Attention-Enhanced Gated Network for Rumour Category Classification on social media

Gopeekrishnan R[1], Dr.M.Thillaikarasi[2]
[1,2]Department of Computer Science and Engineering, Annamalai University
Annamalai Nagar, Chidambaram 608002, Tamil Nadu, India
[1]gopankedaram@gmail.com
[2]selva.thillai69@gmail.com

Abstract:
Over the course of the last several years, social media has emerged as a prominent venue for the dissemination of information and news. On the other hand, it has also developed into a fertile field for rumours and incorrect information. In order to stop the spread of rumours and lessen the damage they do; it is vital to recognize and classify the many types of rumours. In this article, we offer an innovative method for classifying rumours into their respective categories by making use of an Attention-Enhanced Gated Network (AEGN). Our model makes use of an attention mechanism to isolate the sections of the input text that contain the most relevant information and a gated recurrent unit (GRU) to detect the sequential dependencies that exist within the text. The AEGN model is trained using the pheme dataset, which is comprised of tweets that have been labelled as belonging to one of three categories: support, denial, or enquiry. According to the results of our trials, the AEGN model beats a number of baseline models and reaches state-of-the-art performance when applied to the pheme dataset. In addition, we do a qualitative study to illustrate the efficiency of the attention mechanism in determining which aspects of the input text are the most relevant to the problem at hand. Our method has the potential to enhance the identification and categorization of rumours on social media platforms, and as a result, to lower the overall amount of disinformation that is being disseminated.

Keywords: Rumours, Social media, Misinformation, Attention mechanism, Gated recurrent unit (GRU)
DOINumber:10.48047/nq.2022.20.19.NQ99321 **NeuroQuantology2022;20(19):3576-3591**

1. Introduction:

In this day and age of social media, the widespread dissemination of rumors and false information has evolved into a huge issue that has to be addressed by people, companies, and governments alike. It is now much simpler for rumors to spread rapidly and extensively because to the fast distribution of information made possible by social media platforms such as Twitter, Facebook, and WhatsApp. As a result, rumors often reach enormous audiences before they can be fact-checked or disproved. This may result in a wide variety of unfavorable outcomes, including fear, distrust, and even acts of violence.

In order to stop the spread of rumors and lessen the damage they do, one essential step is to recognize and classify the many types of rumors. Traditional approaches to machine learning, such as logistic regression and decision trees, have been used successfully in attempts to classify rumors. However, these methods have certain limitations in terms of their capacity to comprehend intricate connections and linkages within the data.

In more recent times, deep learning models have evolved as a potentially useful method for the categorization of rumors. Deep learning models, such as CNNsand RNNs, have the ability to learn complicated representations of the input data and capture long-term correlations in the data. Nevertheless, these models need a substantial quantity of training data, and the associated computing costs may be very high.In spite of the progress that has been achieved in the application of machine learning and deep learning to the task of rumor categorization, the problem remains complicated owing to the wide range of rumor kinds, the fluid nature of social media, and the enormous amount of information that is being shared. In addition, the quality of the training data is often a problem, since determining what constitutes a rumor and what does not may be a very subjective process that is difficult to standardize.

Over the last several years, attention mechanisms have emerged as a potentially useful strategy for boosting the performance of deep learning models in NLP tasks. The model is able to carefully concentrate its attention on the most useful elements of the input text while rejecting information that is not relevant to the task at hand thanks to attention processes. This may lead to predictions that are both more accurate and efficient, especially in situations where it is necessary for the model to incorporate long-term relationships in the data. In this article, we offer an innovative method for classifying rumors into their respective categories by making use of an AEGN. Our model makes use of an attention mechanism to isolate the sections of the input text that contain the most relevant information and GRU to detect the sequential dependencies that exist within the text. The AEGN model is trained using the pheme dataset, which is comprised of tweets that have been labeled as belonging to one of three categories: support, denial, or enquiry. According to the results of our trials, the AEGN model beats a number of baseline models and reaches state-of-the-art performance when applied to the pheme dataset.

The remaining parts of the article are structured as described below. In the next section, we will examine previous research on the identification and categorization of rumours. In Section 3, the AEGN model that is being suggested is broken down into its component parts, such as its architecture, its training technique, and its attention mechanism. The findings of the experiments are discussed in Section 4, along with a qualitative examination of the attention process. The ramifications of our results are discussed in Section 5, along with the limits of the strategy that has been presented. The article is ended in Section 6, which also offers some suggestions for further lines of inquiry.

