Identifying Ironic Content Spreaders on Twitter using Psychometrics, Contextual and Ironic Features with Gradient Boosting Classifier

Notebook for PAN at CLEF 2022

Ehsan Tavan*[1], Maryam Najafi*[1] and Reza Moradi*[1]

[1]*NLP Department, Part AI Research Center, Tehran, Iran*

Abstract

The study of irony detection on social networks has gained much attention in recent years. As part of this task, a collection of users' tweets is provided, and the goal is to determine if these users are spreaders of irony or not. As we hypothesized that user-generated content is affected by the author's psychometric characteristics, contextual information, and irony features in the text, we investigated the effects of incorporating this information to identify ironic spreaders. Using the emotion and personality detection module, we were able to extract the author's psychometric features. A pre-trained language model based on SBERT and T5-based architecture has been employed to extract context-based features. Our paper describes a framework using the author's psychometric, contextual, and ironic features in a Gradient Boosting classifier based on our theory. Experimental results in this paper demonstrate the importance of this combination in identifying ironic spreader users. As a result, we were able to achieve an accuracy of 95.00% and 93.81% with 5-fold and 10-fold cross-validation respectively on the IROSTEREO training dataset. However, on the official PAN test set, our system attained a 88.89% score.

Keywords

Stereotype, Author Profiling, Irony Detection, Sarcasm Detection, Language Psychology, Deep Learning

1. Introduction

Social media networks have become a platform for people with different intellectual, ideological, and psychological characteristics to share their thoughts, opinions, and interests. In view of the high importance and comprehensiveness of the information published in these networks, it is urgent to develop automatic tools for processing it. Furthermore, users often use informal text in addition to the usual linguistic complexities such as slang, idioms, typos, and grammatical errors. By contrast, people often convey their meaning in complex ways. Using Figurative Language (FL) is one way to emphasize one's point of view.

FL utilizes linguistic features such as sarcasm and irony, which are common in user generated content on platforms including Facebook and Twitter. Despite some attempts to provide a good definition of irony, there is still no consensus in the research community. But almost all

---

researchers agree on two ironic features. First, it is unclear exactly what the author meant by his words. In fact, it uses words contrary to the original meaning. The second is that irony arises from opinions, emotions, feelings, and thoughts [1, 2, 3, 4, 1].

Irony and sarcasm are two related concepts that are sometimes used interchangeably in the literature, but there are subtle differences between them [1, 5]. Sarcasm's rarity, infrequency, difficulty in detection, and ambiguous meaning are its most challenging aspects. A positive word that is used in sarcasm can imply a negative meaning. For example, in "If voting by mail is good enough for Trump, it is good enough for me." despite the use of positive words, polarity is negative due to the presence of ironic content.

Recently, the research community has become more interested in identifying users who consistently publish certain types of textual content, such as Fake News (FN) and Hate Speech (HS). PAN introduces hate speech and fake news spreader profiling tasks in 2020 and 2021 demonstrating the importance of Author Profiling (AP) in social media networks [6, 7]. These tasks focused on identifying spreaders as an AP task rather than the published text to prevent the publication of hateful and fake news. Now, in Profiling Irony and Stereotype Spreaders on Twitter (IROSTEREO) task in PAN@2022 [8, 9], the goal is to identify spreaders of ironic content on Twitter. Especially, those prone to use irony to spread stereotypes against women or the LGBT community [1]. The ultimate goal is to identify and profile spreaders of ironic content to prevent the spread of this type of content. As far as we know, this task has not been done before.

In the IROSTEREO task, the main challenge is to get a comprehensive representation of the user's tweets. User tweets contain meaningful information that indicates the user's personality type, opinion, and thoughts. Using this wide range of information in a user-level representation is challenging. Therefore in this research, we propose a parallel framework for extracting user's representation based on contextualize, psychometric and ironic features. The proposed framework includes three modules: 1) Contextualized embedding module, 2) Psychometric embedding module, and 3) Irony embedding module. The sentence transformers (SBERT) are used in the contextual embedding module to extract context features from user's tweets. Also, in the psychometric embedding module and irony embedding module, we use deep learning architecture by using the Text-To-Text Transfer Transformer (T5) language model. To prepare user-level representation by these modules, we first extract the representation for each user's tweets. By averaging this tweet-level representation embedding, the final user-level representation embedding is obtained. After extracting the user-level embedding vector, the Gradient Boosting classifier is used to classify users into ironic and non-ironic categories. By using the concatenating contextualized, irony, and psychometric embedding vector, the proposed framework has achieved 88.89% accuracy on the official PAN test set. Our code is available at GitHub[2] for researchers.

