Deep learning

Author1, Author2, Author3
Affiliation1, Affiliation2, Affiliation3
Address1, Address2, Address3
author1@xxx.yy, author2@zzz.edu, author3@hhh.com
{author1, author5, author9}@abc.org

Abstract

We are interested in text classification. Rule-based systems are more expensive than machine learning. We hypothesize that a neural network with long short-term memory and attention, M&S, H&M, and RCMP could possibly be used to classify text. We report the state-of-the art results which are significantly higher than the baseline.

Keywords: keyword1, keyword2, keyword3

1. Introduction

We are interested in text classification. Text classification was approached by us by deep learning (Putin et al., 2004) because deep learning has had good results (Orban and Trump, 2019). But, text classification has not been done by deep learning with convolutional neural networks, recurrent neural networks, long short-term memory, M&S, H&M, and RCMP, so we used it. We report the best performance on this task.

1.1. Related Work

(Mandela, 1990) used hidden Markov Models. (Thunberg, 2019) used support vector machines. (Weil, 1975) used naive Bayes. They do not have good performance because we have state-of-the art performance now.

2. Materials and Methods

Texts were collected by hiding microphones in support group meetings for mentally ill children with Acquired Immune Deficiency Syndrome. They were then labelled with identifying information and annotated by crowd-sourcing on Amazon Mechanical Turk, which is very recommended because it is very inexpensive (Fort et al., 2011). No data has ever been annotated so inexpensively. Annotations were then checked by the first author and if the first author disagreed with the annotators, the annotation was corrected (Chomsky, 2009).

A sample of positive instances was then selected. We tokenized (Trieschnigg et al., 2007), part-of-speech tagged (Brill, 1992), structurally parsed (Bikel, 2004), dependency parsed (Trump, 2018), enhanced with synonyms from WordNet (Miller, 1995; Fellbaum, 2010), PropBank (Palmer et al., 2005), and Wikipedia (Trump, 2019), named entity recognized (Nouvel et al., 2015), semantically role labeled (Gildea and Jurafsky, 2002), argumentation labeled, coreferentially resolutioned (Hobbs, 1978), sentiment-tagged (**?**), segmented (Jurafsky and Martin, 2009), and pre-classified with an oracle, and then word embeddings were calculated by word2vec (Jurafsky and Martin, 2009) and GLOVE (Jurafsky and Martin, 2009). Then a convolutional neural network (Jurafsky and Martin, 2009) and a recurrent neural network with bidirectional long short-term memory (Jurafsky and Martin, 2009) and attention () was used. That means that there was no manual feature engineering. The code used in our experiments is freely available at `https://github.com/manishanker/cnn_text_classification`.

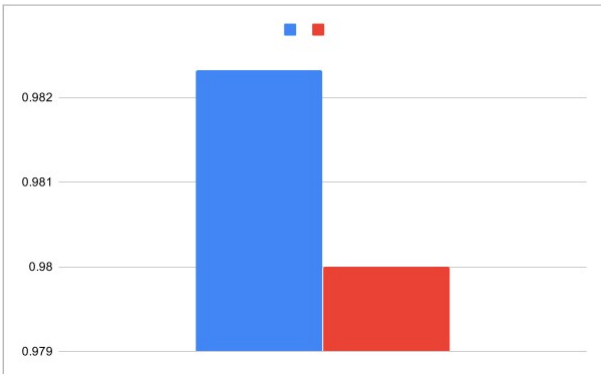

Figure 1: Our system significantly outperforms the baseline system.

3. Results

The baseline system only achieved an accuracy of 0.50. In contrast, our system achieved an accuracy of 0.9823211104 with manual correction of some instances that the system should have gotten right.

System	**Accuracy**
Random assignment	0.50
CNN	0.9823211103
CNN + RNN	0.9813211102
CNN + RNN + LSTM	0.9803211100
CNN + RNN + LSTM + M&S	0.9821211102
CNN + RNN + LSTM + M&S + H&M	0.9823201104
CNN + RNN + LSTM + M&S + H&M + RCMP	**0.9823211104**

Table 1: The bolded number is on the bottom line.

4. Discussion and Conclusions

Research in the developed world suggests that parental alcohol use negatively impacts child mental health. However, little research has examined these relations among children

in the developing world and no studies to date have done so in the context of AIDS-orphanhood. These results suggest that orphan status and caregiver alcohol use may independently relate to mental health problems in children and that the effects of both should be considered in the context of the mental health needs of children in AIDS-affected countries.