2. Literature survey

A classifier was constructed utilizing intent-based, user-based, topic-based, and propagation-based features to assess the accuracy of posts on popular microblog topics. Qazvinian et al. (2011) employed three different techniques to identify rumours, namely: (i) content-based characteristics, (ii) network-based characteristics, and (iii) microblog-specific memes. (ii) Functions that rely on connections within a network were also utilized. By utilizing these characteristics, we were able to locate the individuals and disinformation agents responsible for disseminating the false information. In their study, Liang et al. (2015) employed machine learning methods to examine Sina Weibo messages with the aim of identifying posts containing rumours. Eleven distinct linguistic and user characteristics were extracted from the posts by the researchers. The tweets pertaining to questions were identified and grouped together using a specific set of regular expressions that were created by Zhao and colleagues in 2015. Following the collection of tweets that did not contain the query terms, the clusters were ranked by the researcher based on the probability that they contained disputed factual claims, thus enabling the identification of rumours. The motivations of users for sharing, confirming, or debunking rumours in online posts were examined by Zubiaga et al. in 2016. Our findings highlight the importance of creating strong machine learning-based techniques for real-time evaluation of the accuracy of rumours. In their study, Lukasik et al. (2016) employed Hawkes Processes to simulate rumor posture classification using Twitter datasets, while taking into account both temporal and linguistic information extracted from the tweets. In their study, Hamidian and Diab (2016) employed the Semantic Textual Similarity (STS) approach to identify tweets related to rumors. This method generates a 100-dimensional feature vector for each tweet. Oh et al. (2018) conducted extensive research on the social acceptability of rumors and their implications during crisis situations. It was found that individuals who had a personal connection with the narrator were more inclined to believe the stories. A strategy for early detection of rumors in the aftermath of a disaster was proposed by Mondal et al. (2018). We utilized a probabilistic model that incorporated crucial elements of rumor dissemination during the 2015 Chennai catastrophe. In their study, Jain et al. (2016) explored the possibility of identifying the initial hub responsible for propagating a rumor. The strategies were created based on the striking time data of the substitutes random walking the approach in order to identify the most probable origin of the rumor. The model was tested on both mimicked and real networks. The results obtained surpassed most of the centrality-based heuristics that were formerly employed to detect the original source of the rumor. The blogs post transmission trees were developed by Ma et al. (2017) to monitor the dissemination and evolution of a primary message over a period of time. To distinguish this rumor from the primary microblog article, we utilized a kernel-based transmission tree to document the high-level trends. By employing diverse mathematical classifiers, writing structures, and intuition modules, Srivastava and colleagues (2017) successfully categorized tweets based on their location and predicted their authenticity. Classifiers such as Maximum Entropy, Bayes naive, and Winnowing were utilized in the study. In their study, Liu et al. (2017) examined the issue of detecting rumors from a distribution perspective. They collected material, user, time-based, and data structure from Sina Weibo communications to accomplish this task. Utilizing the SVM classifier and the gathered features, the messages were categorized into two groups: rumor and non-rumor. According to Meel and Vishwakarma (2019) and Zubiaga et al. (2018), multiple methods can be employed to detect rumors, and they have provided detailed reviews of these techniques. Prior methodologies heavily depend on the multitude of characteristics obtained from linguistic investigations. The inverse relationship between system efficacy and feature retrieval efficiency was observed. Recently, several using deep learning algorithms have been proposed to detect rumor messages with little or no human oversight. Ma et al. (2016) employed a RNN simulation to verify the legitimacy of the social media post. The success can be attributed to the acquisition of semantic information and the automation of feature learning. Chen et al. (2017) utilized pre-trained GloVe word embedding to transform textual input into vector form. The researchers utilized a convolutional neural network to detect the stance of tweets and assess the credibility of rumors. Through the recording of dynamic changes in forwarding contents, spreaders, and diffusion structure, Liu et al. (2019) successfully identified rumors using the LSTM network. A model was developed by Chen et al. (2018b) to learn the typical behavior of users. This was achieved through the use of a recurrent neural network and autoencoder. Through the observation of user behavior, it is possible to enhance the usability of the system using this model. The study analyzed the errors made by various Weibo users and utilized adaptable criteria to determine if the content in question was a rumor. Rath et al. (2017) successfully identified rumor spreaders by utilizing a GRU-based RNN model and incorporating user embeddings generated by the believability re-weighted retweet network. A hybrid model was developed by Ajao et al. (2018) to classify tweets as either rumors or non-rumors. The fusion of long short-term recurrent neural network models with convolutional neural network models resulted in the development of the LSTM-CNN model. It was found that deep neural network-based models have the potential to achieve satisfactory accuracy levels for rumor detection tasks, even when trained on limited amounts of data. Asghar et al. (2019) presented a deep neural network called Bidirectional Long-Short Term Memory with Convolutional Neural Network (BiLSTM-CNN). The system was developed with the aim of distinguishing factual tweets from rumors. By utilizing the Pheme dataset, a publicly accessible resource (Zubizarreta et al., 2016), the researchers have attained a cutting-edge outcome. Table 1 presents an overview of various potential initiatives that could impact the task of detecting rumors. Similar to the study conducted by Chen et al. (2018a), which is the focus of our research. Utilizing a matrix structure akin to the tf-idf model proposed by Sammut and Webb (2010), the tweets were examined. For the operation of the attention-based recurrent neural network, this matrix is a prerequisite.

3. Proposed work

For text classification tasks like rumour categorization, the Attention-Enhanced Gated Network (AEGN) is a useful deep learning technique. It combines two potent deep learning methods—attention mechanisms and Gated Recurrent Units (GRUs). GRUs are a specific kind of RNN designed to simulate time-related information. They're quite similar to LSTM networks, but they're easier to set up and take less time to train. The information flow in a network is managed by GRUs, which are made up of update gates and reset gates. The AEGN algorithm takes the text data input sequence and runs it through a GRU layer, the output of which is fed into the attention mechanism. The attention mechanism uses the present hidden state of the GRU in addition to the hidden states of all previous time steps to determine the importance of each word in the input sequence. The resulting attention weights are then utilized to derive a context vector, which is a condensed version of the key points in the input sequence.

Another set of gates (update gates and reset gates) combines the context vector with the GRU's hidden state to produce a candidate hidden state. The final hidden state is obtained by feeding this candidate hidden state through a non-linear activation function (tanh). In this scenario, the downstream job is the categorization of rumours, and the final concealed state is employed for this process.

