

RoBERTa training: cross-entropy loss (avg. over 5 folds)

