

I finally describe you second (but not last!) home task. You have to develop two variants of named entity recognition system for first track of the FactRuEval-2016 competition. This competition is described in the paper <http://www.dialog-21.ru/media/3430/starostinaetal.pdf>. Of course, you can solve tasks for second and third tracks too, but it is not obligatory :blush:

First part of this home task is devoted to do experiments with classical neural architecture of named entity recognizer (NER): word embeddings (Word2Vec or FastText) + Bidirectional LSTM + Conditional Random Fields (CRF) or time-distributed dense layer. You can see example of such NER in the paper <https://arxiv.org/abs/1508.01991>

Second part of this home task is concerned in using of more powerful contextual ELMo embeddings instead of classical word embeddings. Conception of ELMo is proposed in the paper <https://arxiv.org/pdf/1802.05365.pdf>. As you know, ELMo is based on the LSTM language model. You can use TF-Hub ELMo model for English: <https://tfhub.dev/google/elmo/2> But in our hometask you can process texts in Russian, and pre-trained Russian ELMo model is available by this link: http://docs.deeppavlov.ai/en/latest/apiref/models/embedders.html#deeppavlov.models.embedders.elmo_embedder.ELMoEmbedder This model is part of the DeepPavlov library <http://docs.deeppavlov.ai>. You can use this model as frozen feature extractor before the BiLSTM-CRF for final classification, but also you may attempt to finetune this model on large unlabeled Tweet corpus to increase efficiency of language modeling: <http://docs.deeppavlov.ai/en/latest/apiref/models/elmo.html>

Linear-chain CRF is special machine learning model, which can be last layer for neural network. CRF is adapted to classify some objects in context, i.e. in objects sequence (in our hometask you also need to classify objects in their context, where objects are words and context is whole sentence). CRF layer is available in the extension repository for Keras: <https://github.com/keras-team/keras-contrib> Also, you can use CRF as part of the Tensorflow library https://www.tensorflow.org/api_docs/python/tf/contrib/crf/ (for example see <https://guillaumequent.github.io/sequence-tagging-with-tensorflow.html>)

You have to use <https://github.com/dialogue-evaluation/factRuEval-2016/tree/master/devset> as training data and <https://github.com/dialogue-evaluation/factRuEval-2016/tree/master/testset> as data for final testing. You have to use special F1-score for evaluation on test set (see part 6.1 of the paper <http://www.dialog-21.ru/media/3430/starostinaetal.pdf>). You can use special script `t1_eval.py` (see <https://github.com/dialogue-evaluation/factRuEval-2016/tree/master/scripts>) to not write own evaluation function. Example of the FactRuEval data parsing is in https://github.com/bond005/deep_ner/blob/c18dde55e9fa2654f26feb84b6cf1cfbb213bf9/deep_ner/utils.py#L416 or in https://github.com/bond005/factRuEval-2016/blob/3a1b4540b1025fa73118d0e065c526437b37df12/elmo_lstm/create_submit.py#L196

I received F1-score > 89% in my experiments. I hope that you can receive equivalent results, or at least your F1-score will be greater than 80%.