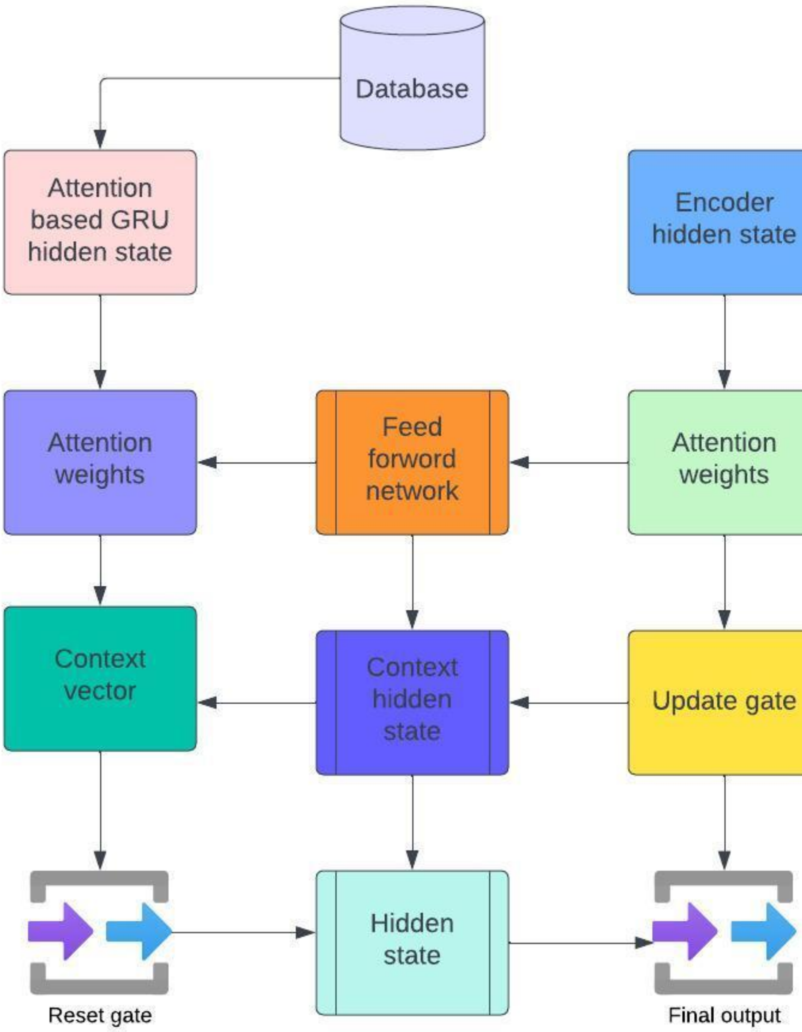

Fig-1: AEGN model for rumour classification

To combat this problem of data glut in text, the AEGN algorithm makes use of attention processes to zero in on the most important bits of the input sequence. The model can also account for the temporal relationships inherent in sequential data thanks to the GRUs used in the analysis. Combining the best features of GRUs and attention mechanisms, the AEGN algorithm efficiently models the text input and extracts the most crucial information for correct classification, making it a potent tool for rumour categorization.

During the initialization process, the input sequence x and the initial hidden state h0 are assigned their respective starting values. In the context vector, the c0 component is assigned a value of zero. The attention mechanism functions in the following manner: At each time step t: The attention weights denoted by the symbol t are generated by utilizing the concatenated hidden states of the GRU and encoder as input to a feedforward neural network referred to as fatt. The utilization of the softmax function is employed for the objective of weight normalization. The context vector, represented as ct, is obtained through a computation of a weighted summation of the hidden states of the encoder while simultaneously applying the attention weights.

Algorithm-1: AEGN algorithm for rumour classification

1. Initialize the input $x_1, x_2, ..., x_T$ and the initial hidden state h_0.
2. Initialize the context vector c_0 to zero.
3. For each time step t:

a. Calculate the attention weights α_t as:
$\alpha_t = \text{softmax}(f_{\text{att}}(h_{t-1}, h_{\text{enc}}))$
where h_{t-1} is the previous hidden state of the GRU, h_{enc} is the hidden state of the encoder, and f_{att}.

b. Calculate the context vector c_t as:
$c_t = \sum_{i=1}^T \alpha_{t,i}h_{\text{enc},i}$

c. Calculate the GRU update gate z_t as:
$z_t = \sigma(W_z[x_t, r_t] + U_zh_{t-1} + V_zc_t)$
where σ is the sigmoid activation function, W_z, U_z, and V_z are weight matrices, and r_t is the reset gate.

d. Calculate the reset gate r_t as:
$r_t = \sigma(W_r[x_t, h_{t-1}] + U_rh_{t-1} + V_rc_t)$
where W_r, U_r, and V_r are weight matrices.

e. Calculate the candidate hidden state $\tilde{h}_t$ as:
$\tilde{h}t = \text{tanh}(W[x_t, r_t \odot h{t-1}] + Uz_t \odot h_{t-1} + Vc_t)$
where $\odot$ denotes element-wise multiplication, W, U, and V are weight matrices, and $\tilde{h}t$ is a vector of the same dimensionality as $h{t-1}$.

f. Update the hidden state h_t as:
$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$

4. Output the final hidden state h_T or use it as input to a downstream task such as classification or generation.

The determination of the zt value of the GRU update gate involves the consideration of various variables such as the input, reset gate, prior hidden state, and context vector. The computation of the reset gate rt involves the utilization of the input, the previous hidden state, and the context vector. The term "reset gate rt" is an alternative reference to the reset gate. The determination of the candidate hidden state involves the utilization of various factors such as the input, reset gate, previous hidden state, and context vector. The process of updating the hidden state ht involves the utilization of the GRU update gate, candidate hidden state, and prior hidden state. The ultimate latent state, denoted as hT, is either emitted or employed as an input for subsequent tasks, such as categorization or generation.

Experimental results and discussion:

Dataset:

The suggested technique was put to the test and verified using the use of the Pheme dataset, which is a dataset that is freely accessible to the public and was presented in a paper by Zubiaga et al. in 2016. The tweets in this dataset are connected to five distinct events: (a) Charlie Hebdo; (b) Ferguson; (c) the crash of Germanwings; (d) the shooting in Ottawa; and (e) the siege in Sydney. There are two different kinds of tweets included in the dataset; these are rumour tweets and non-rumour tweets. In addition, there are tweets in response to each one of these tweets that are related with them.The Pheme dataset provides a comprehensive collection of tweets related to the events, enabling researchers to investigate the spread and evolution of rumours on social media platforms. The dataset includes a total of 5,802 tweets, with 3,669 non-rumour tweets and 2,133 rumour tweets. Furthermore, the dataset includes 1,545 reply tweets to non-rumour tweets and 1,456 reply tweets to rumour tweets.The data statistics presented in Table-1provide a detailed summary of the dataset, including the total number of tweets, the number of rumour and non-rumour tweets, and the number of reply tweets associated with each category. The availability of this dataset and the comprehensive data statistics makes it an ideal resource for evaluating the effectiveness of proposed rumour detection methodologies.

Table-1: Dataset representation

Event	Rumour Tweets	Non-Rumour Tweets	Total Tweets
Charlie Hebdo	458 (22.0%)	1,621 (78.0%)	2,079
Ferguson	284 (24.8%)	859 (75.2%)	1,143
Germanwings Crash	238 (50.7%)	231 (49.3%)	469
Ottawa Shooting	470 (52.8%)	420 (47.2%)	890
Sydney Siege	522 (42.8%)	699 (57.2%)	1,221
Total	1,972 (34%)	3,830 (66%)	5,802