The rest of the paper is organized as follows: In Section 2 we present the related works; in Section 3 we described the characteristics of dataset. In Section 4 we present our proposed method; the experiments and results are presented in Section 5, and finally, our conclusion and future works are in Section 6.

---

[1]https://pan.webis.de/clef22/pan22-web/author-profiling.html
[2]https://github.com/MarSanTeam/Author_Profiling

2. Related Works

IROSTEREO is a challenging interdisciplinary sub-task of NLP that sits between Irony Detection and AP. Therefore, both Irony detection and AP were investigated in the following sections.

2.1. Irony Detection

Due to the high semantic similarity of irony and sarcasm as well as the comprehensiveness of our studies, research on both tasks has been considered. In [10], a hybrid method called CASCADE is proposed to use both context and content-based information. It used user embedding and discourse embedding to capture the stylometric, and personality traits of users, and also topical information. On the other hand, it used content-based features with a CNN-based network to capture local patterns. Some research has shown the remarkable effect of the use of sentiment analysis in irony detection. For instance, [2] used transfer learning from sentiment resources. They found that sentiment knowledge has valuable information for irony detection to overcome their implicit incongruity. The author of [11], used the multi-head attention bidirectional LSTM-based network to recognize sarcasm comments.

Recently, attention is focused on using the pre-trained language model. [12] proposed an end-to-end architecture that uses a pre-trained transformer, namely RCNN-RoBERTa. Moreover, in [13] a RoBERTa-based model is proposed to demonstrate the importance of using contextual information. The [14] proposed a BERT-based model using weight loss techniques to overcome data unbalance problems and ensemble learning for better generalization. Additionally, [15] introduced a pre-trained transformer model and aspect-based sentiment analysis to identify whether response comments in a dialogue are sarcastic or not.

2.2. Author Profiling

Due to the significant increase in the research community's attention to AP, much research has been done in recent years. Some of these works have used deep learning methods but others investigated hybrid methods. In the following sections, these methods are described in more detail.

Deep-learning-based Methods:
In recent years, the two-step method has become the most popular in this field. These methods aim to extract user profiles in a two-step way. In the first and second steps, a tweet-level and user-level representation is extracted respectively. [16] proposed a two-step method. In the first step, they extract the contextualized vector for tweets to make a tweet-level representation, and in the second, aggregate them by averaging to make a user-level representation. For the tweet-level representation, They examined different variants of the BERT model and found that the best result was achieved by the BERT-base model that was fine-tuned by the external corpus of Twitter for sentiment analysis [17]. Finally, they used the user-level representation to identify haters. In [18], two methods were proposed for the Hate Speech task (HS) at PAN 2021 in the first method, tweets were classified based on the hate label in the tweet-level method. Then, users were classified based on the average label of user tweets. In another method, they used the proposed method in [16] with a little modification. They found that the user-level method

with the BERT model outperforms the other baselines. Likewise, [19] introduced a method like what was proposed in [18] that uses BERTweet instead of the BERT model. Although, the CNN-based model is in the literature, in [20] authors proposed a CNN-based model for handling the Hate Speech Profiling in English and Spanish on PAN@2020. During this competition, the proposed method won the competition. It was based on a 1D-Convolutional layer in addition to local and global average pooling.

Hybrid methods:

The main purpose of these methods is to use the techniques of other NLP sub-tasks and sciences such as cognitive science methods. In the study concluded by [21], a hybrid approach was proposed that used contextualized representation (CR) and character-based statistical embedding, named Char-LDSE. For CR, used extractive summarization and the RoBERTa language model to extract the semantically important tweets. Then, used these representations to make an ensemble. These representations were used to make an ensemble. However, in [22] an interesting idea was proposed based on using concepts and coarse-grained to handle the FN problem. They used a variant of ConceptNet [23] by the name ConceptNet Numberbatch to make the conceptual representation. On the other hand, n-gram features and the tfidf weighting method were used to make a coarse-grained representation. These two forms of representation were combined and fed to a CNN-based network. Regardless of these methods' abilities and the promising results they have achieved, they are still far from competing with state-of-the-art methods.

