Multidimensional and Multilingual Emotional Analysis

Sofia Aparicio[1(✉)], Joao T. Aparicio[1(✉)], and Manuela Aparicio[2(✉)]

[1] INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Rua Alves Redol 9, 1000-029 Lisbon, Portugal
{sofia.aparicio,joao.aparicio}@tecnico.ulisboa.pt

[2] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Lisbon, Portugal
manuela.aparicio@novaims.unl.pt

Abstract. In order to monitor informal political online discussions and to lead a better understanding of hate speech on social media, we found that it was necessary to use sentiment quantification for languages with few training datasets. Previous studies mainly rely on languages with enough data to train a model. Several statistical and machine learning models were produced and compared in three languages (English, Portuguese and Polish). This work shows promising results when inferring sentimental dimensions, even in languages other than English.

Keywords: Emotional ratings of text · Affective norms · Long Short-Term Memory · Recurrent Neural Networks · Machine learning

1 Introduction

Sentiment analysis relies on the necessity to extract either negative and positive evaluation or estimate emotion. Two leading families of methods have been developed to represent human emotions [8]. One is categorical, based on six universal basic emotions (BE) [7]. The other is dimensional, advocating continuous numerical values that progress through multiple dimensions [23]. Since it takes significant human resources to annotate words and textual utterances regarding sentiment, it was necessary to produce automatic methods to infer sentiment. Several studies were conducted using convolutional neural networks (CNN) (i.e., that consider the spatial organization of a sentence) and recurrent neural networks (RNN) (i.e. that consider the sequential organization) [1,5,26]. These works were mainly applied to the English language. So, to the best of my knowledge, there is still a gap when using deep learning to quantify sentiment from languages with few or no training resources. The results show that three trained models performed better (Attention Concat, Attention Feature Bassed, and Attention Affine Transformation); however, the Average word-level prediction model also showed promising results. LSTM tends to perform slightly better than CNN models. The difference was more evident in the arousal dimension.

A. Rocha et al. (Eds.): WorldCIST 2023, LNNS 802, pp. 13–22, 2024.
https://doi.org/10.1007/978-3-031-45651-0_2

2 Related Work

2.1 Assigning Emotion to Textual Utterances

In emotion analysis, word-level prediction differs greatly from assigning emotion values to larger linguistic units, such as paragraphs and sentences. [3] recognized three different approaches for emotion detection: keyword-based, learning-based, and hybrid. However, all these methods resort to different linguistic analysis tools (e.g., semantic level, sentence segmentation, parts of speech recognition, token level). However, word-level problem solving cannot solve high-level linguistic prediction because of the way these words are combined [12]. The other use sequential input data, typical for recurrent neural networks (RNN), long short-term memory (LSTM), and general regression neural networks (GRNN). In Alswaidan et al. [1] work, three models were considered, gated recurrent unit followed by CuDNN concatenated with a CNN and a frequency-inverse document frequency (TF-IDF) to better label the text according to emotional categories.

2.2 Spatially and Sequential Architectures

In this study, the primary goals were to evaluate the two models of sentiment representation, namely the dimensional and the categorical models, and determine their applications and expected accuracy. Facebook posts were rated, firstly, considering the valence and arousal dimensions separately. In sum, 2895 messages were evaluated, and VA parameters were compared through the age and gender of the writer, with the authors concluding that female post-writers express more arousal and valence. Later, a two-linear regression model using a BoW representation, on 10-fold cross-validation with this data, reaches a high correlation to the annotated results, obtaining a Pearson correlation of 0.650 and 0.850 for valence and arousal, respectively. With the limited research on the use of sequential input data and the need for more emotionally rated data, [26] started a new investigation. Since CNN! (CNN!) are not ideal for processing sequences and RNN! (RNN!)s perform slowly, the authors induced four sequence-based convolution neural networks (SCNN).

