Expert Systems With Applications 169 (2021) 114400

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Predicting tweet impact using a novel evidential reasoning prediction method

Lucía Rivadeneira *, Jian-Bo Yang, Manuel López-Ibáñez

Alliance Manchester Business School, University of Manchester, Booth St West, M15 6PB, UK

ARTICLE INFO

Keywords:
Evidential reasoning rule
Belief rule-based inference
Maximum likelihood data analysis
Twitter
Retweet
Prediction

ABSTRACT

This study presents a novel evidential reasoning (ER) prediction model called MAKER-RIMER to examine how different features embedded in Twitter posts (tweets) can predict the number of retweets achieved during an electoral campaign. The tweets posted by the two most voted candidates during the official campaign for the 2017 Ecuadorian Presidential election were used for this research. For each tweet, five features including type of tweet, emotion, URL, hashtag, and date are identified and coded to predict if tweets are of either high or low impact. The main contributions of the new proposed model include its suitability to analyse tweet datasets based on likelihood analysis of data. The model is interpretable, and the prediction process relies only on the use of available data. The experimental results show that MAKER-RIMER performed better, in terms of misclassification error, when compared against other predictive machine learning approaches. In addition, the model allows observing which features of the candidates' tweets are linked to high and low impact. Tweets containing allusions to the contender candidate, either with positive or negative connotations, without hashtags, and written towards the end of the campaign, were persistently those with the highest impact. URLs, on the other hand, is the only variable that performs differently for the two candidates in terms of achieving high impact. MAKER-RIMER can provide campaigners of political parties or candidates with a tool to measure how features of tweets are predictors of their impact, which can be useful to tailor Twitter content during electoral campaigns.

1. Introduction

While the volume of users and information continues to expand globally on Twitter, content producers find themselves in an increasingly contested environment when seeking to capture attention and spread their influence. This issue has gained particular relevance in electoral contexts. Simultaneously, in the scholarly realm, existing research has evidenced the critical role that Twitter plays in presidential elections. Indeed, there has been a dramatic increase in the use of Twitter for electoral purposes, and a progressive supplanting of traditional media platforms (Enli, 2017), especially when candidates feature limited political experience or lack support from influential actors from the political realm (Wang, Luo, Niemi, & Hu, 2016). This is evident at least for the period between Barack Obama's victory in the 2008 US Presidential race and the latest 2016 US Presidential election in which Donald J. Trump was elected president (Clarke & Grieve, 2019; Enli, 2017). Of relevance to this, several lines of inquiry have emerged that may contribute to identifying, for instance, who is reached on Twitter, who composes the intended audience, whether a message can have influence on political behaviour and preferences, and what impact a tweet can make.

However, although extensive research has been conducted on the role that Twitter plays during electoral campaigns, disagreements remain among scholars about the causes behind high and low Twitter impact. For example, when aiming for a high volume of retweets, there are disagreements about the convenience of co-producing content with influential Twitter users. Also, the measurement of tweet impact has been subject to controversy, being counts of account-followers, tweet favourites, or retweets the most frequently used. Yet, the link between patterns of tweets and retweet counts remains as an important subject of inquiry, particularly in connection with the widely established claim that viral information reflects public opinions and political preferences (Grover, Kar, Dwivedi, & Janssen, 2019).

Following the work of Grčar, Cherepnalkoski, Mozetič, Kralj, and Novak (2017), this paper assumes that the number of retweets embodies the influence of a tweet, and argues that retweets of a candidate's tweet are driven by a combination of content and name value. Therefore, the ability of a tweet to generate high impact is not a "one-size-fits-all"

* Corresponding author.
E-mail address: lucia.rivadeneirabarreiro@manchester.ac.uk (L. Rivadeneira).

https://doi.org/10.1016/j.eswa.2020.114400
Received 29 October 2019; Received in revised form 24 November 2020; Accepted 27 November 2020
Available online 13 December 2020

approach, but instead is mediated by both the candidate's profile and the content.

In the social media context, previous research has not fully addressed the mechanisms that underpin retweeting behaviour when deliberating about politics. When aiming to model Twitter data, previous work has used traditional machine learning methods such as logistic regression, decision tree, or support vector machine. One limitation of these methods is that they need "sufficiently large sample data to learn predictive models" (Kong et al., 2016, p. 36). However, even in the absence of statistically meaningful data to train a single model, these methods still proceed with the prediction. It is likely, then, that their prediction outcomes might not be fully trusted because of the limitation of sufficient and meaningful data. The second limitation is that these approaches are often considered as black-box systems. This means that they provide very limited information about how the inference processes are carried out. Therefore, often, when using machine learning, explanations are not fully reliable and results can be misleading, which can cause negative impact in real-world situations (Rudin, 2019).

This study seeks to address these issues by proposing a novel evidential reasoning (ER) predictive model grounded on likelihood data analysis, evidence-based probabilistic inference, and interpretable machine learning (Liu, Chen, Yang, Xu, & Liu, 2019; Yang & Xu, 2017). The ER model comprises two approaches called MAKER-RIMER. MAKER stands for maximum likelihood evidential reasoning (Yang & Xu, 2017) and RIMER for belief rule-based inference methodology using the ER approach (Yang, Liu, Xu, Wang, & Wang , 2007). The model aims at maximising the use of available data by splitting a model (MAKER) into sub-models (partial MAKER) for analysis and then combine them back together (RIMER). For this purpose, the MAKER-RIMER model presented in this study aims to examine how different features embedded in personal tweets may function as predictors of the impact that tweets can produce in terms of their number of retweets. Five potential features, including type of tweet, emotion, uniform resource locator (URL), hashtag, and the moment in the timeline (date), are identified and coded to determine the propensity of a tweet to be high or low impact.

The foreseen advantages of using MAKER-RIMER in this study are twofold. First, it is likely that, given the number of tweets available for analysis and that the input variables do not have all value combinations, it performs better than other machine learning approaches as it is recursive in nature and can deal with incomplete datasets without deleting data or imputing data. This will be validated by comparison against other machine learning approaches. Second, the MAKER-RIMER model is purely data driven (Yang & Xu, 2017). This means two things: 1) that it performs only with existing data, even when datasets are incomplete. The other machine learning methods, instead, often deal with incomplete datasets by relying on intuition, or by using data augmentation techniques (Wong, Amer, Maul, Liao, & Ahmed, 2020). And 2), that the weights that MAKER-RIMER would assign to the different parameters can show how the input variables influence the outcome, for which it is said to be a transparent model (Kong, Xu, Yang, Wang, & Jiang, 2020). MAKER-RIMER is also an interpretable approach. Interpretable means that the decisions made by the algorithm during the inference process are explicable to users (Laugel, Lesot, Marsala, Renard & Detyniecki, 2018).

Therefore, in this study, we provide a method to identify relevant features for a particular candidate and we fully expect that those features will change for different electorates, countries and evolve over time. This is one of the reasons that a fully interpretable and transparent model such as MAKER-RIKER is useful to interpret the results when applied to different data.

To validate the MAKER-RIMER model, an analysis of tweets produced by the two most popular candidates to win the 2017 Ecuadorian Presidential election was conducted. In addition to the ER methodology, four traditional machine learning methods, namely logistic regression, Naïve Bayes, decision tree, and support vector machine, are also evaluated for prediction purposes to compare their performance based on misclassification errors (MCE) using the same datasets for training and testing purposes as in the ER model. Lastly, the model described in this study can facilitate the identification of features of tweets that could lead to obtain high number of retweets, which is crucial for politicians to spread influence across their target audience.

The rest of the study takes the form of seven sections: In Section 2 the concepts of the ER rule are introduced, in which the MAKER and RIMER frameworks are presented. Section 3 reviews related work about predictive models using Twitter data. The methodology that leads this study is described in Section 4, where the case study is introduced and the variables that are part of the model are presented. In Section 5 the case study using the MAKER-RIMER approach is conducted to predict the impact of tweets, as well as the application of different machine learning approaches for the same purpose. Finally, Section 6 shows the results and discussion, while the conclusion is presented in Section 7.

2. Brief introduction to the ER rule

The ER rule is based on the Dempster-Shafer (D-S) theory (Dempster, 1967; Shafer, 1976) and Bayesian probability theory. ER means reasoning with evidence (Benalla, Achchab, & Hrimech, 2020). The ER rule is a probabilistic reasoning process to combine multiple pieces of independent evidence considering both reliability and weight of the evidence (Xu et al., 2020). A piece of evidence is independent if the information it contains does not depend on other evidence (Yang & Xu, 2013), and it is defined as a probability distribution over a set of mutually exclusive and collectively exhaustive propositions. Mutually exclusive means that propositions, which are the possible outcomes, cannot occur simultaneously. Collectively exhaustive, on the other hand, means that at least one of the possible events must occur.

Weight and reliability play an important role when considering the ER rule. Evidence weight, denoted by w_j, refers to the relative importance of the evidence, which can depend on the source and the way evidence is acquired (Yang & Xu, 2014). Evidence reliability, represented by r_j, denotes the ability of the information source to provide correct assessment to a problem (Fu, Xue, Chang, Xu, & Yang 2020). If all pieces of evidence, which are the observations obtained from the data, are acquired and measured in the same joint space, weight equals reliability, otherwise both need to be generated independently (Yang & Xu, 2014).

The ER rule consists of two parts: the bounded sum of the individual support of two pieces of independent evidence for each proposition, and the orthogonal sum of their collective support for each proposition, which makes it possible to combine different pieces of evidence regardless of their order and without affecting the final results (Yang & Xu, 2013, 2014).

The ER rule has been applied in different disciplines and applications. For example, Zhu, Yang, Xu, and Xu (2016) have used ER to propose a model for monitoring asthma and manage its treatment in children, Xiaobin Xu et al. (2017) for data classification tasks across different kinds of database. Likewise, ER has been consistently applied in assessing navigational risk (Zhang, Yan, Zhang, Yang, & Wang , 2016) and medical quality assessment (Kong, Xu, Yang, & Ma, 2015). These examples suggest the versatility of ER for working with qualitative and quantitative data. The following subsections elaborate further on MAKER and RIMER frameworks.

2.1. The MAKER framework

MAKER, which is proposed by J.-B. Yang and Xu (2017), is a methodological framework to combine multiple pieces of evidence under condition of uncertainty, such as randomness, inaccuracy, and ambiguity, for inferential modelling and analysis.

The MAKER framework demands the generation of joint frequency tables of the input variables, to then calculate basic probabilities or normalised likelihoods using Eq. (1). In these calculations, the

likelihood principle and the Bayesian principle need to be followed (Yang & Xu, 2014). When given two pieces of evidence $e_{i,l}$ and $e_{j,m}$ acquired from two variables x_l and x_m, at $x_l = x_{i,l}$ and $x_m = x_{j,m}$ respectively, their joint likelihood for proposition θ is represented by $c_{\theta,il,jm}$, which is the probability that both $x_{i,l}$ and $x_{j,m}$ are observed given proposition θ. Note thatθ can be a single proposition or a subset of propositions. Then, the normalised likelihood is defined as follows (Yang & Xu, 2014, 2017)

$$p_{\theta,il,jm} = c_{\theta,il,jm} / \sum_{A \subseteq \Theta} c_{A,il,jm} \forall \theta \subseteq \Theta \tag{1}$$

where $\Theta = \{h_1, h_2, \cdots, h_N\}$ is defined as a frame of discernment and refers to a set of mutually exclusive and collectively exhaustive propositions.