3. Dataset

The IROSTEREO task dataset consists of 420 anonymous authors in English. The XML file is provided for each author containing 200 tweets. Thus, the author is classified as either an ironic or non-ironic spreader. There is an equal distribution for each category in the dataset, with each category having 210 authors. Additionally, it is important to note that all URLs, hashtags, and mentioned or retweeted users were masked by standard tokens.

4. Proposed Method

Our methodology for IROSTEREO uses the proposed dataset in PAN@2022 and includes the following modules: Contextualized Embedding (CE), Psychometric Features Embedding (PFE), Irony Embedding (IRE) module. The proposed framework is shown in Figure 1. CE, PFE, and IRE were developed using deep learning methods and language models. Each module aims to make a user-level representation from its perspective. The user-level representation is derived by averaging the tweet-level representation. An overall user-level representation is obtained by concatenating representations derived from CE, PFE, and IRE. The Gradient Boosting Classifier was used over this representation to classify ironic and non-ironic users based on user representations. Below are details regarding each component.
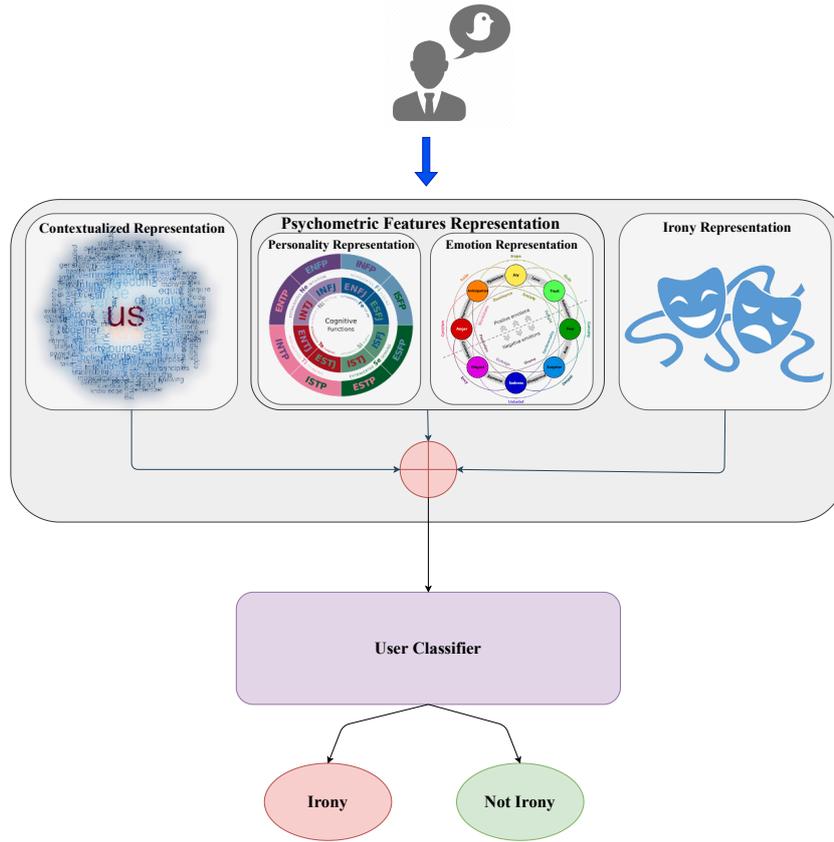
Figure 1: The architecture of the Proposed Method

4.1. Contextualized Embedding Module

Creating a contextualized representation for each tweet is the main goal of the module. The use of contextual information is critical for irony detection because a high degree of irony detection is based on context. The transformer-based language models have excelled in capturing contextualized embedding, they have performed very well in various NLP-based tasks. Thus, Sentence-BERT (SBERT) [24], a variant of BERT that was developed to determine sentence semantic similarity was used in our paper. With SBERT, semantically meaningful sentence embedding is calculated using a siamese architecture. We derive sentence embeddings from word embeddings by performing a mean pooling process.