3 Using Neural Word Embeddings for Extending Lexicons of Emotional Norms

We propose thirteen models based on models described in the related work section. All the studies were conducted in six different languages: English, Spanish, Portuguese, Italian, and Polish. In this section, we will start by describing the need for word embeddings, followed by an explanation of the models created in this study.

3.1 Word Embeddings

Word embeddings are vector representations for words, responsible for capturing their semantic or syntactic meaning. Our model used FastText word vectors pre-trained on Common Crawl and Wikipedia, which are available in 157 languages.

3.2 Models Exploring Statistics

In this study, we assign sentiment to words considering three dimensions, and we compared four different methods to predict the emotion of each word. One of the models that had a good performance was a simple multi-layer perceptron (MLP) [18], a set of neurons fully connected. The MLP model was built through Keras[1], an open-source library integrated on top of TensorFlow to allow building deep learning models. On the output layer, we have a Dense layer with three neurons, one for each emotional dimension we consider. This MLP was trained through 200 epochs, with a batch size of 64 and an Adam [11] optimizer. We adapted the MLP from our previous work and trained it with datasets affective normas for words from six different languages: English [4,21,24], Spanish [17], Portuguese [22], Italian [15], German[19], and Polish [10].

3.3 Models Exploring Machine Learning

Yann LeCun, inspired by the human visual cortex, discovered by Hubel and Weisel [9], developed the Convolution and Polling architecture [13], also known as Convolutional Neural Networks (CNN). LeCun applied this technique to images, and it was years later that CNN was applied to NLP. The main goal of CNN is to detect patterns across space, by firing when a determined pattern of words compared to a determined filter.

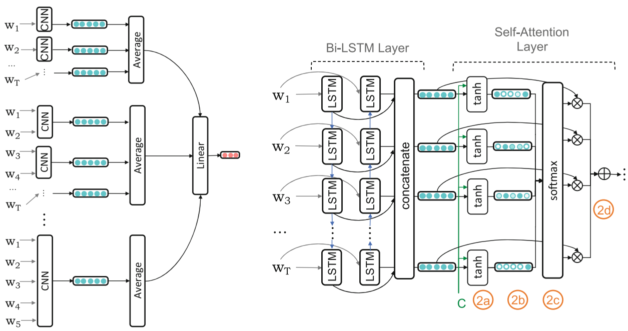

Fig. 1. Convolution and Polling operations applied to a sentence (Left) and BiLSTM and Attention (Right)

With the CNN model, we made four minor alterations. The first model is identical to the one shown in Fig. 1 on the left. In the next three, we applied the MLP model. First, passed the embeddings through the MLP model, and they were the input to the Convolution Layer. Second, we applied the MLP to the output of the Convolution Layer, and after applying the Average Pooling. In third place, we applied the MLP model in the end, after the linear operation.

[1] https://github.com/keras-team/keras.

Long Short-Term Memory

Even though CNN has fast performance, LSTM is more successful when working with natural language processing [25], such as sequences of words expressed as time series. In Eq. 1, we can observe all the operations that an LSTM cell requires.

$$\begin{aligned} s_t = R_{\text{ LSTM }}\left(s_{t-1}, x_t\right) = \left[c_t ; h_t\right] \\ c_t = f \odot c_{t-1} + i \odot z \\ h_t = o \odot \tanh\left(c_t\right) \\ i = \sigma\left(x_t W^{xi} + h_{t-1} W^{hi}\right)] \\ f = \sigma\left(x_t W^{xf} + h_{t-1} W^{hf}\right) \\ o = \sigma\left(x_t W^{xo} + h_{t-1} W^{ho}\right) g \\ = \tanh\left(x_t W^{xz} + h_{t-1} W^{hz}\right) \\ y_t = O_{\text{LSTM}}\left(s_t\right) = h_t \end{aligned} \tag{1}$$

To enhance the position of each word in the sentence [20], we choose to use a Bidirectional LSTM (BiLSTM). The idea is to have two LSTMs traveling through the sentence simultaneously, one that encodes the sentence left to right and, separately, another that travels from the end to the beginning of the sentence. In the end, we concatenate these two representations. This is translated into the BiLSTM Layer of the Fig. 2. However, as Yin et al. [25] referred in their paper, tracing the whole sentence with an LSTM can disregard the keywords. So, align with LSTM, we also used a Self-Attention Layer.