Following the joint basic probability, the interdependence index is calculated to capture the statistical relationship between two pieces of evidence $e_{i,l}(A)$ and $e_{j,m}(B)$, and it is represented by $\alpha_{A,B,i,j}$. The interdependence index measures how strongly one input variable is related to another input variable. Since this index has been obtained from a space where basic probability is acquired as normalised likelihood, it needs to be scaled to ordinary likelihood (Yang & Xu, 2017). The formula to calculate the interdependence index is shown in Eq. (2), while Eq. (3) shows its properties

$$\alpha_{A,B,i,j} = \begin{cases} 0 & \text{if} p_{A,i,l} = 0 \text{or} p_{B,j,m} = 0 \\ p_{A,B,il,jm} / \left(p_{A,i,l} p_{B,j,m}\right) & \text{otherwise} \end{cases} \tag{2}$$

$$\alpha_{A,B,i,j} = \begin{cases} 0 & \text{if} e_{i,l}(A) \text{and} e_{j,m}(B) \text{aredisjoint} \\ 1 & \text{if} e_{i,l}(A) \text{and} e_{j,m}(B) \text{areindependent} \end{cases} \tag{3}$$

After calculating the interdependence index, the next step is to generate the MAKER framework. In the MAKER framework, two pieces of evidence are combined to generate the combined support for proposition θ, as shown next. Suppose two pieces of evidence $e_{i,l}$ and $e_{j,m}$ are independent, the combined probability that proposition θ is jointly supported by both pieces of evidences is denoted by $p(\theta)$ as given by Eq. (4)

$$p(\theta) = \begin{cases} 0 & \theta = \varnothing \\ m_\theta / \sum_{C \subseteq \Theta} m_C & \theta \subseteq \Theta \end{cases} \tag{4}$$

where m_θ measures the combined probability mass for θ from both pieces of evidence and is generated as the bounded sum of the individual support for θ from both $e_{i,l}$ and $e_{j,m}$, and the orthogonal sum of their joint support with their interdependency and joint reliability taken into account, as shown in the recursive formula in Eq. (5)

$$m_\theta = \left[\left(1 - r_{j,m}\right) m_{\theta,i,l} + \left(1 - r_{i,l}\right) m_{\theta,j,m}\right] + \sum_{A \cap B = \theta} \gamma_{A,B,i,j} \alpha_{A,B,i,j} m_{A,i,l} m_{B,j,m} \tag{5}$$

where $r_{i,l}$ is the reliability of the evidence $e_{i,l}$. $\gamma_{A,B,i,j}$ is the ratio of the joint reliability over the product of the individual reliabilities of the two pieces of evidence $e_{i,l}$ and $e_{j,m}$ given that $e_{i,l}$ points to proposition A and $e_{j,m}$ to proposition B with $A \cap B = \theta$. Eq. (5) should be first applied before Eq. (4) is implemented.

2.2. The RIMER framework

RIMER is established as an extension of the traditional IF-THEN rules to belief rules (Yang et al., 2006). A belief rule is defined as a knowledge representation of information under uncertainty of vagueness or incompleteness (Chen et al., 2011). In RIMER, an initial belief rule base (BRB) is constructed consisting in beliefs rules based on the knowledge of experts and experiences from users (Yang et al., 2006). Belief rule, denoted as R_k, is compounded of rule weights, antecedent attribute weights, and consequent belief degrees, and it is described as follows (Kong et al., 2015)

$$\begin{aligned} &R_k : \text{if} A_1^k \wedge A_2^k \wedge \cdots \wedge A_{T_k}^k, \\ &\text{then}\{(D_1, \beta_{1k}), (D_2, \beta_{2k}), \cdots, (D_N, \beta_{Nk})\} \left(\beta_{jk} \geq 0, \sum_{j-1}^{N} \beta_{jk} \leq 1\right), \\ &\text{withruleweight}\theta_k, \text{andattributeweights}\delta_1, \delta_2, \cdots, \delta_{T_k}, \\ &k \in \{1, \cdots, L\} \end{aligned} \tag{6}$$

where $A_i^k (i = 1, \cdots, T_k)$ is the referential category of the ith antecedent attribute in the kth rule, T_k is the number of antecedent attributes used in the kth belief rule, $\beta_{jk} (j = 1, \cdots, N; k = 1, \cdots, L)$ is the assigned belief degree to consequent D_j which is used to describe input information that can be initially given by experts as subjective probability, $\delta_i (i = 1, \cdots, T_k)$ is the antecedent attribute weight that represents the relative importance of the ith attribute, and θ_k is the rule weight representing the relative importance of the kth rule. L represents the number of all belief rules in the rule base, and N is the number of all antecedent attributes used in the kth rule.

The activation weight, denoted by w_k, is calculated for the kth rule. The activation weight measures the degree to which the packet antecedent A^k in the kth rule is activated by the input variables. The weight of each rule and degrees of belief should be considered. w_k is calculated as follows (Kong et al., 2015)

$$w_k = \frac{\theta_k \alpha_k}{\sum_{j=1}^{L} \theta_j \alpha_j} = \frac{\theta_k \prod_{i=1}^{T_k} \left(\alpha_{i,j}^k\right)^{\bar{\delta}_i}}{\sum_{l=1}^{L} \left[\theta_l \prod_{i=1}^{T_l} \left(\alpha_{i,j}^l\right)^{\bar{\delta}_i}\right]} \text{and} \bar{\delta}_i = \frac{\delta_i}{\max_{i=1,\cdots,T_k} \{\delta_i\}} \tag{7}$$

where $\theta_k (\in R^+, k = 1, \cdots, L)$ is the relative weight of the kth rule, and $\delta_i (\in R^+, i = 1, \cdots, T_k)$ is the relative weight of the ith antecedent attribute that is used in the kth rule. The matching degree, $\alpha_{i,j}^k (i = 1, \cdots, T_k)$, is the belief degree to which the input of the ith antecedent attribute belongs to its jth referential value $A_{i,j}^k$ in the kth rule. This degree can be generated from different perspectives, depending on the nature and availability of the attributes (Yang et al., 2006). The final results are generated by aggregating all rules as described below

$$\mu = \left[\sum_{j=1}^{N} \prod_{k=1}^{L} \left(w_k \beta_{j,k} + 1 - w_k \sum_{i=1}^{N} \beta_{i,k}\right) - (N-1) \prod_{k=1}^{L} \left(1 - w_k \sum_{i=1}^{N} \beta_{i,k}\right)\right]^{-1} \tag{8}$$

where μ measures the degree to which the activation weight and belief degrees play in each rule.

$$\beta_j = \frac{\mu^* \left[\prod_{k=1}^{L} \left(w_k \beta_{j,k} + 1 - w_k \sum_{i=1}^{N} \beta_{i,k}\right) - \prod_{k=1}^{L} \left(1 - w_k \sum_{i=1}^{N} \beta_{i,k}\right)\right]}{1 - \mu^* \left[\prod_{k=1}^{L} (1 - w_k)\right]}, j = 1, \cdots, N \tag{9}$$

where β_j is a function of the belief degrees $\beta_{i,k} (i = 1, \cdots, N, k = 1, \cdots, L)$, the rule weights $\theta_k (k = 1, \cdots, L)$, the attribute weights $\delta_i (i = 1, \cdots, T)$, and the input vector x^*.

3. Related work

In its most abstract form, predictive models refer to the use of mathematical tools intending to predict outcomes in the future, based on observed and assumed facts used as input variables. Predicting an output includes, for example, foretelling future trends in behaviour patterns (Iyer, Zheng, Li, & Sycara, 2019). Predictive models increasingly constitute a key decision-making support tool across a wide range of fields, such as marketing, health services, or fraud detection in the security systems industry. Nowadays, following the emergence and

extensive use of social media platforms, vast amounts of data continuously generated and consumed by users, which contain valuable information about demographic aspects, preferences, and behaviours, are increasingly serving as grounds for predictive modelling (Bigsby, Ohlmann, & Zhao, 2019).

3.1. Predictive models using Twitter data: Retweet analysis

In recent years, there have been a growing number of publications focusing on predictive models based on Twitter content, with the retweet measure being one of them. Retweeting refers to the act of sharing others' tweets within users' networks. The importance of the retweet lies in its ability to act as a dissemination tool, and to validate and engage with other Twitter users (Fan, Jiang, Yang, Zhang, & Mostafavi, 2020). Retweet is equivalent to word-of-mouth (WOM) propagation in the Twitter context (Ananda, Hernández-García, Acquila-Natale, & Lamberti, 2019). It also serves as a metric used to determine the effectiveness, popularity, influence, and level of support of a given tweet or Twitter user (Nesi, Pantaleo, Paoli, & Zaza, 2018; Punjabi et al., 2019; Scurlock, Dolsak, & Prakash, 2020).

Studies of retweeting behaviour have been conducted from different perspectives and with different approaches. For example, Lo, Chiong, and Cornforth (2016) focused on ranking audiences on Twitter, while Abdullah, Nishioka, Tanaka, and Murayama (2017) relied on the identification of retweeters – Twitter users that retweet others' tweets – to understand what prompts Twitter users to retweet. Rather than focusing on individual users, this type of analysis demands focusing on the content that becomes widely shared. This perspective leads to the assumption that retweeting behaviour can be triggered by similarity of interests (Ma, Hu, Zhang, Huang, & Jiang, 2019), as a reciprocity action (Yuan et al., 2016), to show support and agreement publicly (Majmundar, Allem, Boley Cruz, & Unger, 2018), for self-enhancement purposes to appear knowledgeable (Yang, Tufts, Ungar, Guntuku, & Merchant, 2018) or to build and engage in an online community (Soboleva, Burton, Mallik, & Khan, 2017). Hence, understanding the motivations behind retweeting behaviour can be a complex task, but it is key when trying to connect with a target audience to disseminate content and gain influence.

Furthermore, another line of inquiry is concerned with what makes some tweets more likely to be retweeted than others. According to Jalali and Papatla (2019), the propensity of retweeting might be influenced by the position or visibility the tweets have, and by the number of followers Twitter users have. The most influential Twitter users have greater probabilities of gaining a higher number of retweets than others. Jalali and Papatla (2019) also include posting time and sharing similar viewpoints in tweets as influential features for propagating tweets. And, Shi, Hu, Lai, and Chen (2018) state that the presence of URLs and hashtags increases the chance of a tweet to be retweeted.