4.2. Psychometric Features Embedding Module

According to language psychology, the words that people use in everyday life are influenced by their psychological processes, including thoughts, feelings, emotional states, and personality [25]. Understanding an author's psychological state can be enhanced by identifying personality and emotional features in the text. [26] shows that identifying these states improves the accuracy of irony recognition models. Therefore, to identify the author's psychological state, we have
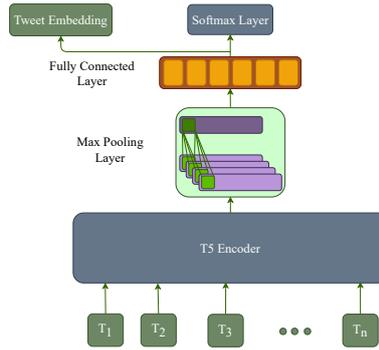
Figure 2: The architecture of the proposed model for the Personality Embedding, Emotion Embedding, and Irony Embedding Module. The input of this model is just a single tweet.

considered a representation of emotion and personality.

4.2.1. Personality Embedding Module

The personality embedding module is trained to recognize the author's personality type based on the Myers Briggs Type Indicator (MBTI)[3] dataset. In MBTI dataset, texts were labeled by 16 distinct personality types along 4 axes. A count of 8675 different samples was included in this dataset, including MBTI type of each author and their last 50 tweets.

We prepare a personality representation for each tweet by using a deep learning model based on the Text-to-Text Transfer Transformer (T5) language model [27]. The T5 language model was shown high performance in various NLP applications. The proposed architecture for this module is shown in Figure 2.

In order to train personality embedding module, we utilize the last hidden-states of the T5 encoder model to prepare the word representation vector. Then, the most appropriate features are extracted using max-pooling and a fully connected layer. The output label is then predicted using a softmax layer. After training thid module personality representation for each tweet is obtained by the fully connected layer output. Then, an average of all tweet-level representations for each user uses as a personality embedding of user.

4.2.2. Emotion Embedding Module

Making user-level emotion embedding is the aim of this module. Studies have shown the high influence of emotional features on irony detection models [28, 29, 30]. The use of emotion features has improved the accuracy of the irony detection task in many studies. Thus, we used the described method in Section 4.2.1 and the CrowdFlower[4] dataset to make user-level emotion representation. CrowdFlower is an emotion detection dataset from Kaggle that was used for training this module. This dataset has 13 different emotion labels and 4000 samples.

---

[3]https://www.kaggle.com/datasets/datasnaek/mbti-type
[4]https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text

4.3. Irony Embedding Module

As described in Section 3, the IROSTEREO dataset is not labeled at the tweet level. However, with the theory that irony spreader users publish more ironic content, we examined the impact of using ironic features on IROSTEREO performance. To accomplish this, we used a model with the same architecture as shown in Figure 2 at the tweet level and train this model with WLV[5] irony detection dataset [31]. Using the method described in Section 4.2.1, we calculated user-level irony representation by averaging tweet-level irony representation.

4.4. User Embedding Module

To create a final user-level embedding, we concatenate Contextualized, Personality, Emotion, and Irony user-level embedding. Eventually, the final user-level embedding was created with dimensions of 4096, which in addition to context-based and irony-based features, also includes emotion-based and personality-based features called psychometrics features.

4.5. Classifier

This study uses the Gradient Boosting Classifier to identify ironic-content spreaders on Twitter using the introduced features. The idea behind gradient boosting is to take a weak hypothesis or weak learning algorithm and make a series of changes to it that will improve the strength of the hypothesis or learning algorithm. we utilized the Log Loss function [32] as the loss function and Friedman MSE [33] as criteria for stopping splitting in the classifier.

5. Experiments and Results

In this section, we review the results of various experiments. we use two metrics accuracy and confidence interval (CI) with 95% confidence, to evaluate different classifiers and features. The main evaluation metric in the IROSTEREO task is accuracy and the systems will be ranked by it. CI is a range of estimation on a parameter that is performed on accuracy here. This is a range of values that will expect the estimation to fall into it if the experiment was repeated.

In our experiments, we evaluate the performance of our proposed framework by using context-base, personality-base, emotion-base, and irony-base features, as well as a combination of theirs on TIRA platform [34]. Also, in the experiments, two classifiers, support vector machine and Gradient Boosting, are used and their performance is compared. Table 1. shows the results of the experiments using 5-fold and 10-fold cross validation.