Attention

In our models, we used a Keras SeqSelfAttention layer with a sigmoid attention activation. This layer can be translated into the Self-Attention Layer from Fig. 2 and the Eq. 2.

$$h_i, j = \tanh\left(x_i^{\top} W_1 + x_j^{\top} W_x + b_i\right) \tag{2a}$$

$$e_{i,j} = \sigma\left(W_a h_{i,j} + ba\right) \tag{2b}$$

$$a_i = \operatorname{softmax}\left(e_i\right) \tag{2c}$$

$$\text{self}_{\text{a}}\text{ttention}_i = \sum_j a_{i,j} x_j \tag{2d}$$

In Self-Attention, it is first necessary to calculate $h_{i,j}$ (2a) by summing the values of the current position and the previous, all previously multiplied by a weight matrix. After multiplying the values by the alignment weights, we get the alignment scores (2b). On 2c, we apply softmax to the attention scores for the values to vary between 0 and 1 and determine the probability of each given word. At the end (2d), a_i, the amount of attention $j^t h$ should pay to i^{th} input, and summing all the results.

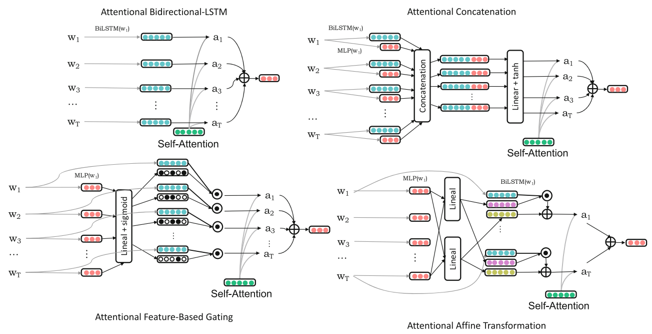

Fig. 2. Proposed models applying Self-Attention and BiLSTM Layers.

LSTM Models

First, we considered models with LSTM layer and a Self-Attention Layer, as shown in Fig. 2. We also analyzed an alteration to this model, instead of receiving the embeddings as the input, we applied the pre-trained MLP to all the embeddings and provided the operation results to the LSTM Layer.

We also produced three models inspired by the work developed by Margatina et all. [14]. These models were given the names they had in this paper.

Attentional Concatenation

The Attentional Concatenation model, Fig. 2 and Eq. 3, we calculate the BiLSTM of each embedding. In parallel, the MLP pre-trained model was applied for every word of the sentence. Then, we proceed to the concatenation of both operations and pass that concatenation through a Self-Attention Layer. In the end, calculate a Dense Layer with three dimensions to predict the three emotional dimensions.

$$x_1 = \tanh\left(W_c\left[\text{BiLSTM}\left(w_1\right) \| MLP\left(w_i\right)\right] + b_c\right)\right) \tag{3a}$$

$$\text{operations 2a -2d}$$

$$d = l \cdot 3 + b \tag{3b}$$

Attentional Feature-Based Gating

The second method, described in Fig. 2 and Eq. 4, we apply the MLP pre-trained model to the word embeddings and later use linear plus sigmoid operations. Appling the gating mechanism, by applying the sigmoid function, we will have a mask-vector where each value varies between 0 and 1 that will later be applied to the embeddings of each word by an element-wise multiplication, ⊙. Lastly, we used a Self-Attention Layer.

$$f_g\left(h_i, \text{MLP}\left(w_i\right)\right) = \sigma\left(W_g\, \text{MLP}\left(w_i\right) + b_g\right) \odot h \tag{4}$$