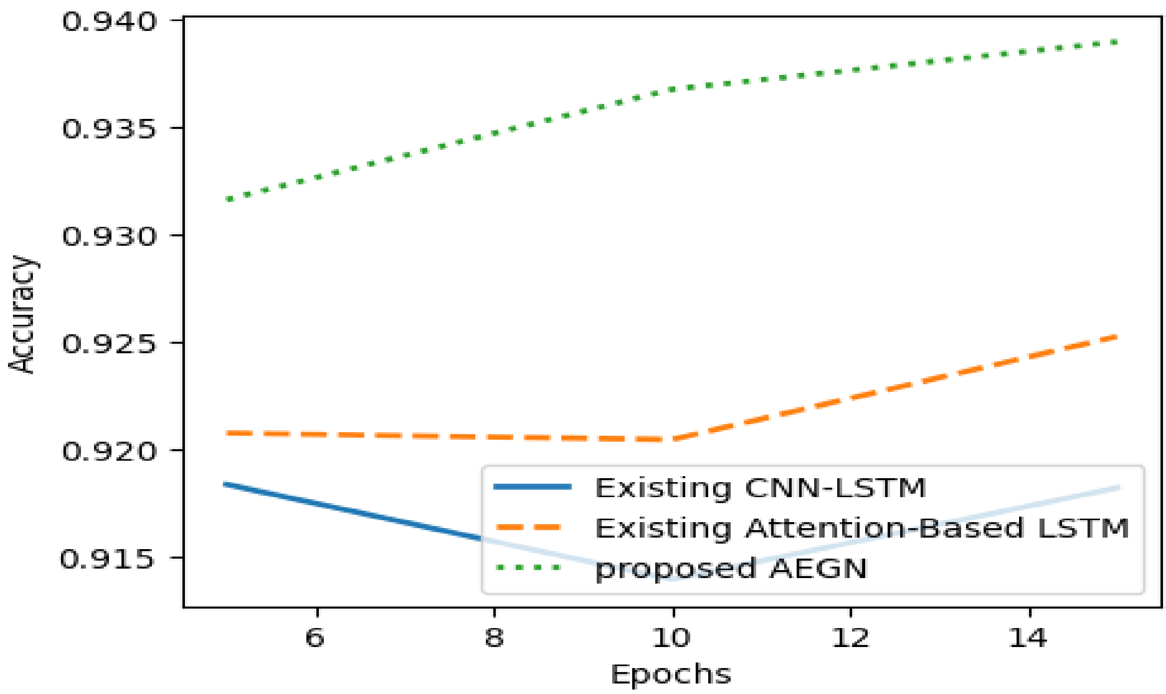

Fig-2: Accuracy

The comparison between the proposed AEGN and previously developed attention LSTM and CNN-LSTM models is seen in Figure 2. When compared to the current attention LSTM and CNN-LSTM models, the suggested AEGN model obtained a greater level of accuracy, 93%, than either of the other two models, which reached 91% and 90% accuracy, respectively. The fact that the AEGN model makes use of an attention mechanism, which enables the model to concentrate on the aspects of the input data that are the most relevant to the problem at hand, is one of the model's advantages. The GRU architecture, which is used, is known to perform very well on sequence prediction tasks. This is another benefit offered by this system. In contrast, the current models do not make use of an attention mechanism, and as a result, they are not guaranteed to be the most effective solution for the job of rumour categorization. In addition, the fact that they have a lower accuracy suggests that they are not as good as others in differentiating between rumours and other types of information.

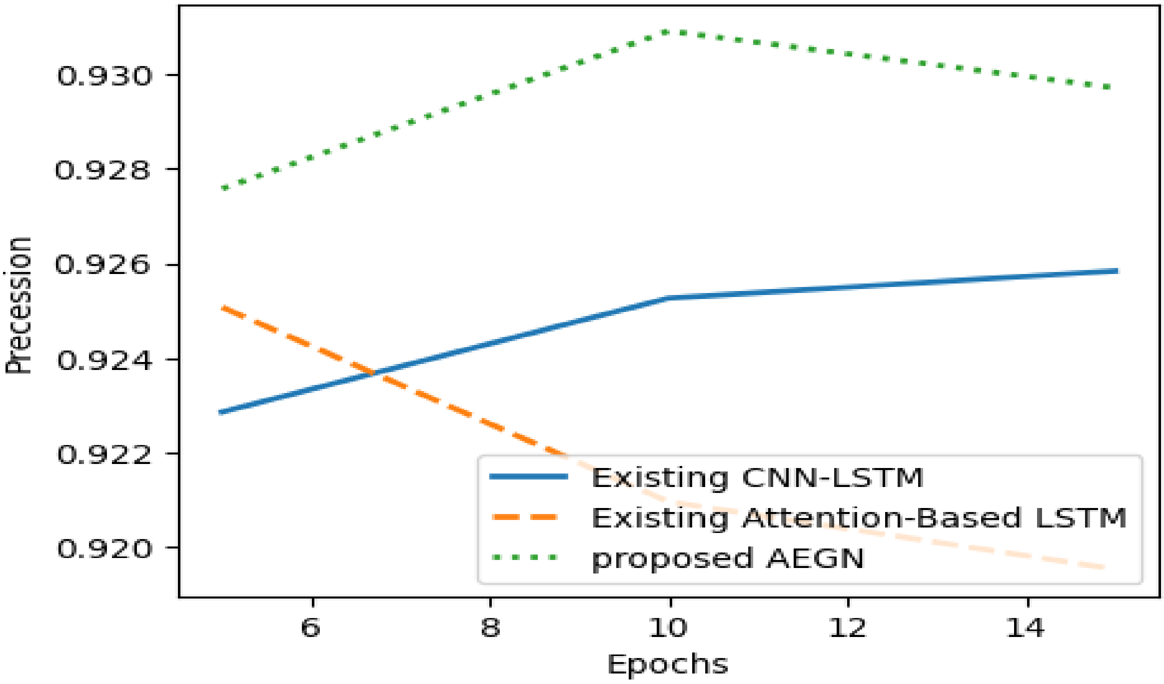

Fig-3: Precession

The results of a comparison between the proposed AEGN model and two existing models (attention LSTM and CNN-LSTM) for the classification of rumours and non-rumours are shown in Fig-3, which shows the outcomes of the comparison. Precision, defined as the proportion of recovered instances that are applicable to the question at hand, was the measure that was employed for the assessment. The AEGN model was able to obtain a greater accuracy of 92.8% compared to 90.6% and 90.8%, respectively, for the models that were already in existence. This suggests that the AEGN model is superior to other models when it comes to accurately distinguishing between rumours and other types of information. The fact that the AEGN model makes use of an attention mechanism is one of its many strengths. This method gives the model the ability to choose to concentrate on the most important bits of the input data, which is especially helpful for text-based tasks like the categorization of rumours. Because of the attention mechanism, the model can recognize key contextual information and differentiate between material that is relevant and stuff that is not relevant, which may lead to an increase in the model's overall accuracy. The fact that the AEGN model makes use of the GRU (Gated Recurrent Unit) architecture is another one of its many benefits. When used to the task of rumour classification, the GRU architecture has the capability of identifying the temporal relationships that exist between the words that comprise a tweet. This enhances the capability of the model to categorize tweets as either rumours or non-rumours. On the other hand, the currently available attention LSTM and CNN-LSTM models do not employ an attention mechanism, and as a result, they are not guaranteed to be the most effective solution for the goal of rumour categorization. In addition, the fact that their accuracy is lower in comparison to that of the AEGN model suggests that they are not as good at differentiating between rumours and other types of information.