Based on the results shown in Table 1, the use of context-base, personality-base, emotion-base, and irony-base features can be used in the IROSTEREO task as a suitable features. Using the Gradient Boosting classifier, which according to the experiments has a much better performance compared to the support vector machine, the accuracy of the model using intorduced features in 5-fold cross validation is 91.43, 91.43, 88.1 and 92.14 percent respectively. The accuracy of the model when using the context-base features that extracted by the SBERT is better than other features. This can can be due to the training of the SBERT model on a large amount of data and

---

power of the SBERT to extract context-base features. Also, the performance of the model using psychometric-base and irony-base features is very acceptable due to the small amount of data to train the appropriate model to extract these features.

Different experiments have been performed to obtain the best performance of the model when concatenation the four introduced features. Table 1 shows these experiments results.

Feature	Methods							
	Gradient Boosting Classifier				SVM(SVC)			
	5-fold		10-fold		5-fold		10-fold	
	ACC	CI	ACC	CI	ACC	CI	ACC	CI
Personality	91.43	5.89	91.19	8.35	61.19	10.42	62.14	14.61
Irony	91.43	5.89	90.71	8.49	61.19	10.42	62.14	14.61
Emotion	88.1	6.88	88.1	9.43	76.19	9	76.67	12.6
Context	92.14	5.62	90.48	8.44	92.86	5.35	93.33	6.8
Context ⊕ Personality	92.38	5.60	93.10	7.31	92.38	10.42	92.41	14.61
Context ⊕ Irony	92.62	5.55	92.86	7.45	92.36	5.52	92.28	7.28
Context ⊕ Emotion	94.14	4.89	93.57	7.04	61.19	5.49	62.14	7.31
Personality ⊕ Irony	91.67	5.84	91.43	8.21	60.22	5.53	62.43	7.33
Personality ⊕ Emotion	91.67	5.82	91.90	8.03	76.67	8.98	77.62	12.16
Emotion ⊕ Irony	91.67	5.82	92.14	7.95	76.67	8.98	77.86	9.53
Context ⊕ Irony ⊕ Personality	92.38	5.62	92.38	7.74	92.14	5.46	92.14	7.78
Context ⊕ Irony ⊕ Emotion	94.76	4.59	93.57	7.11	90.95	6.06	90.71	8.6
Context ⊕ Personality ⊕ Emotion	94.52	4.67	93.33	7.25	90.95	6.06	90.71	8.6
Personality ⊕ Irony ⊕ Emotion	91.9	5.7	92.14	7.95	76.43	9.03	77.14	12.3
Context ⊕ Irony ⊕ Emotion ⊕ Personality	**95.00**	**4.5**	**93.81**	**6.93**	90.95	6.06	91.19	8.35

Table 1
Results of experiments with four introduced features and Gradient Boosting and SVM classifier. 5-fold and 10-fold cross validation are used in experiments.

Because the use of these four features helps the model to extract different aspects of the user's tweets features, it is expected that by concatenation them, the model performance will improve for recognize ironic authors. So we concatenated the features in different modes. In concatenated the features in pairs, the highest accuracy is related to the concatenation of context-base and emotion-base features. The accuracy of the model by concatenating these features and using the Gradient Boosting classifier in 5-fold cross validation is 94.14, which is higher accuracy compared to concatenating other features. The emotion-base feature has lower accuracy than other features when used separately, but by concatenating this feature with a context-based feature, the accuracy of the model is greatly improved, indicating that a concatenating of these two features can be used to improved the performance of model.

In the next step, we concatenated the features in threes. Based on the result of Table 1, the concatenation of context-base and emotion -base features with irony-base and personality-base features has helped to increase the accuracy of the model in 5-fold cross validation. Using a concatenation of context-base, irony-base and emotion-base features, the model has reached 94.76 accuracy in 5-fold cross validation.

In the last step, we concatenated all four introduced features ,and with the Gradient Boosting

in 5-fold cross validation, the accuracy is 95% accuracy. Based on these results, it can be concluded that the concatenation of context-base, irony-base and also psychometric features can be very helpful to identify ironic authors.

One of the metrics that has been examined in all these experiments is the CI. This metric indicates that if another experiment is performed on other data, the accuracy of the model is expected to be within a certain range. The value obtained for this metric in the concatenation of all four introduced features in 5-fold cross validation is 4.5%. So, we can claim that the true classification accuracy of the model is likely between 99.5% and 91.5%. In fact if we repeat this experiment several times and use different data each time, we would find that for approximately 95% (we use this confidence value for all of our experiments) of these experiments, the calculated interval would contain the true accuracy.