Attentional Affine Transformation
In the final model 2, the feature-wise affine transformation is applied; in other words, a normalization layer preserving collinearity and ratios of distances. Primarily, we apply the pretrained MLP model to the word embeddings, and enforce a scaling and shifting vector to the results of the MLP. This model, initially inspired by Perez et al. [16], allow to capture of dependencies between features by a simple multiplicative operation. The results of the linear operation γ over the MLP results are later multiplied element-wise with the results from the BiLSTM Layer over the embeddings. After, we add these values to β, and apply a Self-Attention Layer.

$$f_a\left(h_1, \mathrm{MLP}\left(w_i\right)\right) = \gamma\left(\mathrm{MLP}\left(w_i\right)\right) \odot h_i + \beta\left(\mathrm{MLP}\left(w_i\right)\right) \tag{5a}$$

$$\gamma(x) = W_\gamma x + b_\gamma \tag{5b}$$

$$\beta(x) = W_\beta x + b_\beta \tag{5c}$$

4 Experimental Evaluation

This section describes the experiments conducted to infer sentiment from textual utterances. In the third set of experiments, it was necessary to access the result of statistical models to predict the sentiment of textual utterances. One of the models that had a better performance was the MLP. An MLP was pre-trained with seven datasets in different languages. Figure 3 established a correlation between the dimensional distribution of the datasets (Table 1).

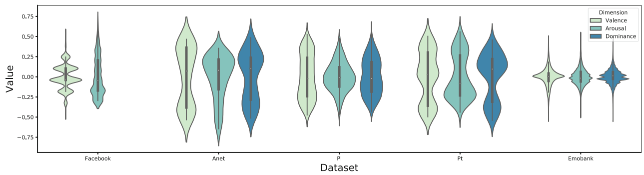

Fig. 3. Comparison of the dimensional distribution of the datasets in several languages.

Table 1. Results obtained for statistical sentiment prediction of textual utterances, in terms of Pearson's correlation coefficient and MAE.

		Pt		Pl		Emobank		ANET		Fb	
		Pearson	MAE	Pearson	MAE	Pearson	MAE	Pearson	MAE	Pearson	MAE
MLP Average	V	**0.686**	0.234	**0.499**	0.227	**0.359**	0.086	0.639	**0.301**	**0.384**	0.154
	A	0.511	0.216	0.222	**0.160**	**0.152**	0.101	**0.542**	0.319	**0.111**	0.234
	D	0.470	**0.238**	0.312	0.187	0.058	0.093	0.261	0.263	–	–
Average MLP	V	0.625	**0.232**	0.429	0.226	0.284	0.073	**0.697**	0.312	0.298	**0.132**
	A	0.342	0.218	0.109	0.187	0.122	0.089	0.433	0.355	0.790	0.237
	D	**0.579**	0.234	**0.436**	0.194	**0.123**	0.122	**0.622**	**0.258**	–	–
Pooling Average MLP	V	0.482	0.256	0.453	0.231	0.201	0.091	0.491	0.323	0.192	0.149
	A	0.187	0.231	0.166	**0.160**	0.110	0.122	0.420	**0.316**	0.79	0.250
	D	0.310	0.257	0.358	**0.183**	0.057	**0.092**	0.397	0.277	–	–
Pooling MLP Average	V	0.537	0.249	0.456	0.231	0.224	0.094	0.492	0.323	0.193	0.148
	A	0.266	0.230	0.168	**0.160**	0.098	0.130	0.420	**0.316**	0.82	0.244
	D	0.405	0.263	0.359	**0.183**	0.068	0.100	0.396	0.360	–	–
MLP Pooling Avg	V	0.339	0.317	0.402	**0.222**	0.083	**0.071**	0.605	0.312	0.137	0.161
	A	0.330	0.253	**0.335**	0.188	0.029	**0.088**	0.515	0.336	0.152	**0.208**
	D	0.219	0.342	0.256	**0.183**	0.039	0.182	0.327	0.268	–	–

The Affective Norms for English Words (Anew) comprised 1,034 unique words. The early work on sentiment analysis is annotated, considering the three dimensions of valence, arousal, and dominance. It was essential to consider other languages and provide richer data such as gender and education level.

