

NLP from Scratch

Alireza Moradi

Most of the state-of-the-art solutions to NLP problems are task-specific and need lots of prior knowledge, So many of the features in these approaches are manually developed and optimized for the specific tasks. The authors of this paper are suggesting a neural network architecture and learning algorithm which is capable of performing close to state-of-the-art task-specific solutions on multiple NLP problems mentioned below:

- Part of Speech Tagging (POS)
- Chunking
- Named Entity Recognition (NER)
- Semantic Role Labelling (SRL)

This approach needs very little linguistic knowledge and we let the network to take the input and learn different layers of feature extraction. The network is a multi-layer neural network (like a series of functions) which can model complex features by adding the non-linear layers (sigmoid, tanh, hardtanh etc.). The networks first layer is a lookup table which maps words into feature vectors (like word embeddings) which improves the performance. There are 2 versions of the model:

- **Window based approach:** which tries to tag a word W based on a window T of words where W is in the center. This approach won't work on SRL because it needs to look at the whole sentence to do the tasks.
- **Sentence based approach:** used in order to handle the problem with SRL mentioned above. Note that this approach doesn't have

any performance/accuracy benefits over window based approach on the other 3 tasks.

In the window based approach, after concatenating the feature vectors from the output of the initial layer they are passed to a linear layer followed by a [non-linear] hard tanh layer. In the final layer which is dependent on the kind of task there are as many outputs as there are tags for the task .

In the sentence based approach, a convolutional layer is used as the second layer to extract features of the sentence and because the word of interest is not in the center, max-pooling is used to select the maximum value in each row. The final 3 layers are the same as window based approach.