6. Conclusions and Future Works

In this paper, we proposed a framework for the Profiling Irony and Stereotype Spreaders on the Twitter task in PAN 2022. As part of our approach, in the first step, we extracted features at the tweet level, and in the second step, we built a user-level representation of each user. The novelty of this work was emphasizing the Psychometric, Ironic, and Context-based features. We demonstrated how context-based, irony-related, and psychometric features affect system performance for distinguishing ironic from non-ironic Twitter authors. Finally, in the 10-fold cross-validation and test sets published by the organizers, we achieved an average accuracy of 93.81% and 88.89% in the train and test sets, respectively. In future works, we would like to examine different language models over the proposed framework and other architectures to achieve the best performance of the current framework. Also, we plan to investigate a variety of features like sentiment analysis to boost the current user-level representation.

References

[1] R. Ortega-Bueno, F. Rangel, D. Hernández Farıas, P. Rosso, M. Montes-y Gómez, J. E. Medina Pagola, Overview of the task on irony detection in spanish variants, in: Proceedings of the Iberian languages evaluation forum (IberLEF 2019), co-located with 34th conference of the Spanish Society for natural language processing (SEPLN 2019). CEUR-WS. org, volume 2421, 2019, pp. 229–256.

[2] S. Zhang, X. Zhang, J. Chan, P. Rosso, Irony detection via sentiment-based transfer learning, Information Processing & Management 56 (2019) 1633–1644.

[3] K. Buschmeier, P. Cimiano, R. Klinger, An impact analysis of features in a classification approach to irony detection in product reviews, in: Proceedings of the 5th workshop on computational approaches to subjectivity, sentiment and social media analysis, 2014, pp. 42–49.

[4] C. Van Hee, Can machines sense irony?: exploring automatic irony detection on social media, Ph.D. thesis, Ghent University, 2017.

[5] E. Filatova, Irony and sarcasm: Corpus generation and analysis using crowdsourcing., in: Lrec, Citeseer, 2012, pp. 392–398.

[6] F. Rangel, G. Sarracén, B. Chulvi, E. Fersini, P. Rosso, Profiling hate speech spreaders on twitter task at pan 2021, in: CLEF, 2021.

[7] F. Rangel, A. Giachanou, B. H. H. Ghanem, P. Rosso, Overview of the 8th author profiling task at pan 2020: Profiling fake news spreaders on twitter, in: CEUR Workshop Proceedings, volume 2696, Sun SITE Central Europe, 2020, pp. 1–18.

[8] J. Bevendorff, B. Chulvi, E. Fersini, A. Heini, M. Kestemont, K. Kredens, M. Mayerl, R. Ortega-Bueno, P. Pezik, M. Potthast, F. Rangel, P. Rosso, E. Stamatatos, B. Stein, M. Wiegmann, M. Wolska, E. Zangerle, Overview of PAN 2022: Authorship Verification, Profiling Irony and Stereotype Spreaders, and Style Change Detection, in: M. D. E. F. S. C. M. G. P. A. H. M. P. G. F. N. F. Alberto Barron-Cedeno, Giovanni Da San Martino (Ed.), Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Thirteenth International Conference of the CLEF Association (CLEF 2022), volume 13390 of *Lecture Notes in Computer Science*, Springer, 2022.

[9] O.-B. Reynier, C. Berta, R. Francisco, R. Paolo, F. Elisabetta, Profiling Irony and Stereotype Spreaders on Twitter (IROSTEREO) at PAN 2022, in: CLEF 2022 Labs and Workshops, Notebook Papers, CEUR-WS.org, 2022.

[10] D. Hazarika, S. Poria, S. Gorantla, E. Cambria, R. Zimmermann, R. Mihalcea, Cascade: Contextual sarcasm detection in online discussion forums, arXiv preprint arXiv:1805.06413 (2018).

[11] A. Kumar, V. T. Narapareddy, V. A. Srikanth, A. Malapati, L. B. M. Neti, Sarcasm detection using multi-head attention based bidirectional lstm, Ieee Access 8 (2020) 6388–6397.