3.2. Predicting retweets

Retweeting behaviour has been used for testing different predictive models. Some of these models have used tweets retrieved randomly, while others have applied a specific retrieving criterion. Similarly, Twitter-based predictive models have used different machine learning approaches. For example, Oliveira, Costa, Silva, & Ribeiro (2018) and Shi, Lai, Hu, and Chen (2017) attempted to examine retweeting behaviour without specific retrieving criteria. The former study used random forest, and the latter both logistic regression and support vector machine algorithms. Concerning studies based on specific retrieving criterion, retweeting behaviour has been studied in different fields. For example, in marketing Walker. Baines, Dimitriu, & Macdonald (2017) used decision tree, in health Kim, Hou, Han, and Himelboim (2016) applied logistic regression, in journalism Trilling, Tolochko, & Burscher (2017) applied negative binomial regression, and in politics Vijayan and Mohler (2018) used a neural network algorithm. Aiming to focus on an individual perspective, Lee and Xu (2018) and Houston et al. (2020) agreed that tweet propagation is more affected by the way tweets are written than by their topic. And Lee & Lim (2016) concluded that the content of a tweet can have an influence on user reactions.

In addition, Twitter data have been analysed using hybrid intelligent system approaches. Hybrid systems combine different knowledge and learning strategies to solve tasks involving uncertainty and vagueness (Abraham, Han, Al-Sharhan, & Liu, 2016). For retweeting behaviour prediction, hybrid systems have been used to predict the popularity of tweets through a two-layered approach using a Hawkes process (Gao et al., 2019; Mishra, Rizoiu, & Xie, 2016), by combining knowledge representation and reasoning and machine learning approaches (Gallo, Simari, Martinez, & Falappa, 2020), or by proposing a neural network hybrid model (Roy, Suman, Chandra, & Dandapat, 2020).

Previous studies have attempted to predict retweeting behaviour using methods different than machine learning approaches. For example, Ca, Oktay, and Manmatha (2013) sought general correlations between the features of random tweets and the retweet counts. Similarly, Pancer and Poole (2016), during an electoral race, used regression analysis to measure how features in tweets correlate with retweet count for each candidate. Also, Tumasjan, Sprenger, Sandner, & Welpe (2011) attempted to predict election results using Twitter but using a sentiment analysis approach. While these papers address similar problems, our work is different in that it develops a model with unstructured and incomplete data to predict the impact of tweets. Thus, we decided to frame our study as a classification problem instead.

As is observed, there is not a unique or straightforward mechanism to analyse the propensity to retweet. Indeed, approaches seem to differ based on the context and field of application. Although these models provide a deeper understanding of retweeting behaviour, this is mostly restricted to a number of features such as number of followers, URLs, hashtags, or mentions, for which there is a need to continue developing retweeting behaviour models by coding these features in a different way.

3.3. Contribution of this paper

Twitter data can yield effective and powerful indicators of future behaviour for a range of situations and applications. However, much uncertainty still exists around the feasibility of intervening continuously and systematically on social media towards desired outcomes of influence. A primary concern of predictive models is still the capacity to deliver information in a dynamic fashion, which is useful to intervene opportunely in the controllable elements that affect the output variable. So far, predictive models have used metrics related to tweets or their authors in terms of numbers of URLs, hashtags, or followers to predict the likelihood of retweeting. Meanwhile, there are features of tweets that remain unexplored, which can influence the propensity of the public to retweet.

This study contributes to the growing body of research on Twitter data for prediction purposes by proposing a novel model called MAKER-RIMER to predict retweeting behaviour based on the ER rule, which is footed on evidence-based probabilistic inference. The main contribution of the proposed ER model is perhaps its interpretability. It means that the inference process is transparent during the application of the MAKER-RIMER model, and the results are interpretable in the sense that the probability for each outcome, based on its circumstances, is openly readable for decision makers. This model involves the use of codified characteristics embedded in tweets to analyse their influence on the distribution of tweets. The implementation of MAKER-RIMER has been applied to predict traumatic injury outcomes using structured data (Almaghrabi, Xu, & Yang, 2019). However, as far as we are aware, there are no other studies applying MAKER-RIMER on Twitter or any other source of unstructured data.

In addition, for decision-making purposes, the model allows the identification of the characteristics of tweets that make the target audience more susceptible to sharing and distributing tweets. As a

Table 1
Total number of tweets generated by the two candidates, total number of retweets generated by other Twitter users during the elections, and descriptive statistics of retweets.

Candidates	1st round		2nd round	
	Lenin Moreno	Guillermo Lasso	Lenin Moreno	Guillermo Lasso
Total tweets	415	745	235	443
Total retweets	302,026	124,642	149,019	324,560
Min. retweets	13	9	6	54
Max. retweets	3275	1517	2448	5681
Median retweets	646	130	530	554
Mean retweets	727.77	167.30	634.12	732.64
No. followers	125,000	298,000	254,000	313,000
No. followees	25	1,456	26	1,457

result, social media campaigns can be tailored and adapted for political purposes to maximise the effectiveness of their campaigns and increase Twitter engagement with the audience. So, identification of drivers or stimuli in reference to what makes a high impact tweet in terms of retweets seems crucial to developing successful Twitter campaigns.

4. Methodology

This section presents the case study used in this research and provides details of the data collection and analysis processes. The description of the model that constitutes the basis for the prediction of impact of tweets is also detailed.

4.1. Case study

The case study of the 2017 Ecuadorian Presidential election was conducted, in which the aim was to predict the impact of tweets in terms of number of retweets from the two most voted candidates, based on different features embedded in their own tweets. The features of tweets and their effect on voter reactions have been investigated in elections, and the findings of the investigations suggest that voters can be influenced by the way a tweet is written and that Twitter strategies draw huge public and media attention in the political arena (Lee & Xu, 2018). Therefore, the personalisation of tweets seems to be strategic when engaging with the target audience, since it can affect users' behavioural reactions.

4.2. Data sampling

All tweets produced by the two most voted candidates, Lenin Moreno from the ruling party, and Guillermo Lasso from the opposition, were extracted and collected during the two rounds of the electoral campaign – 16 weeks – by using the Twitter usernames of *@Lenin* and *@Lasso-Guillermo* respectively. Data extraction was performed by means of the Twitter API (application programming interface) and R Core Team (2013). Before starting the analysis of the data, tweets were cleaned by replacing special Spanish accents. Four datasets were generated comprising the tweets each candidate generated during the first and second round of elections. These files also contained information about number of retweets, which is the focus of this study, and other metrics such as date of creation and number of favourites. This comprised in total 650 tweets for Lenin Moreno and 1188 tweets for Guillermo Lasso, as shown in Table 1. This table also includes the number of retweets that candidates' tweets produced, descriptive statistics, and the number of followers and followees at the end of each round of voting. Although Lasso sent almost twice as many tweets as Moreno, the latter candidate achieved more retweets. This might indicate that, in terms of drawing attention from users, Moreno's Twitter campaign was more successful, especially given that the number of followers Lasso had was always higher than Moreno.

In addition, 1.3 million tweets generated by Twitter users about the two candidates were collected to build the hashtag predictor in the model as detailed in Section 4.3.2. The criteria for this data collection were to retrieve tweets including mentions (*@lenin*, *@LassoGuillermo*), hashtags (*#leninmoreno*, *#guillermolasso*), and keywords including candidates' names.

4.3. Data analysis

With the assistance of R, tweets were randomly split for training and testing purposes. For each candidate, datasets from the first and second rounds were each split into five groups, 80% for training and 20% for testing purposes, accounting for a total of ten groups for each candidate. Thus, eight groups out of ten were used as training set (520 and 952 tweets for Lenin Moreno and Guillermo Lasso, respectively). And the remaining two groups for each candidate were used for testing purposes (130 and 236 tweets for Lenin Moreno and Guillermo Lasso, correspondingly).

4.3.1. Description of the model: Output

The number of retweets achieved by each of the candidates' tweets is the metric used to measure the impact of tweets. The output of the model

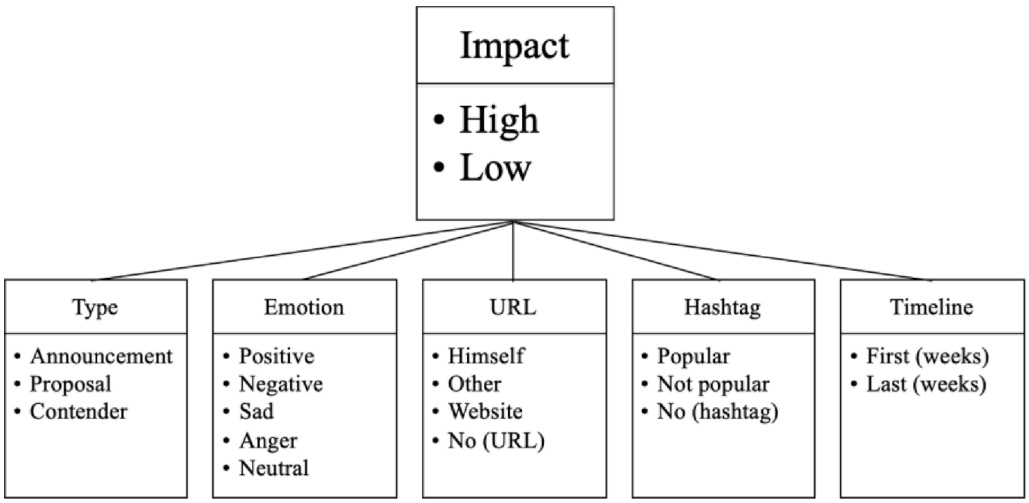

Fig. 1. Original model developed to predict the impact of a tweet based on the number of.

can take two categorical values, "high" or "low", which represents the impact of a tweet. For each candidate, a tweet is classified as high impact if the number of retweets it achieved is above the median of all the candidate's retweets achieved throughout the campaign. Otherwise, the tweet is labelled as low impact. Definition of this classification is consistent with previous works measuring the impact of Twitter in the academic field (Weiss and Davis, 2019) and retweeting behaviour (Rudat & Buder, 2015).

By using the median as threshold, class balance is maintained, meaning that the outputs, high and low, are proportionally distributed across the datasets. In addition, the output of the model was purposely structured as a binary classification precisely for the integration of the MAKER-RIMER model. Also, the analysis used standard retweets. This means retweets of the original tweet as it is. Quote retweets, which are retweets in the form of URLs including a personal comment, were not included in the analysis because the Twitter Search API classifies them as independent tweets instead of retweets. This means that quote retweets do not increase the retweet counts of the shared tweets (Lu, 2019). The use of standard retweets for predicting retweeting behaviour is consistent with previous studies such as Kim et al. (2016), Keib et al. (2018), and Guerrero-Solé and Lopez-Gonzalez (2019).