4.1 Models Exploring Statistics

Despite the simplicity of the model Average (i.e., the MLP model is applied to each word of the text, and an average of all the outputs is calculated to deliver a final output), it was the model that showed a better performance of the word-level solutions in almost every dataset. All the models tested in these experiments are described in the section above.

4.2 Models Exploring Machine Learning

The set of experiments considering text-level sentiment prediction was conducted using cross-validation. This method allows for the validation of a model (e.g., by calculating its precision) by dividing a dataset into splits, usually between 2 and 5. A number of those splits are used to train the model, and the other is used to validate it. The table displays the results for each model through each dataset, considering Pearson's correlation, MAE and MSE.

The LSTM performs better with classification tasks than the CNN and the BiLSTM with regression tasks. It is possible to see a great improvement in more extensive datasets, such as Facebook. The dimension that was more difficult to tackle was arousal, especially in the Facebook dataset. This work shows better values for the dimension arousal than the work from other teams. The results were obtained using a model composed of Bi-LSTM+MP+Attention.

Table 2. The prediction of valence, arousal and dominance with several models. The training and testing data are textual utterances form datasets English, Polish and Portuguese.

		Pt			Pl			Emobank			ANET			Fb		
		Pearson	MAE	MSE	Pearson	MAE	MSE	Pearson	MAE	MSE	Pearson	MAE	MSE	Pearson	MAE	MSE
LSTM	V	0.641	0.184	0.059	**0.507**	**0.184**	0.055	**0.536**	0.070	0.009	**0.769**	0.207	0.059	0.547	0.100	0.018
	A	0.608	**0.164**	0.047	0.333	0.166	0.034	**0.333**	0.088	0.013	**0.617**	0.188	0.053	0.494	0.177	0.060
	D	0.576	**0.164**	0.056	0.445	0.149	0.042	0.092	0.120	0.065	0.439	0.231	0.082	–	–	–
MLP + LSTM	V	0.319	0.246	0.087	0.258	0.225	0.073	0.150	0.276	0.012	0.236	0.316	0.120	0.065	0.126	0.026
	A	0.232	0.241	0.071	0.108	0.146	0.034	0.016	0.297	0.013	0.254	0.282	0.097	0.126	0.235	0.081
	D	0.345	0.232	0.069	0.296	0.192	0.054	0.022	0.552	0.093	0.112	0.375	0.252	–	–	–
CNN	V	0.632	0.228	0.062	0.415	0.211	0.072	0.434	0.070	0.021	0.672	0.261	0.092	0.495	0.102	0.020
	A	0.312	0.241	0.050	0.241	0.148	0.059	0.170	.0.087	0.032	0.493	0.221	0.168	0.260	0.212	0.058
	D	0.427	0.234	0.063	0.247	0.235	0.109	0.040	0.258	0.075	0.261	0.329	0.091	–	–	–
MLP + CNN	V	0.584	0.236	0.076	0.397	0.212	0.067	0.466	**0.069**	0.009	0.657	0.249	0.087	0.501	0.109	0.019
	A	0.345	0.221	0.063	0.281	0.146	0.034	0.136	0.089	0.013	0.536	0.204	0.060	0.316	0.215	0.065
	D	0.419	0.227	0.078	0.282	0.202	0.067	0.040	0.251	0.085	0.167	0.410	0.302	–	–	–
CNN + MLP	V	0.552	0.223	0.066	0.395	0.215	0.066	0.449	0.071	0.009	0.523	**0.080**	0.066	0.485	0.107	0.019
	A	0.343	0.219	0.034	0.197	0.145	0.034	0.214	0.088	0.013	0.393	**0.114**	0.058	0.315	0.215	0.067
	D	0.342	0.227	0.061	0.243	**0.147**	0.061	0.066	0.182	0.076	0.408	0.291	0.149	–	–	–
Attention Concat	V	**0.691**	**0.177**	0.057	0.435	0.202	0.064	0.507	0.073	0.010	0.649	0.238	0.007	**0.561**	0.101	0.019
	A	**0.620**	0.165	0.046	0.297	0.144	0.035	0.302	0.089	0.014	0.481	0.209	0.051	**0.565**	0.176	0.052
	D	**0.663**	0.167	0.049	0.348	0.182	0.050	**0.363**	0.122	0.074	0.283	0.276	0.004	–	–	–
Attention Fracture Based	V	0.641	0.184	0.050	0.501	0.192	0.059	0.531	**0.069**	0.001	0.680	0.226	0.056	0.557	**0.098**	0.021
	A	0.608	0.164	0.042	**0.391**	**0.137**	0.031	0.320	**0.083**	0.014	0.538	0.198	0.051	0.545	**0.174**	0.057
	D	0.576	0.173	0.058	**0.470**	0.160	0.043	0.082	**0.116**	0.065	0.479	**0.217**	0.084	–	–	–
Attention Affine Transformation	V	0.569	0.206	0.072	0.434	0.206	0.065	0.501	0.074	0.010	0.728	0.225	0.070	0.523	0.108	0.057
	A	0.540	0.177	0.050	0.268	0.148	0.036	0.270	0.092	0.015	0.608	0.189	0.056	0.491	0.184	0.436
	D	0.473	0.218	0.075	0.338	0.180	0.051	0.075	0.143	0.067	**0.481**	0.266	0.119	–	–	–