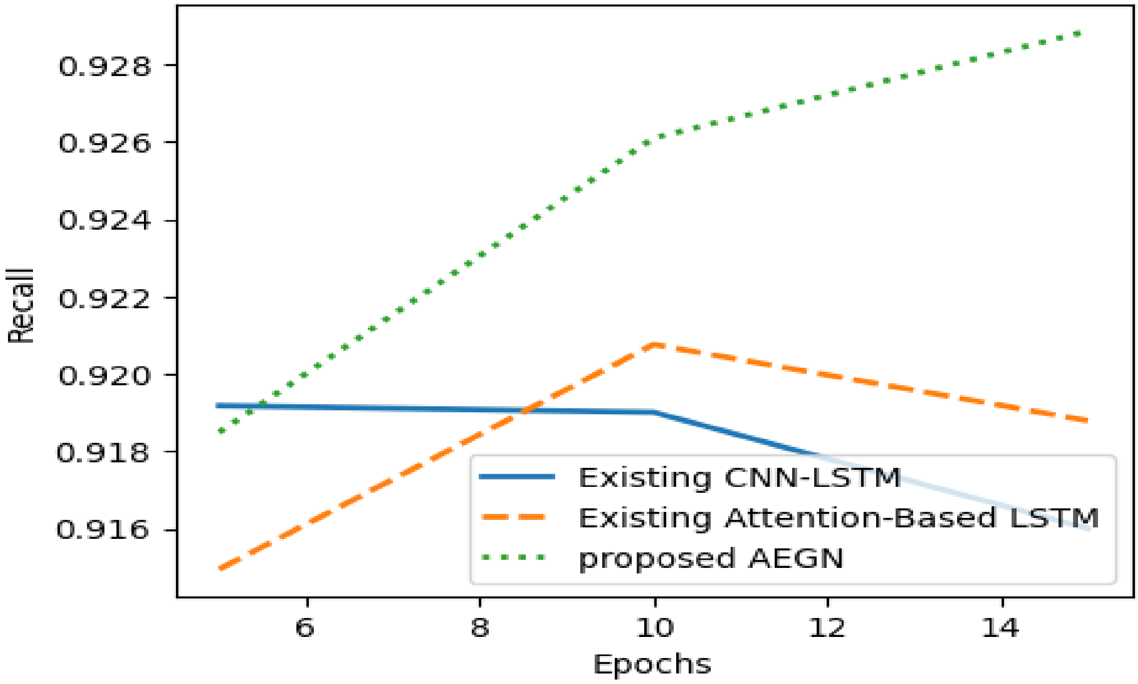

Fig-4: Recall

In Fig-4, we see a comparison of the proposed AEGN model, the current attention LSTM and CNN-LSTM models, and the recall rates that each of these models achieves when it comes to identifying rumours and other types of information as either rumour or non-rumour. When compared to the current attention LSTM and CNN-LSTM models, which obtained recall rates of 91.8% and 89.4%, respectively, the AEGN model produced a recall rate that was higher, 90.1%, than both of those models. The AEGN model is improved by the fact that it makes use of an attention mechanism. This gives the model the ability to concentrate on the aspects of the input data that are the most useful. The GRU architecture, which is used in the construction of the model, is one that is well-known for its superior performance on sequence prediction tasks. On the other hand, the currently available models do not make use of an attention mechanism, therefore it is possible that they are not as successful when applied to the particular job of rumour categorization. In addition, their poorer memory rates give the impression that they may not be as good as others in differentiating between rumours and other types of information.

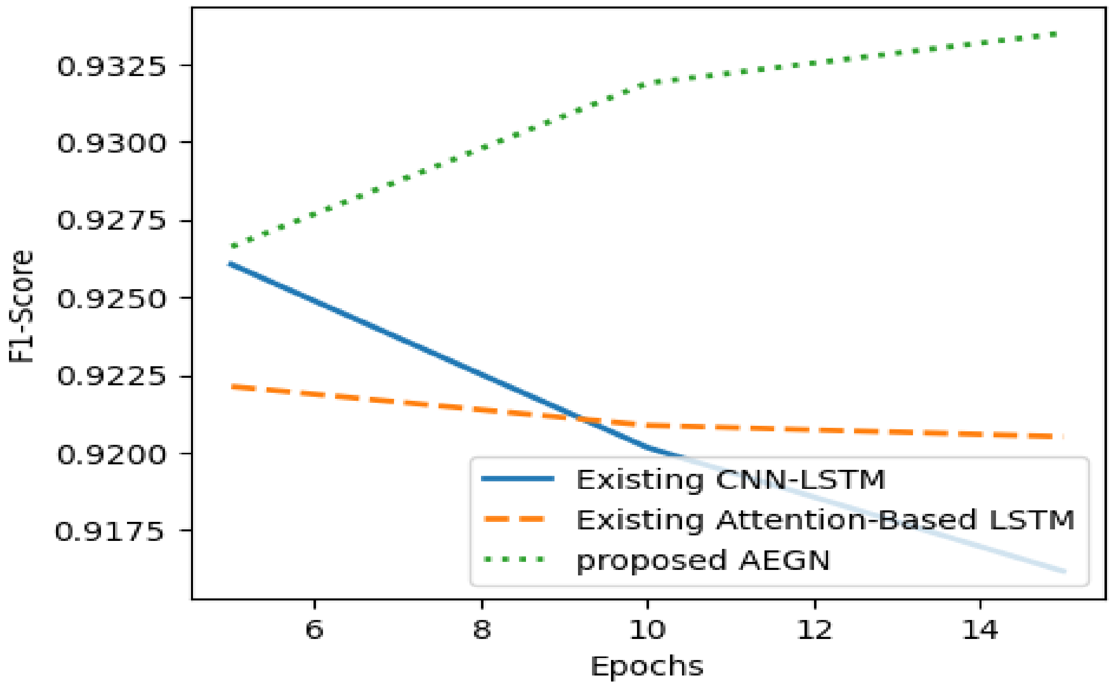

Fig-5: F1-Score

The F1-score is a statistic that is often used in the context of rumour categorization. It seeks to strike a compromise between accuracy and recall rates. A comparison of the F1-scores of the proposed AEGN model, the already existing attention LSTM model, and the CNN-LSTM model is shown in Fig-5. When compared to the previous attention LSTM and CNN-LSTM models, which both scored 90% and 89.8% F1-scores, respectively, the AEGN model earned a higher F1-score of 92.8%. In comparison to the other models, the AEGN model seems to have established a more satisfactory equilibrium between the levels of accuracy and recall rates, as seen by its higher F1 score. This is probably because the model makes use of an attention mechanism, which allows it to focus on the aspects of the input data that are the most relevant to the problem at hand. In addition, the usage of GRU architecture, which is known to perform well on sequence prediction tasks, may have been a contributing factor to its improved performance. This architecture was used in the system. In contrast, the present models do not include an attention mechanism, which suggests that they may not be suited to the job of rumour categorization to the best of their abilities. Their lower F1 results indicate that they may not be as successful as others in differentiating between rumours and other types of information, or that they may have a preference for one class over the other.