4.3.2. Description of the model: Inputs

The ideas contained in tweets may influence the propensity of retweeting, as previously demonstrated by Walker (2016). Therefore, content of tweets can help attract new followers and engage them over time (Lalicic, Huertas, Moreno, & Jabreel, 2019). For this purpose, the information about the features of the tweets embedded in the candidates' tweets is extracted and classified into five variables as shown in Fig. 1, which constitute the input of the model, as described below:

The five input variables presented in this study are not exhaustive. There are surely other factors that may also affect the retweet count, which are not considered in this study. For example, Majmundar et al. (2018) proposed a set of extrinsic and intrinsic factors of Twitter users that can drive retweeting behaviour. However, although these factors could provide further insights, for the purpose of developing the MAKER-RIMER model, we focused only on features that are directly observable in the content of tweets.

Type of tweet: This variable represents the type of messages that the candidates posted during the campaign in terms of its purpose. Based on observation of candidates' tweets – not only Moreno and Lasso, but also the tweets of candidates in the last US Presidential election, Donald Trump and Hillary Clinton – a pattern can be seen in which in most cases, these tweets carried any of the following three purposes: 1) promote campaign's proposals; 2) diminish the sentiment towards a contender; and 3) announce details of their daily agendas. While further research is needed to validate the relevance of this approach to categorising tweets, it fitted well for the purpose of this study. To proceed with this classification, candidates' personal tweets are manually analysed and categorised into three groups: "contender", "proposal", or "announcement". A tweet is classified as "contender" if it contains any type of information about the other candidate through hashtags, mentions, names or any other reference, for example the following tweet written by Guillermo Lasso: *"It is comprehensible that @Lenin does not know how to create jobs because he has never done so. I have experience in the private sector"*. If a tweet has information about their own campaign proposals, they are classified as "proposal", for example *"I will derogate the Communication Law"*. Finally, if a tweet does not contain information about their agendas or topics that are considered either contender or proposal, it is labelled as "announcement", and this type of tweet could include tweets such as "*Good morning! We are starting the interview with @desayunos24 in @teleamazonasec."*.

Emotion: The next step is to categorise tweets based on emotion, which is known as emotion analysis. Emotion analysis aims to detect moods or traits based on a specific text, such as trust, sadness, fear, anger, disgust, and disgust (Li, Wu, Zhu, & Xu , 2018). It provides decision makers with a deeper analysis of what emotion could have more impact on their target audience on social media (Chatzakou, Vakali, & Kafetsios, 2017). Emotion differs from sentiment analysis in that emotion relates to people's mood and is determined by a multi-class classifier that includes, for instance, anger, fear, and surprise. Sentiment, on the other hand, is associated with users' feelings and opinions, and is usually measured using a binary classification of positive and negative (Morente-Molinera et al., 2019). For this classification, a text analytic software tool, the Linguistic Inquiry and Word Count software (LIWC2015) (Pennebaker, Boyd, Jordan, & Blackburn, 2015) is used to analyse the emotions of tweets. Five emotions have been detected in tweets: "positive", including words such as love, happy, and nice; "negative", containing words such as hurt, ugly, and nasty; "sadness", embracing words such as crying, grief, and sad; "anger", including words such as hate, annoyed, and pissed. Finally, if LIWC2015 cannot detect any emotion, it is labelled as "neutral".

URL: Due to Twitter's character limitation, the use of URL as a link to an external website can offer deeper content for other Twitter users. When the data are extracted from Twitter using the API search, information such as images, videos, or GIFs are also converted into internal Twitter URLs. In this study, URLs are manually classified as follows: "himself", "other", "website", or "no". "Himself" is assigned if the internal URL contains information about the candidate under analysis, such as images or videos; "other", if the internal URL contains information about other people or situations not directly related to candidate under analysis; "website" if the URL is redirected to an external website; or "no", meaning that a tweet does not contain URLs.

Hashtag: On Twitter, hashtag refers to a word or a phrase preceded by the hash sign "#", which is used as a keyword to identify specific topics. This study assumes that candidates use hashtags to gain awareness from others and to build a political identity (Masroor, Khan, Aib, & Ali , 2019). Hence, the popularity of hashtags is relevant for that purpose. For this study, hashtags are treated as follows. If a candidate's tweets have the presence of hashtags, information about the number of observations in the data generated from other Twitter users about the candidates are weekly classified into two possible values: "popular" if the hashtag the candidate mentions in his tweet appears within the twenty most popular hashtags in the general dataset, which covers almost 80% of the tweets, and otherwise "not popular". If a tweet contains more than one hashtag, the most popular hashtag, in terms of number of occurrences, is considered. If a tweet does not have the presence of hashtags, "no" is assigned to this variable.

Timeline: Previous work has demonstrated that the date tweets are written can affect the retweeting behaviour (Lee & Xu, 2018). Due to the level of polarisation in political campaigns, it is expected that polarisation during the last period of the campaign increases as well as the attention to candidates. For this study, this variable refers to the week that the tweet is written. Since the campaign lasted for 16 weeks, "first" is assigned if tweets are written during the first eight weeks of the campaign; otherwise, they are labelled as "last".

5. Application of the ER rule to predict impact of tweets based on the number of retweets for the 2017 Ecuadorian Presidential election

This section implements the MAKER-RIMER prediction model. The following subsections deal with the implementation of MAKER and RIMER, the method used to train the different parameters, and the comparison of the model against other machine learning approaches.

5.1. Implementing the MAKER framework

If the datasets contained all value combination of input variables, meaning a situation in which frequencies exist for all possible combination of the parameters introduced in Fig. 1, a single model with all five input variables could be used to develop a predictive model for each one

Table 2
Estimates and normalised likelihoods for the first partial MAKER model of Guillermo Lasso comprising the variables type and emotion.

Input variables		Frequencies Output Impact		Estimates likelihood $c_{\theta,il,jm}$ Output Impact		Normalised likelihood $p_{\theta,il,jm}$ Output Impact	
		High	Low	High	Low	High	Low
Positive	Announcement	142	109	0.2971	0.2300	0.5637	0.4363
	Proposal	241	240	0.5042	0.5063	0.4989	0.5011
	Contender	15	1	0.0314	0.0021	0.9370	0.0630
Negative	Announcement	8	5	0.0167	0.0105	0.6134	0.3866
	Proposal	5	0	0.0105	0.0000	1.0000	0.0000
	Contender	13	0	0.0272	0.0000	1.0000	0.0000
Sadness	Announcement	1	7	0.0021	0.0148	0.1241	0.8759
	Proposal	0	2	0.0000	0.0042	0.0000	1.0000
	Contender	0	0	0.0000	0.0000	0.0000	0.0000
Anger	Announcement	9	4	0.0188	0.0084	0.6905	0.3095
	Proposal	4	1	0.0084	0.0021	0.7987	0.2013
	Contender	7	0	0.0146	0.0000	1.0000	0.0000
Neutral	Announcement	25	91	0.0523	0.1920	0.2141	0.7859
	Proposal	5	13	0.0105	0.0274	0.2761	0.7239
	Contender	3	1	0.0063	0.0021	0.7484	0.2516

[1]For this combination of parameters, data were not available for calculating joint probabilities and prediction

Table 3
Interdependence index applied for the first partial MAKER model of Guillermo Lasso.

Input variables		Interdependence index			
		Normalised likelihood Output		Ordinary likelihood Output	
		High	Low	High	Low
Positive	Announcement	2.3158	1.7168	5.8175	4.3128
	Proposal	1.8945	2.1191	3.1620	3.5369
	Contender	1.8618	2.6593	7.3116	10.4433
Negative	Announcement	1.5946	4.4016	3.1985	8.8288
	Proposal	2.4027	0.0000	16.0140	0.0000
	Contender	1.2573	0.0000	0.2513	0.0000
Sadness	Announcement	2.7223	1.7983	2.8683	1.8948
	Proposal	0.0000	2.2068	0.0000	11.8356
	Contender	0.0000	0.0000	0.0000	0.0000
Anger	Announcement	1.8826	2.8425	3.0482	4.6025
	Proposal	2.0124	1.9878	10.8021	10.6699
	Contender	1.3186	0.0000	0.3949	0.0000
Neutral	Announcement	1.9620	1.9063	1.9666	1.9108
	Proposal	2.3384	1.8874	19.2526	15.5395
	Contender	3.3170	6.5472	9.6216	18.9914

Table 4
First partial MAKER model results after training weights of variables (type and emotion) and their parameters for Guillermo Lasso.

Input variables		MAKER 1 results after training weights Output	
		High	Low
Positive	Announcement	0.5877	0.4123
	Proposal	0.5026	0.4974
	Contender	0.9741	0.0259
Negative	Announcement	0.6280	0.3720
	Proposal	0.9484	0.0516
	Contender	0.9035	0.0965
Sadness	Announcement	0.2224	0.7776
	Proposal	0.0672	0.9328
	Contender	0.0000	0.0000
Anger	Announcement	0.6967	0.3033
	Proposal	0.8656	0.1344
	Contender	0.9253	0.0747
Neutral	Announcement	0.2601	0.7399
	Proposal	0.1807	0.8193
	Contender	0.8013	0.1987

of the two candidates. In this study, however, developing such a single model could lead to misleading results because of the absence of sufficient data for taking into account the five variables together (McDonald, 2014). Since the datasets for this study do not contain all possible combinations of input variables, two partial MAKER models for each candidate, needed to be trained to generate the prediction of the impact of tweets. The combination of variables in each partial MAKER model needs to have sufficient data to calculate the joint probabilities. Each partial MAKER model comprises two and three variables which need to be closely correlated to each other to calculate joint probabilities.

To select the combination of variables to be grouped into each partial MAKER model, the rationale is to choose the combination of variables with the highest number of records evenly distributed in the space model, which shows the relationship between the input and output variables (Yang & Xu, 2017). This is done by registering the frequencies for all possible combinations of two and three parameters of the input variables. Using two partial MAKER models instead of a single one is justified in that it allows increasing the use of the available data. For this study, some combinations of variables do not have records available if a single model were to be developed.

From each combination of input variables selected, joint probability tables are generated. This is done by recording the frequencies for the inputs and outputs variables from the observations. Then, the interdependent index is calculated, which shows the statistical interrelationship between each group of variables. From these results, the partial MAKER model is trained to predict the impact of the tweets. The weights of the five variables and their parameters are trained for optimal prediction. The parameters of the variables are the sub-categories of each input variable in the partial MAKER models. For instance, the input variable "type" can have three possible parameters namely "announcement", "proposal" or "contender".

