Dilated LSTM with ranked units for Classification of Suicide Note

UNIVERSITY OF HULL

BIG DATA ANALYTICS

A.M. Schoene, A.P. Turner and N.D. Dethlefs
A.M.Schoene@2017.hull.ac.uk | A.Turner@hull.ac.uk | N.Dethelfs@hull.ac.uk

Abstract

Recent statistics in suicide prevention show that people are increasingly posting their last words online and with the unprecedented availability of textual data from social media platforms researchers have the opportunity to analyse such data. Furthermore, psychological studies have shown that our state of mind can manifest itself in the linguistic features we use to communicate. In this paper, we investigate whether it is possible to automatically identify suicide notes from other types of social media blogs in a document-level classification task. We present a learning model for modelling long sequences, achieving an f1-score of 0.84 over the baselines of 0.53 and 0.80 (best competing model). We also show through visualisations which features the learning model identifies.

Introduction

- The use of social media platforms has become part of everyday life and there is increasing evidence that social media can influence both suicide-related behaviour and other mental health conditions.
- Whilst there are efforts to tackle suicide and other mental health conditions online by social media platforms, there are still concerns that there is not enough support and protection, especially for younger users.
- This has led to a notable increase in research of suicidal and depressed language usage [9], triggering the development of new healthcare applications that aid detection of concerning posts on social media platforms [8].

References

[1] Pirina et al. Identifying depression on reddit: The effect of training data. In *Proceedings of the 2018 EMNLP Workshop SMM4H: The 3rd Social Media Mining for Health Applications Workshop*

[2] Schler et al. Effects of age and gender on blogging. In *2006 AAAI spring symposium: Computational approaches to analyzing weblogs*

[3] Chang et al. Dilated recurrent neural networks. In *NIPS '17*

[4] Schoene et al. Automatic Identification of suicide notes from linguistic and sentiment features. In *Proceedings of the 10th SIGHUM Workshop*

[5] Chang et al. Dilated recurrent neural networks. In *NIPS '17*

[6] Hochreiter et al. Long short-term memory. In *Neural Computation 1997*

[7] Yang et al. Hierachical attention networks for document classification. In *2016 Proceedings of NAACL*

[8] Calvo et al. Natural Language Processing in mental health applications using non-clinical texts. In *Natural Language Engineering*

[9] Pestian et al. Sentiment Analysis of suicide notes: A shared task. In *Biomedical Informatics insights*

Learning Model

Dilated LSTM with ranked units Each document D contains i sentences S_i, where w_i represents the words in each sentence. Firstly, we embed the words to vectors through an embedding matrix W_e, which is then used as input into the dilated LSTM. The most important part of the dilated LSTM is the dilated recurrent skip connection, where $LSTM_t^{(l)}$ is the cell in layer l at time t:

$$LSTM_t^{(l)} = f(x_t^{(l)}, c_{t-s^{l-1}}^{(l)}). \quad (1)$$

$s^{(l)}$ is the skip length; or dilation of layer l; $x_t^{(l)}$ as the input to layer l at time t; M and L denote dilations at different layers: $s^{(l)} = M^{(l-1)}, l = 1, \dots L$.

The dilated LSTM alleviates the problem of learning long sequences, but not each document has the same sequence length, so in order to overcome this variability we provide fixed boundaries to each layer by reducing the number of hidden units per sub-LSTM hierarchically. Therefore larger sub-LSTMs focus on learning long-term dependencies, whilst smaller sub-LSTMs focus on more frequently occurring short-term dependencies.

Attention Layer We extended the earlier implementation with an attention mechanism ATTENTION, using attention to find words that are most important to the meaning of a sentence at document level. We use the output of the dilated LSTM as direct input into the attention layer, where O denotes the output of final layer L of the Dilated LSTM at time t_{+1}.

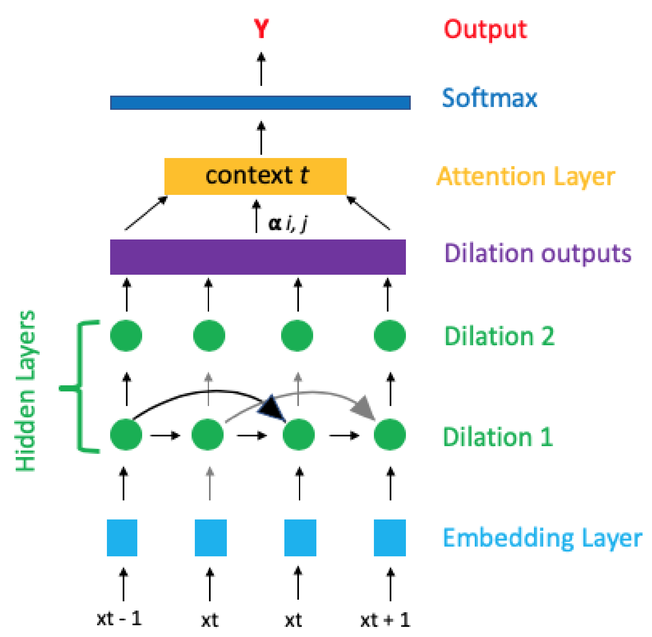

Figure 1: A 2-layer dilated LSTM with ranked units and attention layer.

Experiments and Results

Model	Aver. Precision	Aver. Recall	Aver. F1-score
Maximum Entropy	0.763	0.60	0.53
LSTM	0.50	0.46	0.30
BiLSTM	0.83	0.76	0.74
BiLSTM Attention	0.84	0.81	0.80
DLSTM Attention	0.82	0.81	0.81
DLSTM ranked units	**0.89**	**0.84**	**0.84**

Table 1: F1-score of different learning models in %

Data and Visualisation

my dearest family am terribley sick and it is all my fault blame no one but myself know it is going to hard with william and sister please see that charles gets a mickey mouse watch for his birthday jane am counting on you to take care of mother please do not follow in my footsteps elinor my darling know you did everything possible to avoid this but please forgive me as think it was the only way out god forgive me and help take care of my family

Figure 1: Attention weights (Suicide Note)

Data We used 211 'neutral' posts [2], 211 suicide notes [4, 5] and 211 depressed notes [1] for our experiments.

Conclusion

We have introduced the Dilated LSTM with ranked units and shown that the learning model is able to successfully distinguish suicide notes from other types of corpora. Therefore demonstrating that accurate classification is possible solely on linguistic patterns in this type of textual data. However, additional research is needed to understand if, for example, these language patterns generalise over larger datasets and which role emotions expressed in textual data could help further to identify suicidal ideation. Given further research is conducted such work could be useful in a number of scenarios, including but not limited to assessing the seriousness of a suicide attempt in a clinical setting, distinguishing forged from genuine notes or help suicide prevention charities in flagging up emails that indicate high risk of suicide.