Similar to my results using the Attention Feature Based model, which obtained results of 0.531 for Emobank, 0.557, and 0.545 for Facebook. Comparing the results and considering that my model performed lower, but being trained with several idioms, the lower performance can be justified. Possible applications of this model can be used in multilingual context such as tweets [2,6] (Table 2).

5 Conclusions and Future Work

This research provides three trained models and one word-level model that show promising results compared to the state-of-the-art. An MLP was pre-trained with lexicons from six different languages. Four models that do take into consideration the syntactic structure and do require training were created. LSTMs tend to perform slightly better than CNN models, and this difference was more evident in the arousal dimension.

This study provides, as theoretical implications, a comparison between statistical models and machine learning models. Possible practical applications to the findings in this study could be to monitor informal political online discussions and to lead to a better understanding of hate speech on social media. It could also be interesting to experiment with word embeddings trained on different types of corpora for future work.

Acknowledgement. We gratefully acknowledge financial support from FCT - Fundação para a Ciência e a Tecnologia (Portugal), national funding through research grant UIDB/04152/2020. This work is also supported by national funds through PhD grant (UI/BD/153587/2022) supported by FCT.

References

1. Alswaidan, N., Menai, M.E.B.: KSU at SemEval-2019 Task 3: hybrid features for emotion recognition in textual conversation. In: Proceedings of the 13th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2019, Minneapolis, MN, USA, 6–7 June 2019, pp. 247–250 (2019)
2. Aparicio, J.T., de Sequeira, J.S., Costa, C.J.: Emotion analysis of portuguese political parties communication over the COVID-19 pandemic. In: 2021 16th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6. IEEE (2021)
3. Binali, H., Wu, C., Potdar, V.: Computational approaches for emotion detection in text. In: 4th IEEE International Conference on Digital Ecosystems and Technologies, pp. 172–177. IEEE (2010)
4. Bradley, M.M., Lang, P.J.: Affective norms for English words (ANEW): instruction manual and affective ratings. Technical report, Technical report C-1, the center for research in psychophysiology ... (1999)
5. Buechel, S., Hahn, U.: Word emotion induction for multiple languages as a deep multi-task learning problem. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, 1–6 June 2018, Volume 1 (Long Papers), pp. 1907–1918 (2018)
6. Costa, C., Aparicio, M., Aparicio, J.: Sentiment analysis of Portuguese political parties communication. In: The 39th ACM International Conference on Design of Communication, pp. 63–69 (2021)
7. Ekman, P.: An argument for basic emotions. Cogn. Emot. **6**, 169–200 (1992)
8. Ekman, P., Friesen, W.V.: Constants across cultures in the face and emotion. J. Pers. Soc. Psychol. **17**, 124 (1971)
9. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. **160**, 106–154 (1962)
10. Imbir, K.K.: Affective norms for 4900 polish words reload (ANPW_R): assessments for valence, arousal, dominance, origin, significance, concreteness, imageability and age of acquisition. Front. Psychol. **7**, 1081 (2016)
11. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
12. LaBrie, R.C., Louis, R.D.S.: Information retrieval from knowledge management systems: using knowledge hierarchies to overcome keyword limitations. In: 9th Americas Conference on Information Systems, AMCIS 2003, Tampa, FL, USA, 4–6 August 2003, p. 333 (2003)
13. LeCun, Y., Bengio, Y., et al.: Convolutional networks for images, speech, and time series. Handb. Brain Theory Neural Netw. **3361**, 1995 (1995)
14. Margatina, K., Baziotis, C., Potamianos, A.: Attention-based conditioning methods for external knowledge integration. arXiv preprint arXiv:1906.03674 (2019)
15. Montefinese, M., Ambrosini, E., Fairfield, B., Mammarella, N.: The adaptation of the affective norms for English words (ANEW) for Italian. Behav. Res. Methods **46**, 887–903 (2014)

16. Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A.: FiLM: visual reasoning with a general conditioning layer. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)
17. Redondo, J., Fraga, I., Padrón, I., Comesaña, M.: The Spanish adaptation of anew (affective norms for English words). Behav. Res. Methods **39**, 600–605 (2007)
18. Ruck, D.W., Rogers, S.K., Kabrisky, M., Oxley, M.E., Suter, B.W.: The multilayer perceptron as an approximation to a bayes optimal discriminant function. IEEE Trans. Neural Netw. **1**, 296–298 (1990)
19. Schmidtke, D.S., Schröder, T., Jacobs, A.M., Conrad, M.: ANGST: affective norms for German sentiment terms, derived from the affective norms for English words. Behav. Res. Methods **46**, 1108–1118 (2014)
20. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. IEEE Trans. Signal Process. **45**, 2673–2681 (1997)
21. Scott, G.G., Keitel, A., Becirspahic, M., Yao, B., Sereno, S.C.: The glasgow norms: ratings of 5,500 words on nine scales. Behav. Res. Methods **51**, 1258–1270 (2019)
22. Soares, A.P., Comesaña, M., Pinheiro, A.P., Simões, A., Frade, C.S.: The adaptation of the affective norms for English words (ANEW) for European Portuguese. Behav. Res. Methods **44**, 256–269 (2012)
23. Wang, J., Yu, L.-C., Lai, K.R., Zhang, X.: Community-based weighted graph model for valence-arousal prediction of affective words. IEEE/ACM Trans. Audio Speech Lang. Process. **24**, 1957–1968 (2016)
24. Warriner, A.B., Kuperman, V., Brysbaert, M.: Norms of valence, arousal, and dominance for 13,915 English lemmas. Behav. Res. Methods **45**, 1191–1207 (2013)
25. Yin, W., Kann, K., Yu, M., Schütze, H.: Comparative study of CNN and RNN for natural language processing. arXiv preprint arXiv:1702.01923 (2017)
26. Zahiri, S.M., Choi, J.D.: Emotion detection on TV show transcripts with sequence-based convolutional neural networks. arXiv preprint arXiv:1708.04299 (2017)