An example using the data from the first partial MAKER model from Guillermo Lasso, comprising the input variables type and emotion, is following presented in Tables 2, 3, and 4. The first step in the implementation of MAKER is the creation of the joint frequency tables (i.e. columns 3 and 4 in Table 2) while the result of the output variable can be high or low impact. The estimates likelihood that the pieces of evidence of the input variables are high or low are presented in $c_{\theta,il,jm}$ as shown in columns 5 and 6 from Table 2. Then, the normalised likelihoods are calculated as shown in columns 7 and 8 in Table 2 using Eq. (1) to estimate the joint basic probability. For the partial MAKER models presented in this case study, the pieces of evidence are the observations of the parameters of the input variables and the output. For example, the pieces of evidence of the first partial MAKER model for Guillermo Lasso comprise a parameter for emotion (positive, negative, sadness, anger, neutral), a parameter of type of tweet (announcement, proposal, contender), and an impact for the tweet (high or low). Each piece of

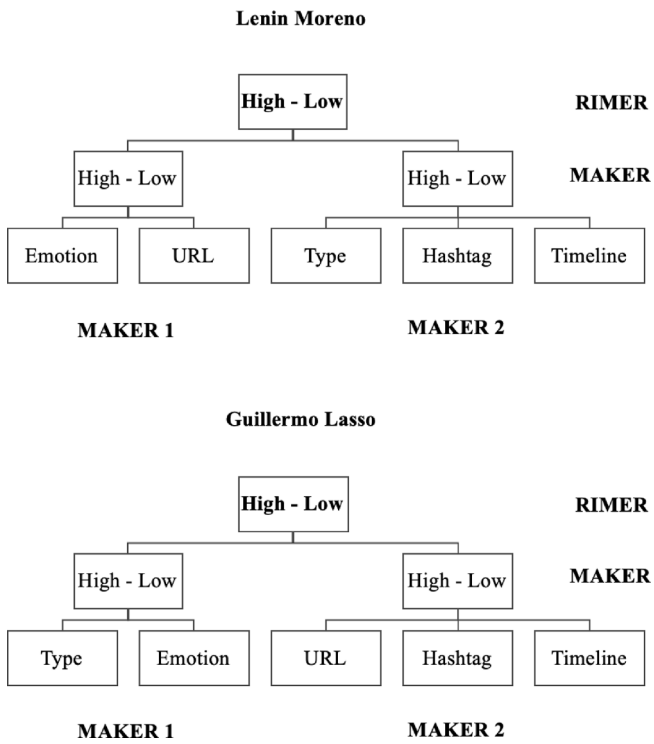

Fig. 2. Hierarchical structures of the models for both candidates.

evidence is represented as an extended probability distribution or belief distribution, with probabilities assigned to propositions, e.g. singleton propositions such as high and low, or non-singleton propositions such as the set of high or low. This way, ambiguity (or unknown) caused by missing data can be represented as probabilities assigned to non-singleton propositions such as the set of high or low.

After obtaining the joint basic probability, the interdependence index is calculated between each pair of evidence and it is presented in columns 3 and 4 of Table 3. In addition, normalised likelihood is scaled to ordinary likelihood as shown in columns 5 and 6 in the same table using Eq. (2).

The model starts assuming that initial weights are the same for all the input variables and their parameters, but these weights are later trained or optimised as will be explained in Section 5.3. In addition, since the data come from the same source, weight is equal to reliability. In the MAKER framework, two pieces of evidence are combined to generate the combined support for proposition θ and it is defined by using Eqs. (5) and (4), and the results are shown in Table 4. This table presents the results of the first partial MAKER model after training the weights of the parameters of MAKER. The table shows, for Guillermo Lasso, the probabilities of a tweet achieving high or low impact for each combination of inputs type and emotion. For example, a type of tweet coded as "announcement" with an emotion coded as "positive", has a probability of 0.5877 of being high impact, and 0.4123 of being low impact. In this sense, MAKER shows a transparent process on how weights are assigned and trained.

This process is repeated with the second group of variables that form the second partial MAKER model, which for Guillermo Lasso comprises the three following input variables: URL, hashtag, and timeline (Table A.2 in the Appendix A, while abbreviations of variables and parameters are shown in Table A.1). A similar calculation process was conducted for the two partial MAKER models for Lenin Moreno, and the results are presented in Table A.3 and Table A.4 in Appendix A.

Table 5
Illustration of the possible belief rules and belief degrees used for this case study.

Output	Four belief rules			
	High/High	High/Low	Low/High	Low/Low
High	Belief degree 1	Belief degree 3	Belief degree 5	Belief degree 7
Low	Belief degree 2	Belief degree 4	Belief degree 6	Belief degree 8

Table 6
Rule base using RIMER with updated belief degrees considering MAKER 1 and MAKER 2 partial model results for Lenin Moreno.

Antecedent	Consequent
(MAKER 1 is high $\wedge$ MAKER 2 is high)	Impact of tweet is $\{(\text{high}, 0.9371), (\text{low}, 0.0629)\}$
(MAKER 1 is high $\wedge$ MAKER 2 is low)	Impact of tweet is$\{(\text{high}, 0.3400), (\text{low}, 0.6600)\}$
(MAKER 1 is low $\wedge$ MAKER 2 is high)	Impact of tweet is$\{(\text{high}, 0.3313), (\text{low}, 0.6687)\}$
(MAKER 1 is low $\wedge$ MAKER 2 is low)	Impact of tweet is$\{(\text{high}, 0.2508), (\text{low}, 0.7492)\}$

Table 7
Rule base using RIMER with updated belief degrees considering MAKER 1 and MAKER 2 partial model results for Guillermo Lasso.

Antecedent	Consequent
(MAKER 1 is high $\wedge$ MAKER 2 is high)	Impact of tweet is$\{(\text{high}, 0.9651), (\text{low}, 0.0349)\}$
(MAKER 1 is high $\wedge$ MAKER 2 is low)	Impact of tweet is$\{(\text{high}, 0.6745), (\text{low}, 0.3255)\}$
(MAKER 1 is low $\wedge$ MAKER 2 is high)	Impact of tweet is$\{(\text{high}, 0.3203), (\text{low}, 0.6797)\}$
(MAKER 1 is low $\wedge$ MAKER 2 is low)	Impact of tweet is$\{(\text{high}, 0.1037), (\text{low}, 0.8963)\}$

5.2. *Implementing the RIMER framework*

After completing the partial MAKER models for each candidate, a RIMER model is developed to combine the results generated by the two partial MAKER models, as shown in Fig. 2, where MAKER 1 is the first partial MAKER model comprising two input variables, and MAKER 2 is the second partial MAKER model comprising three input variables for each candidate. Thus, RIMER combines the two partial MAKER models back together.

In the RIMER model, an initial belief rule base (BRB) is constructed, consisting of belief rules established on the basis of the types of outputs of the two partial MAKER models. For this reason, the parameters, which for RIMER are composed of the attribute weights of the two MAKER models and four belief rules, and the eight belief degrees of the four belief rules, need to be trained. The four belief rules come from the possible combination of the output, which can be High/High, High/Low, Low/High, and Low/Low. And the eight belief degrees are the results of all the possible consequents of a rule as shown in Table 5.

From the outputs of the two partial MAKER models, RIMER can be implemented as follows. The activation weight for each belief rule is calculated using Eq. (7). Then, the degrees of belief β_{ik} are generated implementing Eq. (9). Following the training of the RIMER parameters, as will be explained in Section 5.3, the final RIMER results are presented in Table 6 for Lenin Moreno, and Table 7 for Guillermo Lasso. For example, from Table 7 about Guillermo Lasso, when MAKER 1 (i.e. type: announcement, and emotion: positive) is high, and MAKER 2 (i.e. URL: website, hashtag: not popular, and timeline: first) is low, the probability of a tweet being high impact is 0.6745, and of being low impact is 0.3255.

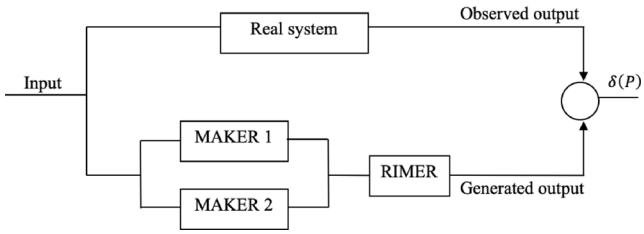

Fig. 3. Illustration of the MAKER-RIMER generic training process.

Fig. 3 presents a generic model of the optimal learning process adapted from Yang et al. (2007). This model is used to represent the process to predict outcomes from input variables, which includes the implementation of the two partial MAKER models and RIMER methodology for each candidate.

5.3. Training the parameters of MAKER and RIMER

To begin with, weights are randomly assigned for both, parameters of MAKER and those of the RIMER. The parameters for MAKER, for this study, are composed of the five input variables and their subcategories as presented in Fig. 1. The parameters of the RIMER are the attributes, belief rules, and belief degrees. Then, weights from the training dataset need to be trained to improve the performance of the model (Xu, Zheng, Yang, Xu, & Chen, 2017). An optimisation model to minimise the misclassification error (MCE) is applied using the Eq. (10), where P represents the vector of the parameters for MAKER and RIMER to be trained. The model above is optimised by means of Differential Evolution (Ardia, Mullen, Peterson, & Ulrich, 2016) as implemented in the R package "DEoptimR" (Conceicao & Maechler, 2016). The stopping criterion was based on the maximum number of iterations to be performed before the optimisation process is stopped, which was set to 5,000 iterations.

$$\delta(P) = \frac{1}{S}\sum_{s=1}^{S}\sum_{\theta\subseteq\Theta}\left(p^{(s)}(\theta) - \widehat{p}^{(s)}(\theta)\right)^2$$

$$s.t. 0 \le P \le 1 \tag{10}$$

where $\delta(P)$ refers to the objective function aiming to reduce the MCE,S is the total number of observations, $p^{(s)}(\theta)$ is the expected score of the output generated by using the MAKER and RIMER models for the sth observation, and $\widehat{p}^{(s)}(\theta)$ is the observed output of the sth observation. The constraints of the training model for both MAKER and RIMER encompass normalisation of weights, so that they are between zero and one (J.-B. Yang et al., 2007).

5.4. Comparison with other machine learning approaches

Four machine learning approaches were also applied to compare the results obtained using MAKER-RIMER namely logistic regression (LR), Naïve Bayes (NB), decision tree (DT), and support vector machine (SVM). The four machine learning approaches described above were also implemented in R, with the same training and testing datasets used to develop the MAKER-RIMER model. For this purpose, the following R packages were used: "nnet" (Venables & Ripley, 2002) for LR, "e1071" (Meyer, Dimitriadou, Hornik, Weingessel, & Leisch, 2017) for both NB and SVM, and "tree" (Ripley, 2012) for DT. One difference when implementing these approaches compared to MAKER-RIMER is that for the former approaches the input variables were considered all at once in one single model and not hierarchically, as suggested in Fig. 2. In this sense, the MAKER-RIMER model maximises the use of data available by splitting the input variables into two groups for each candidate because the case study does not have sufficient data for combining the five variables together.

