

The model metrics for mountain NER are similar to the BERT metrics for all categories of named entities on supervised [Few-NERD](#) dataset. So it seems that by changing the training parameters we can slightly improve the model performance, but a more significant performance improvement is only available by changing the base model from DistilBERT to another language model, such as RoBERTa or XLNet.