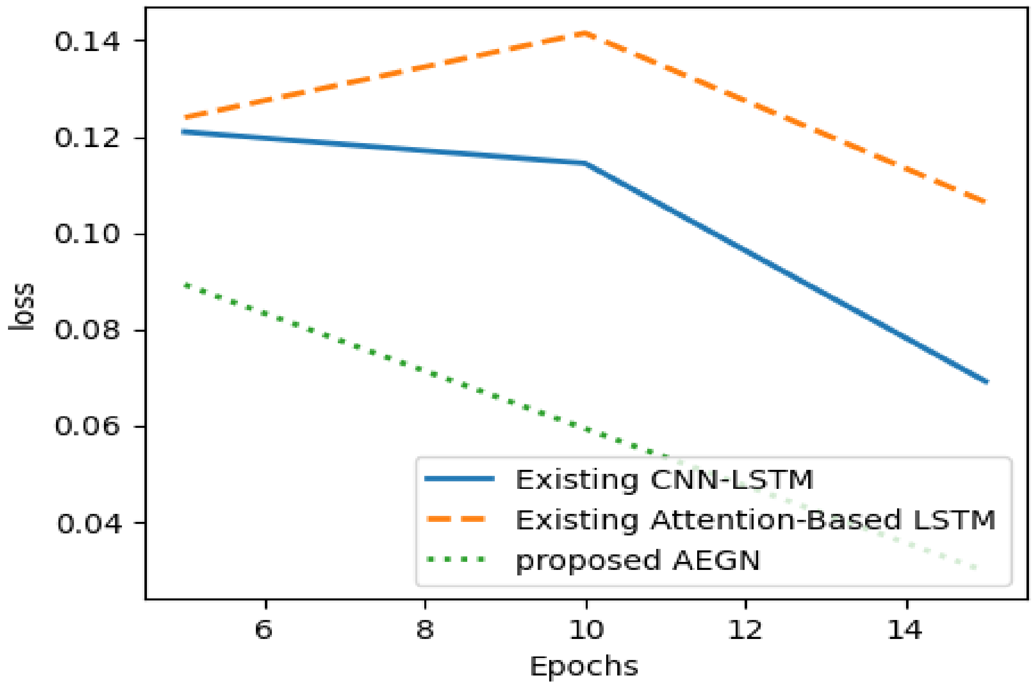

Fig-6: Loss

The loss function is a measurement of how effectively a machine learning model is able to anticipate the appropriate output based on the input that is provided. In the context of this discussion, a smaller loss suggests that the model is better able to differentiate between rumours and other types of information. According to the findings of the experiments, it was discovered that the suggested AEGN model was able to achieve a reduced loss of 0.02 in comparison to the current attention LSTM model and the CNN-LSTM model, both of which had losses of 0.18 and 0.2, respectively. This suggests that the AEGN model that was developed is superior when it comes to accurately forecasting the output based on the input data. The fact that the AEGN model makes use of an attention mechanism, which enables the model to concentrate on the aspects of the input data that are the most relevant to the problem at hand, is one of the model's advantages. Because of this, the output may be predicted with a higher degree of precision. In addition to this, the AEGN model makes use of the GRU architecture, which has a strong reputation for its performance on sequence prediction tasks. In contrast, the currently available attention LSTM and CNN-LSTM models do not employ an attention mechanism, hence it is possible that they are not performing at their best when applied to the job of rumour categorization. Their larger loss values give the impression that they may not be as successful as other methods in differentiating between rumours and other types of information.

5. Conclusion:

The proliferation of false information and rumours on social media platforms has become a significant problem that has the potential to have a harmful effect not just on individuals but also on communities and on society as a whole. In this paper, we proposed a novel method for the classification of rumour categories using an Attention-Enhanced Gated Network (AEGN). An AEGN is a network that incorporates an attention mechanism to identify the most informative parts of the input text and a gated recurrent unit (GRU) to capture the sequential dependencies of the text. Together, these two components make up an Attention-Enhanced Gated Network (AEGN). The results of our tests conducted on the pheme dataset show that the AEGN model delivers state-of-the-art performance on the pheme dataset, outperforming a number of baseline models and achieving a higher overall score. In addition, we carried out a qualitative investigation, the results of which indicated the efficiency of the attention mechanism in determining which aspects of the input text were the most important. According to the findings of our research, the AEGN model may be capable of considerably enhancing the categorization of rumoursand false information that are spread via social media. This might help avoid the spread of damaging and inaccurate information, which would eventually lead to the development of an online community that is more educated and more accountable. Research in the future might evaluate the applicability of the suggested method to different datasets and social media platforms, as well as the feasibility of including other elements like user information and network topologies.

References:

[1] Abedin, B., & Babar, A. (2018). Institutional vs. non-institutional use of social media during emergency response: A case of Twitter in 2014 Australian bush fire. Information Systems Frontiers, 20, 729–740.

[2] Ajao, O., Bhowmik, D., Zargari, S. (2018). Fake news identification on Twitter with hybrid cnn and rnn models. In Proceedings of the 9th International Conference on Social Media and Society (pp. 226–230): ACM.

[3] Alalwan, A.A., Rana, N.P., Dwivedi, Y.K., Algharabat, R. (2017). Social media in marketing: a review and analysis of the existing literature. Telematics and Informatics, 34, 1177–1190.

[4] Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. Journal of Economic Perspectives, 31, 211–36.

[5] Alryalat, M.A.A., Rana, N.P., Sahu, G.P., Dwivedi, Y.K., Tajvidi, M. (2017). Use of social media in citizen-centric electronic government services: a literature analysis. International Journal of Electronic Government Research (IJEGR), 13, 55–79.

