

Paper: *Convolutional Neural Networks for Sentence Classification*

Razwan Ahmed Tanvir

Date: 03/25/2022

Quote “... Our work is philosophically similar to Razavian et al. (2014) which showed that for image classification, feature extractors obtained from a pretrained deep learning model perform well on a variety of tasks—including tasks that are very different from the original task for which the feature extractors were trained. ”

Overview In this paper the author experimented with Convolutional Neural Networks (CNN) in various ways using the pre-trained word embedding vectors which is word-to-vec, to achieve sentence classification like sentiment analysis, question classification etc. The author mentioned that they used some hyper-parameter tuning on a basic Convolutional Neural Network and claimed that they produced state-of-the-art benchmarks in several tasks. Moreover, they proposed a simple modification to the model architecture so that they can use task-specific vectors and static vectors in their model. In this work, the authors trained a basic CNN which has one layer of convolution and word-vectors obtained from a pre-trained model. The author described that their model represents the sentences with two channels, static and non-static channels. After that these representations are filtered with feature maps. The model also incorporates max-over-time pooling and dropout along with softmax in fully connected layer. By these filters, the features were extracted and the final output from the model is a probability distribution over the labels. Authors claimed that the dropout as a regularization method works pretty well to regularize the proposed model. The usage of this dropout prevents the co-adaptation of hidden units. This is achieved by setting a hidden unit to zero at the time of forward-backpropagation.

The authors compared the results of their model in several benchmark tasks. While some of the tasks did not do very well but some performance gain was observed in the proposed model. The usage of the pre-trained vectors significantly enhanced the model performance. The proposed model performs better than several complex models which implements complex pooling mechanisms.

Intellectual Merit Author of this paper showed that, even a simple Convolutional Neural Network with a little bit of hyper-parameter tuning can outperform complex neural models. In this research the author experimented with several variants of the model and produced better results. The author explained each part of the model clearly and mentioned the model architectures and parameter values clearly. They measured their results with the current benchmarks for each of the tasks.

Broader Impact This research delved into the idea that a simple but effective model can be produced using some modifications on simple CNN and some hyper-parameter tuning. The proposed multi-channel and single channel models perform well to some extent. The author has a good command over NLP and he is a professor of MIT and he did his PhD from Harvard University.

Keywords Natural Language Processing, CNN, Hyper-parameter, Word Vectors

- Discussion • The authors did not discuss broadly about how they selected the specific values for
Questions their hyper-parameters.
- Though this research produced mentionable output, however, more discussion is need
to explain why some word-vectors provide better results.

Table 1: Grade deductions by section

Overview	Intellectual M.	B. Impact	Keywords	Questions	Is Online?