhang Y, Lu J (2021) Shortwalk: an approach to network embedding on directed graphs. Soc Netw Anal Min 11(1):1–12. https://doi.org/10.1007/s13278-020-00714-y

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.