4.1. Future Work

We will annotate more data in order to improve peformance.

4.2. Conclusions

In conclusion, our contribution of state-of-the art performance is the best reported performance. The novelty is that we used a neural network with long short-term memory, M&S, H&M, and RCMP. This proves that deep learning is the best way to natural language process (see also our earlier work (Erdogan and Trump, 2019c; Erdogan and Trump, 2019b; Erdogan and Trump, 2019a)). There is no other way to interpret the scores, and we hope that everyone will use deep learning. (Putin et al., 2004) has observed very forcefully that deep learning is the best way to natural language process. We argue that it is.

5. Acknowledgements

The authors gratefully deny having received funding from the CHEKA, NKVD, KGB, GRU, and Moscow Judo Academy, but if we had received funding from them, which we didn't, it would be fine.

6. Bibliographical References

Bikel, D. M. (2004). Intricacies of collins' parsing model. *Computational Linguistics*, 30(4):479–511.

Brill, E. (1992). A simple rule-based part of speech tagger. In *Proceedings of the third conference on Applied natural language processing*, pages 152–155. Association for Computational Linguistics.

Chomsky, N. (2009). *Cartesian linguistics: A chapter in the history of rationalist thought*. Cambridge University Press.

Erdogan, R. T. and Trump, D. J. (2019a). Tentative preliminary proposal for processing of old low norse epic poetry from the lower southwestern shore of lake faenmeg with recurrent convolutional long short-term memory neural networks and attention. In *Skažite emu, čto on očn stabilnyj genij, i ètot idiot peredast vam kljuçi ot svoej demokratii*.

Erdogan, R. T. and Trump, D. J. (2019b). Tentative preliminary proposal for processing of old low norse epic poetry from the lower southwestern shore of lake faenmeg with recurrent convolutional long short-term memory neural networks. In *Počemu my okazali davlenie na Kommunističeskuju partiju Ispanii, čtoby podavit' profsojuzy, socialistov i anarhistov vo vremja graždanskoj vojny v Ispanii*.

Erdogan, R. and Trump, D. J. (2019c). Tentative preliminary proposal for processing of old low norse epic poetry from the lower southwestern shore of lake faenmeg with recurrent convolutional neural networks. In *Očen' horošaja konferencija po osetrovoj ikre i èkskursija po moej dače*.

Fellbaum, C. (2010). Wordnet. In *Theory and applications of ontology: computer applications*, pages 231–243. Springer.

Fort, K., Adda, G., and Cohen, K. B. (2011). Amazon mechanical turk: Gold mine or coal mine? *Computational Linguistics*, 37(2):413–420.

Gildea, D. and Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational linguistics*, 28(3):245–288.

Hobbs, J. R. (1978). Resolving pronoun references. *Lingua*, 44(4):311–338.

Jurafsky, D. and Martin, J. (2009). Speech and language processing.

Mandela, N. (1990). Evaluating theories of syntactic structure via a large sample of unrelated languages. In *Association for Computational Linguistics*.

Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Nouvel, D., Ehrmann, M., and Rosset, S. (2015). *Les entités nommées pour le traitement automatique des langues*. ISTE Group.

Orban, V. and Trump, D. J. (2019). Tentative preliminary proposal for processing of old low norse epic poetry from the lower southwestern shore of lake faenmeg with recurrent neural networks. In *Stoletnjaja meždunarodnaja konferencija po pisêvaritel'nym preimuŝestvam borŝa*.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.

Putin, V. V., Shestakov, V., and Levitsky, A. (2004). *Judo: History, theory, practice*. North Atlantic Books.

Thunberg, G. (2019). Visual presentation of quantitative data on global warming. In *Proceedings of the United Nations*.

Trieschnigg, D., Kraaij, W., and de Jong, F. (2007). The influence of basic tokenization on biomedical document retrieval. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 803–804. ACM.

Trump, D. J. (2018). Opinions about natural processing of language. In *arXiv*.

Trump, D. J. (2019). Deep learning of old low norse epic poetry from the southwestern shore of lake faenmeg. In *Very International Colloquium On Freedom of the Very Mean and Unfair Made-up Failing Press*.

Weil, S. (1975). Agency and universalism. In *Éditions du seuil*.