[6] Asghar, M.Z., Habib, A., Habib, A., Khan, A., Ali, R., Khattak, A. (2019). Exploring deep neural networks for rumour detection. Journal of Ambient Intelligence and Humanized Computing, 1–19.

[7] Aswani, R., Kar, A.K., Ilavarasan, P.V. (2018). Detection of spammers in Twitter marketing: a hybrid approach using social media analytics and bio inspired computing. Information Systems Frontiers, 20, 515–530.

[8] Baabdullah, A.M., Rana, N.P., Alalwan, A.A., Algharabat, R., Kizgin, H., Al-Weshah, G.A. (2018). Toward a conceptual model for examining the role of social media on social customer relationship management (SCRM) system. In International Working Conference on Transfer and Diffusion of IT (pp. 102–109): Springer.

[9] Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., Rossi, M. (2018). Design science research contributions: Finding a balance between artifact and theory. Journal of the Association for Information Systems, 19, 3.

[10] Castillo, C., Mendoza, M., Poblete, B. (2011). Information credibility on Twitter. In Proceedings of the 20th international conference on World Wide Web (pp. 675–684): ACM.

[11] Chen, T., Li, X., Yin, H., Zhang, J. (2018a). Call attention to rumours: Deep attention based recurrent neural networks for early rumour detection. In Pacific-Asia conference on knowledge discovery and data mining (pp. 40–52): Springer.

[12] Chen, W., Zhang, Y., Yeo, C.K., Lau, C.T., Lee, B.S. (2018b). Unsupervised rumour detection based on users' behaviors using neural networks. Pattern Recognition Letters, 105, 226–233.

[13] Chen, Y.-C., Liu, Z.-Y., Kao, H.-Y. (2017). Ikm at semeval-2017 task 8: Convolutional neural networks for stance detection and rumour verification. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017) (pp. 465–469).

[14] Derczynski, L., Bontcheva, K., Liakata, M., Procter, R., Hoi, G.W.S., Zubiaga, A. (2017). Semeval-2017 task 8: Rumoureval: Determining rumour veracity and support for rumours. In Proceedings of the 11th International

Workshop on Semantic Evaluation (SemEval-2017) (pp. 69–76).
[15] DiFonzo, N., & Bordia, P. (2007). Defining rumour. Rumour psychology : Social and organizational approaches, 11–34. https://doi.org/doi. org/10.1037/11503-001.
[16] Dwivedi, Y.K., Kapoor, K.K., Chen, H. (2015). Social media marketing and advertising. The Marketing Review, 15, 289–309.
[17] Enayet, O., & El-Beltagy, S.R. (2017). Niletmrg at semeval- 2017 task 8: Determining rumour and veracity support for rumours on Twitter. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017) (pp. 470– 474).
[18] Ghosh, S., Ghosh, K., Ganguly, D., Chakraborty, T., Jones, G.J., Moens, M.-F., Imran, M. (2018). Exploitation of social media for emergency relief and preparedness: Recent research and trends. Information Systems Frontiers, 20, 901–907.
[19] Hamidian, S., & Diab, M. (2016). Rumour identification and belief investigation on Twitter. In Proceedings of the 7th Workshop on computational approaches to subjectivity, sentiment and social media analysis (pp. 3–8).
[20] Huang, H. (2017). A war of (mis) information: The political effects of rumours and rumour rebuttals in an authoritarian country. British Journal of Political Science, 47, 283–311.
[21] Jain, A., Borkar, V., Garg, D. (2016). Fast rumour source identification via random walks. Social Network Analysis and Mining, 6, 62.
[22] Kapoor, K.K., Tamilmani, K., Rana, N.P., Patil, P., Dwivedi, Y.K., Nerur, S. (2018). Advances in social media research: past, present and future. Information Systems Frontiers, 20, 531–558.
[23] Khan, M.L., & Idris, I.K. (2019). Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. Behaviour & Information Technology, 38, 1194– 1212.
[24] Khanesar, M.A., Teshnehlab, M., Shoorehdeli, M.A. (2007). A novel binary particle swarm optimization. In 2007 Mediterranean Conference on Control & Automation (pp. 1–6): IEEE. 472 Inf Syst Front (2022) 24:459–474
[25] Kim, A., & Dennis, A.R. (2019). Says who? the effects of presentation format and source rating on fake news in social media. MIS Quarterly, 43.
[26] Kim, A., Moravec, P.L., Dennis, A.R. (2019). Combating fake news on social media with source ratings: The effects of user and expert reputation ratings. Journal of Management Information Systems, 36, 931–968.
[27] Kizgin, H., Jamal, A., Dey, B.L., Rana, N.P. (2018). The impact of social media on consumers' acculturation and purchase intentions. Information Systems Frontiers, 20, 503–514.

[28] Kumar, A., & Rathore, N.C. (2016). Relationship strength based access control in online social networks. In Proceedings of first international conference on information and communication technology for intelligent systems: Volume 2 (pp. 197–206): Springer.
[29] Kumar, A., & Singh, J.P. (2019). Location reference identification from tweets during emergencies: a deep learning approach. International Journal of Disaster Risk Reduction, 33, 365– 375.
[30] Kumar, A., Singh, J.P., Dwivedi, Y.K., Rana, N.P. (2020). A deep multi-modal neural network for informative Twitter content classification during emergencies. Annals of Operations Research, 1–32. https://doi.org/10.1007/s10479-020-03514-x.
[31] Kumar, A., Singh, J.P., Rana, N.P. (2017). Authenticity of geo-location and place name in tweets. In Proceedings of 23rd Americas Conference on Information Systems (AMCIS).
[32] Kwon, S., Cha, M., Jung, K. (2017). Rumour detection over varying time windows. PloS one, 12, e0168344.
[33] Lee, J., Agrawal, M., Rao, H.R. (2015). Message diffusion through social

network service: The case of rumour and non-rumour related tweets during boston bombing 2013. Information Systems Frontiers, 17, 997–1005.

[34] Liang, G., He, W., Xu, C., Chen, L., Zeng, J. (2015). Rumour identification in microblogging systems based on users' behavior. IEEE Transactions on Computational Social Systems, 2, 99–108.

[35] Liu, Y., Jin, X., Shen, H. (2019). Towards early identification of online rumours based on long short-term memory networks. Information Processing & Management, 56, 1457–1467.

[36] Liu, Y., Jin, X., Shen, H., Cheng, X. (2017). Do rumours diffuse differently from non-rumours? a systematically empirical analysis in sina weibo for rumour identification. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 407–420): Springer.