To test the performance of the different classifiers, MCE was the metric used for comparison purposes. MCE, also known as error rate, refers to the total proportion of observations that are incorrectly classified across all the classes. This measure is used for evaluation since it works appropriately in predictive models with balanced outcome classes as suggested by Ballabio, Grisoni & Todeschini (2018). To calculate this metric, it is necessary to obtain the false positive (FP) and false negative (FN) results, which refer to the proportion of instances incorrectly classified for the considered classes as positive and negative respectively, while true positive (TP) and true negative (TN) refer to the proportion of instances correctly classified for the considered classes as positive and negative correspondingly, as shown in Eq. (11).

$$MCE = \frac{FP + FN}{TP + FN + FP + TN} \tag{11}$$

6. Results and discussion

The results from MAKER-RIMER after training the parameters are presented in Table 6 for Lenin Moreno, and Table 7 for Guillermo Lasso. The belief structures presented in both tables can provide decision makers with a tool to predict the outcome of the impact of tweets based on their own features. So, campaigners could estimate the possible outcomes and their probability of occurrence and understand what is indeed triggering such outcomes. From an intuitive line of thinking, it can be assumed that when both partial MAKER models are high impact, the result for RIMER should be also high. Similarly, when both MAKER models are low impact, the intuitive RIMER result is expected to be low. However, when the MAKER models are high/low or low/high, results from RIMER are uncertain. Hence, intuitive thinking might not be precise, so a robust model needs to be trained based on the actual data and weights. In this sense, the proposed MAKER-RIMER model overcomes this limitation by providing a robust model grounded on the evidence of the data. For example, using Guillermo Lasso's results, high impact in tweets is achieved either when both MAKER models are high (0.9651), or when the partial MAKER 1 model is high (0.6745). Similarly, low impact is obtained either when both MAKER results are low impact (0.8963), or when the partial MAKER 1 model is low (0.6797). These results show that the process is interpretable because is openly readable for decision-makers in the sense that they can fully understand how the input variables are affecting the outcome.

The MAKER-RIMER model proposed in this study provides several advantages. In terms of data availability and statistical significance, this case study does not offer sufficient data to combine the five input variables together; thus, not all the combinations of variables are available for prediction purposes. These limitations are overcome when grouping input variables to form the two partial MAKER models. This is achieved by aggregating variables that are more closely interrelated, which form the lowest part of the structure as presented in Fig. 2, for which is said the model is transparent. The results from both partial MAKER models are considered when applying RIMER methodology, as shown in the top of the structure in the Fig. 2. This shows that, even when there are not statistically significant data from all the variables together, the prediction is still evidence-based and the reasoning is based on the knowledge of the data, and not on intuition. On the other hand, other data-driven modelling approaches might attempt to intuitively perform the prediction with all the variables together, even in the absence of statistically meaningful data to train the whole model. As a result, the model might not be fully interpretable or be trusted because of the limitations of the data, or because the inference process is like a black box, in which these cannot provide enough detailed to understand how the inference process is carried out be not fully explained (Rudin, 2019).

In terms of interpretability, the model presented in this research provides a robust procedure to map and represent the inputs and outputs (Kong et al., 2016; Yang et al., 2007). This is helpful when interpreting

Table 8
Comparison of performance of machine learning methods based on the MCE.

Approaches	Lenin Moreno		Guillermo Lasso	
	MCE Train	MCE Test	MCE Train	MCE Test
MAKER-RIMER	0.4115	0.3385	0.2489	0.2373
LR	0.4250	0.3538	0.2574	0.2585
NB	0.4385	0.3923	0.2489	0.2500
DT	0.4635	0.4000	0.2532	0.2415
SVM	0.4269	0.4308	0.2595	0.2585

the results from the model relating to how to predict the impact of tweets based on the characteristics embedded in their own tweets. Unlike other machine learning approaches that do not provide a clear revelation of how the inference process unfolds (Rudin, 2019), the MAKER-RIMER model presents a more interpretable and transparent process. This not only considers the reasoning behind how variables are grouped together, but also provides details (in the MAKER models) for the weights of the variable inputs and their parameters. Also, it contemplates belief degrees, weights of antecedent attributes, and rules in RIMER. As explained in Section 5.3, the weights of parameters for MAKER-RIMER were first randomly assigned and later trained. In this sense, the initial assumption of equally weighted parameters is challenged because after the optimisation, the weights of the parameters for the model are trained to predict the impact output. In sum, the MAKER-RIMER model overcomes intuitive thinking, and provides campaigners with a knowledgeable tool, which is a transparent and interpretable reasoning process that determines the outputs based on the available inputs from the data. Limitations of the model might arise especially when relationships between predictors and outcomes are not available, or prior knowledge is limited, so constructing the initial knowledge base represents a challenge (Kong et al., 2016; Yang et al., 2007). In addition, if the datasets are affected by noise, the generation of rules and the overall outcomes of MAKER-RIMER represents a challenge (Yang et al., 2006).

In addition, the comparison of the performance of the different classification methods is shown in Table 8. These results suggest that the approach with the best performance in terms of minimum MCE is MAKER-RIMER for both candidates. MCE values for Guillermo Lasso were consistently smaller than those for Lenin Moreno because Lasso posted almost twice as many tweets as Moreno.

After obtaining results from the performance of the models, the next step is to identify which characteristics of tweets affect the impact of tweets for each candidate. From the MAKER-RIMER results, the two candidates share similar patterns when achieving the high impact of tweets. For example, the prominent high impact tweets should include information about the contender, either with a positive or with a negative connotation. Tweets written during the last period have higher impact. The presence of hashtags for both candidates is not associated with high impact. The difference between the candidates lies in URLs, since for Lenin Moreno high impact is linked with URLs about himself, while for Guillermo Lasso it is linked with URLs about other people or situations not directly related to his image. Concerning those tweets with low impact on retweets, it is observed that for Lenin Moreno, only one combination of the characteristics of tweets generated low impact that comprises tweets containing announcements with sad connotations, having URLs about himself, without hashtags, and written in the last period of the campaign. However, for Guillermo Lasso, 25 possible combinations of the parameters of variables resulted in low impact. The most prevalent combinations included announcements with a neutral emotion, URLs containing information about himself or without URLs, with positive hashtags, and written in the first weeks of the campaign.

The results of this study are supported by previous research conducted in the political field. Concerning the diffusion of tweets based on type, those containing attacks on the contender tend to attract more attention from users (Darwish, Magdy, & Zanouda, 2017; Lee & Xu, 2018). In addition, emotional content is more viral than non-emotional content (Kim et al., 2016). In this sense, a high diffusion of information is achieved when the content of tweets involves positive emotion (Wang, Zhou, Jin, Fang, & Lee, 2017), but even higher if it conveys negative connotations (Lee & Xu, 2018). This study showed that high impact is associated with tweets involving either positive or negative emotions. Concerning URLs and hashtags, most studies have focused on the number thereof in tweets and their positive impact when retweeting, but surprisingly, for this study at least, the presence of hashtags is not prominently seen as vital when retweeting. Concerning the timeline, since the level of polarization and Twitter traffic in the last period of the campaign increases, it is expected that tweets will gain more attention and diffusion (Cram, Hill, & Magdy, 2017; Darwish et al., 2017).

7. Conclusion

In the task of dynamically shaping the political image of a candidate, information continuously emerging on Twitter may provide sources for effective and opportune actions, which would usually involve adjustments in rhetoric, symbols, and language. The information produced herein should be useful for dynamic decision making in managing a campaign and constructing or shaping the images of candidates.

This study has presented an ER based predictive model, MAKER-RIMER, to predict the impact of tweets in terms of the number of the retweets they achieved. The tweets posted by the two most voted candidates of the 2017 Ecuadorian Presidential election were used to develop the model. This model is based on likelihood data analysis and probabilistic inference via evidence combination. The proposed model provides a better interpretability of the reasoning process and results. It also presents and compares the performances of different machine learning approaches for prediction.

In addition, findings show that the MAKER-RIMER model performed better than other machine learning approaches in predicting the impact of tweets, as it showed a smaller MCE. A smaller MCE is relevant since errors continue to be a barrier for machine learning approaches to be comprehensively adopted in prediction of human behaviour. The model presented in this study also allows the identification of features of a tweet that are predictors of its impact for each candidate. The results have shown that for both candidates, high impact is obtained when their tweets include information about the contender, have either a positive or negative emotion, with URLs comprising information about the candidates themselves or about other people or situations not directly related to them, without the presence of hashtags, and written in the last period of the campaign.

The generalisability of the results of this study is subject to certain limitations. This study only used tweets generated by the two most voted candidates to build the predictive model, with other users' tweets disregarded. The impact of the limited choice of data on the quality of the generated predictive model remains uninvestigated. In addition, the model is appropriate depending upon Twitter penetration among users/voters and upon candidates' participation on Twitter. So, the model is appropriate when both parts generate and consume information on Twitter. In addition, the list of predictors in the model used in this study is not exhaustive. They depend on the fields and contexts of case studies, meaning that new variables can be added or adapted for consideration, which could be critical to outcomes, but can serve as a starting point for developing a retweeting model. For example, the model does not include the presence of emoticons or lists as predictors because political campaigns on Twitter usually lack them.

In terms of future research, the proposed MAKER-RIMER model could be tested in different fields to analyse how the predictors in the model work, and how new variables can be adapted in different contexts. In addition, to enhance Twitter campaigns, future work could build retweeting predictive models which include tweets by other relevant users that generate high impact in terms of retweets. In this sense, candidates would be able to learn what works well for these users so that they can adapt their own tweets accordingly. Another need for future

research is the automation of variable classification, which was labelled manually in this study, by using machine learning, either via unsupervised (clustering) or supervised (classification) approaches. Finally, the model could be tested using unbalanced datasets for classifying high and low impact, for example assigning 30% of the data to high impact, while the remaining is assigned to low impact, to analyse the overall performance of the model with imbalanced classes.

CRediT authorship contribution statement

Lucía Rivadeneira: Investigation, Conceptualization, Methodology, Software. **Jian-Bo Yang:** Investigation, Supervision. **Manuel López-Ibáñez:** Investigation, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The Secretariat of Higher Education, Science, Technology, and Innovation (SENESCYT) for supporting L. Rivadeneira with a PhD scholarship. SENESCYT did not have any involvement in the different phases of this study.

Appendix A

Table A.1
Abbreviations of variables and parameters.

Abbreviations			Description
Type	A	Announcement	If the tweet contains information about facts or occurrences
	P	Proposal	If the tweet contains information about proposals
	C	Contender	If the tweet contains information about the other candidate
Emotion	POS	Positive	If the emotion contained in the tweet is positive
	NEG	Negative	If the emotion contained in the tweet is negative
	SAD	Sadness	If the emotion contained in the tweet is sad
	ANG	Anger	If the emotion contained in the tweet is angry
	NEU	Neutral	If no emotion can be detected in the tweet
URL	H	Himself	If the URL contain information about the candidate himself
	O	Other	If the URL contain information about other people/situations
	W	Website	If the URL redirect to an external website
	N	No URL	If no URL can be found in the tweet
Hashtag	P	Popular	If the hashtag is among the 20 most popular hashtags
	NP	Not popular	If the hashtag is not among the 20 most popular hashtags
	N	No hashtag	If no hashtag can be found in the tweet
Timeline	F	First	If the tweet was written between weeks 1 and 8
	L	Last	If the tweet was written between weeks 9 and 16

Table A.2
MAKER 2: Second partial model involving URL, hashtag, and timeline for Guillermo Lasso.