[12] R. A. Potamias, G. Siolas, A.-G. Stafylopatis, A transformer-based approach to irony and sarcasm detection, Neural Computing and Applications 32 (2020) 17309–17320.

[13] T. Dadu, K. Pant, Sarcasm detection using context separators in online discourse, in: Proceedings of the Second Workshop on Figurative Language Processing, 2020, pp. 51–55.

[14] S. Jiang, C. Chen, N. Lin, Z. Chen, J. Chen, Irony detection in the portuguese language using bert, Proceedings http://ceur-ws. org ISSN 1613 (2021) 0073.

[15] S. Javdan, B. Minaei-Bidgoli, et al., Applying transformers and aspect-based sentiment analysis approaches on sarcasm detection, in: Proceedings of the Second Workshop on Figurative Language Processing, 2020, pp. 67–71.

[16] E. Finogeev, M. Kaprielova, A. Chashchin, K. Grashchenkov, G. Gorbachev, O. Bakhteev, Hate speech spreader detection using contextualized word embeddings, in: CLEF, 2021.

[17] A. Go, R. Bhayani, L. Huang, Twitter sentiment classification using distant supervision, CS224N project report, Stanford 1 (2009) 2009.

[18] T. Anwar, Identify hate speech spreaders on twitter using transformer embeddings features and automl classifiers (2021).

[19] R. L. Tamayo, D. C. Castro, R. O. Bueno, Deep modeling of latent representations for twitter profiles on hate speech spreaders identification task (2021).

[20] M. Siino, E. Di Nuovo, I. Tinnirello, M. La Cascia, Detection of hate speech spreaders using convolutional neural networks, in: CLEF, 2021.

[21] H. B. Giglou, T. Rahgooy, J. Razmara, M. Rahgouy, Z. Rahgooy, Profiling haters on twitter using statistical and contextualized embeddings, in: CLEF, 2021.

[22] H. B. Giglou, J. Razmara, M. Rahgouy, M. Sanaei, Lsaconet: A combination of lexical and conceptual features for analysis of fake news spreaders on twitter., in: CLEF (Working

Notes), 2020.

[23] R. Speer, J. Chin, C. Havasi, Conceptnet 5.5: An open multilingual graph of general knowledge, in: Thirty-first AAAI conference on artificial intelligence, 2017.

[24] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, arXiv preprint arXiv:1908.10084 (2019).

[25] S. Argamon, S. Dhawle, M. Koppel, J. W. Pennebaker, Lexical predictors of personality type, in: Proceedings of the 2005 Joint Annual Meeting of the Interface and the Classification Society of North America, 2005, pp. 1–16.

[26] S. Poria, E. Cambria, D. Hazarika, P. Vij, A deeper look into sarcastic tweets using deep convolutional neural networks, arXiv preprint arXiv:1610.08815 (2016).

[27] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu, Exploring the limits of transfer learning with a unified text-to-text transformer, arXiv preprint arXiv:1910.10683 (2019).

[28] S. Frenda, V. Patti, Computational models for irony detection in three spanish variants, in: 2019 Iberian Languages Evaluation Forum, IberLEF 2019, volume 2421, CEUR-WS, 2019, pp. 297–309.

[29] D. I. H. Farías, V. Patti, P. Rosso, Irony detection in twitter: The role of affective content, ACM Transactions on Internet Technology (TOIT) 16 (2016) 1–24.

[30] A. Reyes, P. Rosso, D. Buscaldi, From humor recognition to irony detection: The figurative language of social media, Data & Knowledge Engineering 74 (2012) 1–12.

[31] O. Rohanian, S. Taslimipoor, R. Evans, R. Mitkov, Wlv at semeval-2018 task 3: Dissecting tweets in search of irony, Association for Computational Linguistics, 2018.

[32] V. Vovk, The fundamental nature of the log loss function, in: Fields of Logic and Computation II, Springer, 2015, pp. 307–318.

[33] J. H. Friedman, Greedy function approximation: a gradient boosting machine, Annals of statistics (2001) 1189–1232.

[34] M. Potthast, T. Gollub, M. Wiegmann, B. Stein, TIRA Integrated Research Architecture, in: N. Ferro, C. Peters (Eds.), Information Retrieval Evaluation in a Changing World, The Information Retrieval Series, Springer, Berlin Heidelberg New York, 2019. doi:10.1007/978-3-030-22948-1_5.