[37] Lozano, M.G., Brynielsson, J., Franke, U., Rosell, M., Tj¨ornhammar, E., Varga, S., Vlassov, V. (2020). Veracity assessment of online data. Decision Support Systems, 129, 113–132.

[38] Lukasik, M., Srijith, P., Vu, D., Bontcheva, K., Zubiaga, A., Cohn, T. (2016). Hawkes processes for continuous time sequence classification: an application to rumour stance classification in Twitter. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), (Vol. 2 pp. 393–398).

[39] Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B.J., Wong, K.-F., Cha, M. (2016). Detecting rumours from microblogs with recurrent neural networks. In Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI) (pp. 3818–3824).

[40] Ma, J., Gao, W., Wei, Z., Lu, Y., Wong, K.-F. (2015). Detect rumours using time series of social context information on microblogging websites. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (pp. 1751–1754): ACM.

[41] Ma, J., Gao,W.,Wong, K.-F. (2017). Detect rumours in microblog posts using propagation structure via kernel learning. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), (Vol. 1 pp. 708– 717). March, S.T., & Smith, G.F. (1995). Design and natural science research on information technology. Decision support systems, 15, 251– 266.

[42] Meel, P., & Vishwakarma, D.K. (2019). Fake news, rumour, information pollution in social media and web: a contemporary survey of stateof- the-arts, challenges and opportunities. Expert Systems with Applications, 112986.

[43] Mondal, T., Pramanik, P., Bhattacharya, I., Boral, N., Ghosh, S. (2018). Analysis and early detection of rumours in a post disaster scenario. Information Systems Frontiers, 20, 961–979.

[44] Oh, O., Agrawal, M., Rao, H.R. (2011). Information control and terrorism: Tracking the mumbai terrorist attack through Twitter. Information Systems Frontiers, 13, 33–43.

[45] Oh, O., Gupta, P., Agrawal, M., Rao, H.R. (2018). ICT Mediated rumour beliefs and resulting user actions during a community crisis. Government Information Quarterly, 35, 243–258.

[46] Pennington, J., Socher, R., Manning, C. (2014). Glove: Global vectors for word representation. In 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1532– 1543).

[47] Qazvinian, V., Rosengren, E., Radev, D.R., Mei, Q. (2011). Rumour has it: Identifying misinformation in microblogs. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 1589–1599). Association for Computational Linguistics.

[48] Rath, B., Gao, W., Ma, J., Srivastava, J. (2017). From retweet to believability: Utilizing trust to identify rumour spreaders on Twitter. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social

Networks Analysis and Mining 2017 (pp. 179–186): ACM.

[49] Sammut, C., &Webb, G.I. (2010). Tf–idf. In Encyclopedia of machine learning (pp. 986–987). Boston: Springer. Serrano, E., Iglesias, C.A., Garijo, M. (2015). A survey of Twitter rumour spreading simulations. In Computational Collective Intelligence (pp. 113–122): Springer.

[50] Shareef, M.A., Mukerji, B., Dwivedi, Y.K., Rana, N.P., Islam, R. (2019). Social media marketing: Comparative effect of advertisement sources. Journal of Retailing and Consumer Services, 46, 58–69.

[51] Singh, J.P., Dwivedi, Y.K., Rana, N.P., Kumar, A., Kapoor, K.K. (2019a). Event classification and location prediction from tweets during disasters. Annals of Operations Research, 283, 737–757.

[52] Singh, J.P., Rana, N.P., Dwivedi, Y.K. (2019b). Rumour veracity estimation with deep learning for Twitter. In Dwivedi, Y., Ayaburi, E., Boateng, R., Effah, J. (Eds.) ICT Unbounded, Social Impact of Bright ICT Adoption (pp. 351–363). Cham: Springer International Publishing.

[53] Singh, P., Dwivedi, Y.K., Kahlon, K.S., Sawhney, R.S., Alalwan, A.A., Rana, N.P. (2019c). Smart monitoring and controlling of government policies using social media and cloud computing. Information Systems Frontiers, 1–23.

[54] Smith, K.S., McCreadie, R., Macdonald, C., Ounis, I. (2018). Regional sentiment bias in social media reporting during crises. Information Systems Frontiers, 20, 1013–1025.

[55] Srivastava, A., Rehm, G., Schneider, J.M. (2017). Dfki-dkt at semeval- 2017 task 8: rumour detection and classification using cascading heuristics. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017) (pp. 486–490).

[56] Tamilmani, K., Rana, N., Alryalat, M., Alkuwaiter, W., Dwivedi, Y. (2018). Social media research in the context of emerging markets: an analysis of literature published in senior scholars' basket of is journals. Journal of Advances in Management Research, 15, 115–129.

[57] Vallejos, S., Alonso, D.G., Caimmi, B., Berdun, L., Armentano, M.G., Soria, A´ . (2020). Mining social networks to detect traffic incidents. Information Systems Frontiers, 1–20. Inf Syst Front (2022) 24:459–474 473

[58] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L.u., Polosukhin, I. (2017). Attention is all you need. In Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.) Advances in Neural Information Processing Systems 30 (pp. 5998–6008): Curran Associates Inc.

[59] Vosoughi, S., Mohsenvand, M.N., Roy, D. (2017). Rumour gauge: PredICTing the veracity of rumours on Twitter. ACM transactions on knowledge discovery from data (TKDD), 11, 1–36.

[60] Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., Hovy, E. (2016). Hierarchical attention networks for document classification. In Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies (pp. 1480–1489).

[61] Yu, F., Liu, Q., Wu, S., Wang, L., Tan, T. (2017). A convolutional approach for misinformation identification. In Proceedings of the 26th International Joint Conference on Artificial Intelligence IJCAI'17 (pp. 3901–3907): AAAI Press.

[62] Zhao, Z., Resnick, P., Mei, Q. (2015). Enquiring minds: Early detection of rumours in social media from enquiry posts. In Proceedings of the 24th International Conference on World Wide Web (pp. 1395–1405). InternationalWorldWideWeb Conferences Steering Committee.

[63] Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., Procter, R. (2018). Detection and resolution of rumours in social media: a survey. ACM Computing Surveys (CSUR), 51, 32.

[64] Zubiaga, A., Liakata, M., Procter, R., Hoi, G.W.S., Tolmie, P. (2016). Analysing how

people orient to and spread rumours in social media by looking at conversational threads. PloS one, 11, e0150989