Variables	MAKER 2 results after training weights	
	High	Low
H-P-F	0.1295	0.8705
O-P-F	0.0900	0.9100
W-P-F	0.0107	0.9893
N-P-F	0.0712	0.9288
H-NP-F	0.0187	0.9813
O-NP-F	0.3496	0.6504
W-NP-F	0.1059	0.8941
N-NP-F	0.0485	0.9515
H-N-F	0.0756	0.9244
O-N-F	0.3149	0.6851
W-N-F	0.0153	0.9847
N-N-F	0.1348	0.8652
H-P-L	0.5561	0.4439
O-P-L	0.7942	0.2058
W-P-L	0.8208	0.1792
N-P-L	0.7739	0.2261
H-NP-L	0.4640	0.5360
W-NP-L	0.3495	0.6505
N-NP-L	0.8081	0.1919
H-N-L	0.6749	0.3251
O-N-L	0.8798	0.1202
W-N-L	0.4283	0.5717
N-N-L	0.7124	0.2876

Table A.3
MAKER 1: First partial model involving emotion and URL for Lenin Moreno.

Variables	MAKER 1 results after training weights	
	High	Low
POS-H	0.4963	0.5037
NEG-H	0.7400	0.2600
SAD-H	0.0376	0.9624
NEU-H	0.4081	0.5919
POS-O	0.5254	0.4746
NEG-O	0.9604	0.0396
SAD-O	0.5032	0.4968
NEU-O	0.5013	0.4987
POS-W	0.2112	0.7888
POS-N	0.5100	0.4900
NEG-N	0.6849	0.3151
SAD-N	0.6450	0.3550
NEU-N	0.6717	0.3283

Table A.4
MAKER 2: Second partial model involving type, hashtag, and timeline for Lenin Moreno.

Variables	MAKER 2 after training weights	
	High	Low
A-P-F	0.4061	0.5939
P-P-F	0.6349	0.3651
P-NP-F	0.9711	0.0289
A-N-F	0.4616	0.5384
P-N-F	0.6608	0.3392
A-P-L	0.6456	0.3544
P-P-L	0.7625	0.2375
A-NP-L	0.6803	0.3197
P-NP-L	0.4475	0.5525
A-N-L	0.4970	0.5030
P-N-L	0.4073	0.5927
C-N-L	1.0000	0.0000

References

Abdullah, N. A., Nishioka, D., Tanaka, Y., & Murayama, Y. (2017). Why I retweet? Exploring user's perspective on decision-making of information spreading during disasters. Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS-50). https://doi.org/10.24251/hicss.2017.053.

Abraham, A., Han, S. Y., Al-Sharhan, S., & Liu, H. (Eds.). (2016). *Hybrid Intelligent Systems* (1st ed.). Cham: Springer.

Almaghrabi, F., Xu, D.-L., & Yang, J.-B. (2019). A new machine learning technique for predicting traumatic injuries outcomes based on the vital signs. In *Proceedings of the 25th International Conference on Automation and Computing (ICAC)* (pp. 1–5).

Ananda, A. S., Hernández-García, Á., Acquila-Natale, E., & Lamberti, L. (2019). What makes fashion consumers "click"? Generation of eWoM engagement in social media. *Asia Pacific Journal of Marketing and Logistics, 31*(2), 398–418. https://doi.org/10.1108/APJML-03-2018-0115

Ardia, D., Mullen, K. M., Peterson, B. G., & Ulrich, J. (2016). DEoptim: Differential Evolution in R. Version 2.2-4. https://CRAN.R-project.org/package=DEoptim.

Ballabio, D., Grisoni, F., & Todeschini, R. (2018). Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems, 174*, 33–44. https://doi.org/10.1016/j.chemolab.2017.12.004

Benalla, M., Achchab, B., & Hrimech, H. (2020). Improving driver assistance in intelligent transportation systems: An agent-based evidential reasoning approach. *Journal of Advanced Transportation, 2020*, 1–14. https://doi.org/10.1155/2020/4607858

Bigsby, K. G., Ohlmann, J. W., & Zhao, K. (2019). The turf is always greener: Predicting decommitments in college football recruiting using Twitter data. *Decision Support Systems, 116*, 1–12. https://doi.org/10.1016/j.dss.2018.10.003

Can, E. F., Oktay, H., & Manmatha, R. (2013). Predicting retweet count using visual cues. Proceedings of the 22nd ACM International Conference on Information & Knowledge Management (pp. 1481–1484). https://doi.org/10.1145/2505515.2507824.

Chatzakou, D., Vakali, A., & Kafetsios, K. (2017). Detecting variation of emotions in online activities. *Expert Systems with Applications, 89*, 318–332. https://doi.org/10.1016/j.eswa.2017.07.044

Chen, Y.-W., Yang, J.-Bo., Xu, D.-L., Zhou, Z.-J., Tang, D.-W., et al. (2011). Inference analysis and adaptive training for belief rule-based systems. *Expert Systems with Applications, 38*, 12845–12860. https://doi.org/10.1016/j.eswa.2011.04.077

Clarke, I., & Grieve, J. (2019). Stylistic variation on the Donald Trump Twitter account: A linguistic analysis of tweets posted between 2009 and 2018. *PLoS One, 14*(9). https://doi.org/10.1371/journal.pone.0222062

Conceicao, E., & Maechler, M. (2016). Differential Evolution Optimization in Pure R. R package version 1.0-8. https://CRAN.R-project.org/package=DEoptimR.

Cram, L., Llewellyn, C., Hill, R., & Magdy, W. (2017). UK General election 2017: A Twitter analysis. *ArXiv Preprint*. ArXiv:1706.02271.

Darwish, K., Magdy, W., & Zanouda, T. (2017). *Trump vs. Hillary: What went viral during the 2016 US presidential election.* (pp. 143–161) Cham: Springer.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics, 38*(2), 325–339. https://doi.org/10.1214/aoms/1177698950

Enli, G. (2017). Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. *European Journal of Communication, 32*(1), 50–61. https://doi.org/10.1177/0267323116682802

Fan, C., Jiang, Y., Yang, Y., Zhang, C., & Mostafavi, A. (2020). Crowd or Hubs: Information diffusion patterns in online social networks in disasters. *International Journal of Disaster Risk Reduction, 46*, 101498. https://doi.org/10.1016/j.ijdrr.2020.101498

Fu, C., Xue, M., Chang, W., Xu, D., & Yang, S. (2020). An evidential reasoning approach based on risk attitude and criterion reliability. *Knowledge-Based Systems, 199*, 105947. https://doi.org/10.1016/j.knosys.2020.105947

Gallo, F. R., Simari, G. I., Martinez, M. V., & Falappa, M. A. (2020). Predicting user reactions to Twitter feed content based on personality type and social cues. *Future Generation Computer Systems, 110*, 918–930. https://doi.org/10.1016/j.future.2019.10.044

Gao, X., Zheng, Z., Chu, Q., Tang, S., Chen, G., & Deng, Q. (2019). Popularity prediction for single tweet based on heterogeneous Bass model. In *IEEE Transactions on Knowledge and Data Engineering* (pp. 1–14). https://doi.org/10.1109/TKDE.2019.2952856

Grčar, M., Cherepnalkoski, D., Mozetič, I., & Kralj Novak, P. (2017). Stance and influence of Twitter users regarding the Brexit referendum. *Computational Social Networks, 4*(1), 1–25. https://doi.org/10.1186/s40649-017-0042-6

Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2019). Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences. *Technological Forecasting and Social Change, 145*, 438–460. https://doi.org/10.1016/j.techfore.2018.09.009

Guerrero-Solé, F., & Lopez-Gonzalez, H. (2019). Government formation and political discussions in Twitter: An extended model for quantifying political distances in multiparty democracies. *Social Science Computer Review, 37*(1), 3–21. https://doi.org/10.1177/0894439317744163

Houston, J. B., McKinney, M. S., Thorson, E., Hawthorne, J., Wolfgang, J. D., & Swasy, A. (2020). The twitterization of journalism: User perceptions of news tweets. *Journalism, 21*(5), 614–632. https://doi.org/10.1177/1464884918764454

Iyer, R. R., Zheng, R., Li, Y., & Sycara, K. (2019). Event outcome prediction using sentiment analysis and crowd wisdom in microblog feeds. *ArXiv Preprint*. ArXiv:1912.05066.

Jalali, N. Y., & Papatla, P. (2019). Composing tweets to increase retweets. *International Journal of Research in Marketing, 36*(4), 647–668. https://doi.org/10.1016/j.ijresmar.2019.05.001

Keib, K., Himelboim, I., & Han, J.-Y. (2018). Important tweets matter: Predicting retweets in the #BlackLivesMatter talk on twitter. *Computers in Human Behavior, 85*, 106–115. https://doi.org/10.1016/j.chb.2018.03.025

Kim, E., Hou, J., Han, J. Y., & Himelboim, I. (2016). Predicting retweeting behavior on breast cancer social networks: Network and content characteristics. *Journal of Health Communication, 21*(4), 479–486. https://doi.org/10.1080/10810730.2015.1103326

Kong, G., Xu, D.-L., Yang, J.-B., & Ma, X. (2015). Combined medical quality assessment using the evidential reasoning approach. *Expert Systems with Applications, 42*(13), 5522–5530. https://doi.org/10.1016/j.eswa.2015.03.009

Kong, G., Xu, D.-L., Yang, J.-B., Wang, T., & Jiang, B. (2020). Evidential reasoning rule-based decision support system for predicting ICU admission and in-hospital death of trauma. In *IEEE Transactions on Systems, Man, and Cybernetics: Systems* (pp. 1–12). https://doi.org/10.1109/TSMC.2020.2967885

Kong, G., Xu, D.-L., Yang, J.-B., Yin, X., Wang, T., Jiang, B., & Hu, Y. (2016). Belief rule-based inference for predicting trauma outcome. *Knowledge-Based Systems, 95*, 35–44. https://doi.org/10.1016/j.knosys.2015.12.002

Lalicic, L., Huertas, A., Moreno, A., & Jabreel, M. (2019). Which emotional brand values do my followers want to hear about? An investigation of popular European tourist destinations. *Information Technology & Tourism, 21*(1), 63–81. https://doi.org/10.1007/s40558-018-0134-7

Laugel, T., Lesot, M.-J., Marsala, C., Renard, X., & Detyniecki, M. (2018). Comparison-based inverse classification for interpretability in machine learning. 17th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2018), 100–111. https://doi.org/10.1007/978-3-319-91473-2_9.

Lee, J., & Lim, Y.-S. (2016). Gendered campaign tweets: The cases of Hillary Clinton and Donald Trump. *Public Relations Review, 42*(5), 849–855. https://doi.org/10.1016/j.pubrev.2016.07.004

Lee, J., & Xu, W. (2018). The more attacks, the more retweets: Trump's and Clinton's agenda setting on Twitter. *Public Relations Review, 44*(2), 201–213. https://doi.org/10.1016/j.pubrev.2017.10.002

Li, J., Wu, Z., Zhu, B., & Xu, K. (2018). Making sense of organization dynamics using text analysis. *Expert Systems with Applications, 111*, 107–119. https://doi.org/10.1016/j.eswa.2017.11.009

Liu, F., Chen, Y.-W., Yang, J.-B., Xu, D.-L., & Liu, W. (2019). Solving multiple-criteria R&D project selection problems with a data-driven evidential reasoning rule. *International Journal of Project Management, 37*(1), 87–97. https://doi.org/10.1016/j.ijproman.2018.10.006

Lo, S. L., Chiong, R., & Cornforth, D. (2016). Ranking of high-value social audiences on Twitter. *Decision Support Systems, 85*, 34–48. https://doi.org/10.1016/j.dss.2016.02.010

Lu, Y. Y. (2019). #VoteLeave or #StrongerIn: Resonance and rhetoric in the EU referendum [Doctoral dissertation, Oxford Internet Institute]. University of Oxford.

Ma, R., Hu, X., Zhang, Q., Huang, X., & Jiang, Y.-G. (2019). Hot topic-aware retweet prediction with masked self-attentive model. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 525–534). https://doi.org/10.1145/3331184.3331236

Majmundar, A., Allem, J.-P., Boley Cruz, T., & Unger, J. B. (2018). The why we retweet scale. *PloS One, 13*(10). https://doi.org/10.1371/journal.pone.0206076

Masroor, F., Khan, Q. N., Aib, I., Ali, Z., & .. (2019). Polarization and ideological weaving in Twitter discourse of politicians. *Social Media + Society, 5*(4), 1–14. https://doi.org/10.1177/2056305119891220

McDonald, J. H. (2014). *Small numbers in chi-square and G-tests* (3rd ed.). Sparky House Publishing.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2017). E1071: Misc functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R Package Version 1.6-8, 1.7-0.

Mishra, S., Rizoiu, M.-A., & Xie, L. (2016). Feature driven and point process approaches for popularity prediction. Proceedings of the 25th ACM International Conference on Information and Knowledge Management, 1069–1078. https://doi.org/10.1145/2983323.2983812.

Morente-Molinera, J. A., Kou, G., Samuylov, K., Ureña, R., & Herrera-Viedma, E. (2019). Carrying out consensual Group Decision Making processes under social networks using sentiment analysis over comparative expressions. *Knowledge-Based Systems, 165*, 335–345. https://doi.org/10.1016/j.knosys.2018.12.006

Nesi, P., Pantaleo, G., Paoli, I., & Zaza, I. (2018). Assessing the reTweet proneness of tweets: Predictive models for retweeting. *Multimedia Tools and Applications, 77*(20), 26371–26396. https://doi.org/10.1007/s11042-018-5865-0

Oliveira, N., Costa, J., Silva, C., & Ribeiro, B. (2018). Retweet predictive model for predicting the popularity of tweets. In A. Madureira, A. Abraham, N. Gandhi, C. Silva, & M. Antunes (Eds.), Proceedings of the Tenth International Conference on Soft Computing and Pattern Recognition (SoCPaR 2018) (Vol. 942, pp. 185–193). Springer, Cham.

Pancer, E., & Poole, M. (2016). The popularity and virality of political social media: Hashtags, mentions, and links predict likes and retweets of 2016 U.S. presidential nominees' tweets. *Social Influence, 11*(4), 259–270. https://doi.org/10.1080/15534510.2016.1265582

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas at Austin.

Punjabi, V. D., More, S., Patil, D., Bafna, V., Shah, S., & Bachhav, H. (2019). A survey on trend analysis on Twitter for predicting public opinion on ongoing events. *International Journal of Computer Applications, 180*(26), 13–17. https://doi.org/10.5120/ijca2018916596

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (1.0.136) [Computer software]. http://www.r-project.org/.

Ripley, B. (2012). Classification and regression trees. *R package version, 1*. https://CRAN.R-project.org/package=tree.

Roy, S., Suman, B. K., Chandra, J., & Dandapat, S. K. (2020). Forecasting the future: Leveraging RNN based feature concatenation for tweet outbreak prediction. Proceedings of the 7th ACM IKDD CoDS and 25th COMAD (CoDS COMAD 2020), 219–223. https://doi.org/10.1145/3371158.3371190.

Rudat, A., & Buder, J. (2015). Making retweeting social: The influence of content and context information on sharing news in Twitter. *Computers in Human Behavior, 46*, 75–84. https://doi.org/10.1016/j.chb.2015.01.005

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Scurlock, R., Dolsak, N., & Prakash, A. (2020). Recovering from Scandals: Twitter Coverage of Oxfam and Save the Children Scandals. *Voluntas, 31*(1), 94–110. https://doi.org/10.1007/s11266-019-00148-x

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press.

Shi, J., Hu, P., Lai, K. K., & Chen, G. (2018). Determinants of users' information dissemination behavior on social networking sites: An elaboration likelihood model perspective. *Internet Research, 28*(2), 393–418. https://doi.org/10.1108/IntR-01-2017-0038

Shi, J., Lai, K. K., Hu, P., & Chen, G. (2017). Understanding and predicting individual retweeting behavior: Receiver perspectives. *Applied Soft Computing, 60*, 844–857. https://doi.org/10.1016/j.asoc.2017.08.044

Soboleva, A., Burton, S., Mallik, G., & Khan, A. (2017). 'Retweet for a Chance to…': An analysis of what triggers consumers to engage in seeded eWOM on Twitter. *Journal of Marketing Management, 33*(13-14), 1120–1148. https://doi.org/10.1080/0267257X.2017.1369142

Trilling, D., Tolochko, P., & Burscher, B. (2017). From Newsworthiness to Shareworthiness: How to Predict News Sharing Based on Article Characteristics. *Journalism & Mass Communication Quarterly, 94*(1), 38–60. https://doi.org/10.1177/1077699016654682

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Predicting elections with Twitter: What 140 characters reveal about political sentiment. *Social Science Computer Review, 29*(4), 402–418. https://doi.org/10.1177/0894439310386557

Venables, W., & Ripley, B. (2002). *Modern applied statistics with S* (4th ed.). Springer.

Vijayan, R., & Mohler, G. (2018). Forecasting retweet count during elections using graph convolution neural networks. In *IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 256–262). https://doi.org/10.1109/dsaa.2018.00036

Walker, L. (2016). What factors influence whether politicians' tweets are retweeted? Using CHAID to build an explanatory model of the retweeting of politicians' tweets during the 2015 UK General Election campaign [Doctoral dissertation, Cranfield University]. http://dspace.lib.cranfield.ac.uk/handle/1826/11156.

Walker, L., Baines, P. R., Dimitriu, R., & Macdonald, E. K. (2017). Antecedents of Retweeting in a (Political) Marketing Context: Antecedents of retweeting. *Psychology & Marketing, 34*(3), 275–293. https://doi.org/10.1002/mar.20988

Wang, C., Zhou, Z., Jin, X.-L., Fang, Y., & Lee, M. K. O. (2017). The influence of affective cues on positive emotion in predicting instant information sharing on microblogs: Gender as a moderator. *Information Processing & Management, 53*(3), 721–734. https://doi.org/10.1016/j.ipm.2017.02.003

Wang, Y., Luo, J., Niemi, R., Li, Y., & Hu, T. (2016). Catching fire via "likes": Inferring topic preferences of Trump followers on Twitter. *ArXiv Eprint*. ArXiv:1603.03099v1.

Weiss, G. J., & Davis, R. B. (2019). Discordant financial conflicts of interest disclosures between clinical trial conference abstract and subsequent publication. *PeerJ, 7*, e6423. https://doi.org/10.7717/peerj.6423

Wong, W. S., Amer, M., Maul, T., Liao, I. Y., & Ahmed, A. (2020). Conditional generative adversarial networks for data augmentation in breast cancer classification. In R. Ghazali, N. Nawi, M. Deris, & J. Abawajy (Eds.), Recent Advances on Soft Computing and Data Mining. SCDM 2020. Advances in Intelligent Systems and Computing (Vol. 978, pp. 392–402). Cham: Springer.

Xu, X., Zheng, J., Yang, J.-b., Xu, D.-L., & Chen, Y.-W. (2017). Data classification using evidence reasoning rule. *Knowledge-Based Systems, 116*, 144–151. https://doi.org/10.1016/j.knosys.2016.11.001

Xu, X., Zhao, Z., Xu, X., Yang, J., Chang, L., Yan, X., & Wang, G. (2020). Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models. *Knowledge-Based Systems, 190*, 105324. https://doi.org/10.1016/j.knosys.2019.105324

Yang, J.-B., Liu, J., Wang, J., Sii, H.-S., & Wang, H.-W. (2006). Belief rule-base inference methodology using the evidential reasoning approach - RIMER. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 36*(2), 266–285. https://doi.org/10.1109/TSMCA.2005.851270

Yang, J.-B., Liu, J., Xu, D.-L., Wang, J., & Wang, H. (2007). Optimization Models for Training Belief-Rule-Based Systems. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 37*(4), 569–585. https://doi.org/10.1109/TSMCA.2007.897606

Yang, J.-B., & Xu, D.-L. (2013). Evidential reasoning rule for evidence combination. *Artificial Intelligence, 205*, 1–29. https://doi.org/10.1016/j.artint.2013.09.003

Yang, J.-B., & Xu, D.-L. (2014). A study on generalising Bayesian inference to evidential reasoning. In F. Cuzzolin (Ed.), Belief Functions: Theory and Applications (Vol. 8763, pp. 180–189). Cham: Springer. https://doi.org/10.1007/978-3-319-11191-9_20.

Yang, J.-B., & Xu, D.-L. (2017). Inferential modelling and decision making with data. In *23rd International Conference on Automation and Computing (ICAC)* (pp. 1–6).

Yang, Q., Tufts, C., Ungar, L., Guntuku, S., & Merchant, R. (2018). To Retweet or not to retweet: Understanding what features of cardiovascular tweets influence their retransmission. *Journal of Health Communication, 23*(12), 1026–1035. https://doi.org/10.1080/10810730.2018.1540671

Yuan, N. J., Zhong, Y., Zhang, F., Xie, X., Lin, C.-Y., & Rui, Y. (2016). Who will reply to/retweet this tweet? The dynamics of intimacy from online social interactions. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining* (pp. 3–12). https://doi.org/10.1145/2835776.2835800

Zhang, D.i., Yan, X., Zhang, J., Yang, Z., & Wang, J. (2016). Use of fuzzy rule-based evidential reasoning approach in the navigational risk assessment of inland waterway transportation systems. *Safety Science, 82*, 352–360. https://doi.org/10.1016/j.ssci.2015.10.004

Zhu, H., Yang, J.-B., Xu, D.-L., & Xu, C. (2016). Application of evidential reasoning rules to identification of asthma control steps in children. In *22nd International Conference on Automation and Computing (ICAC)* (pp. 444–449). https://doi.org/10.1109/IConAC.2016.7604960