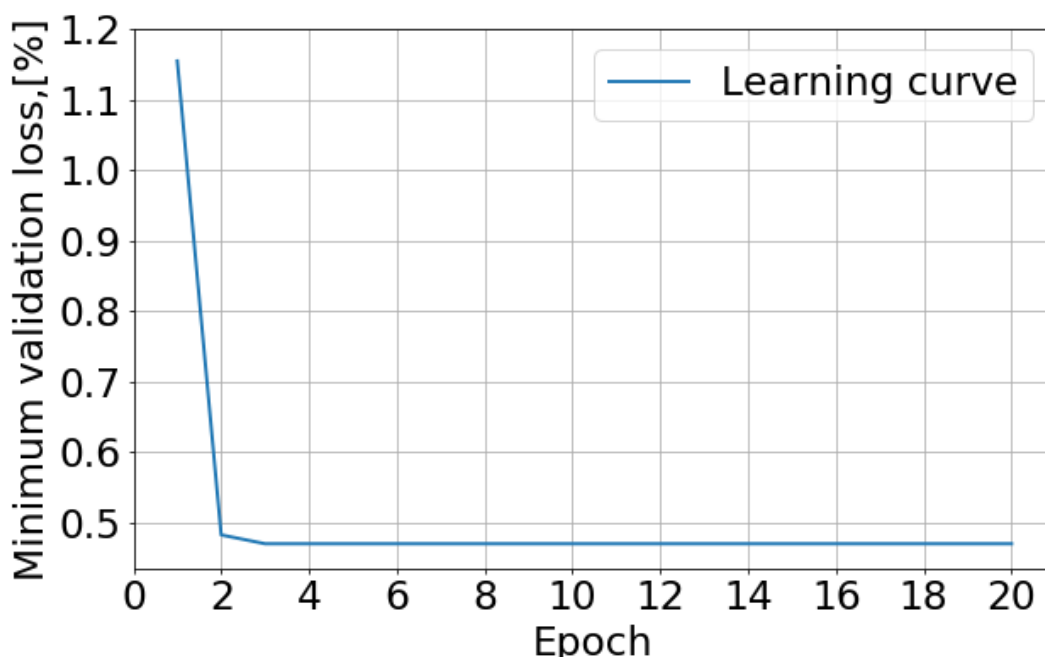
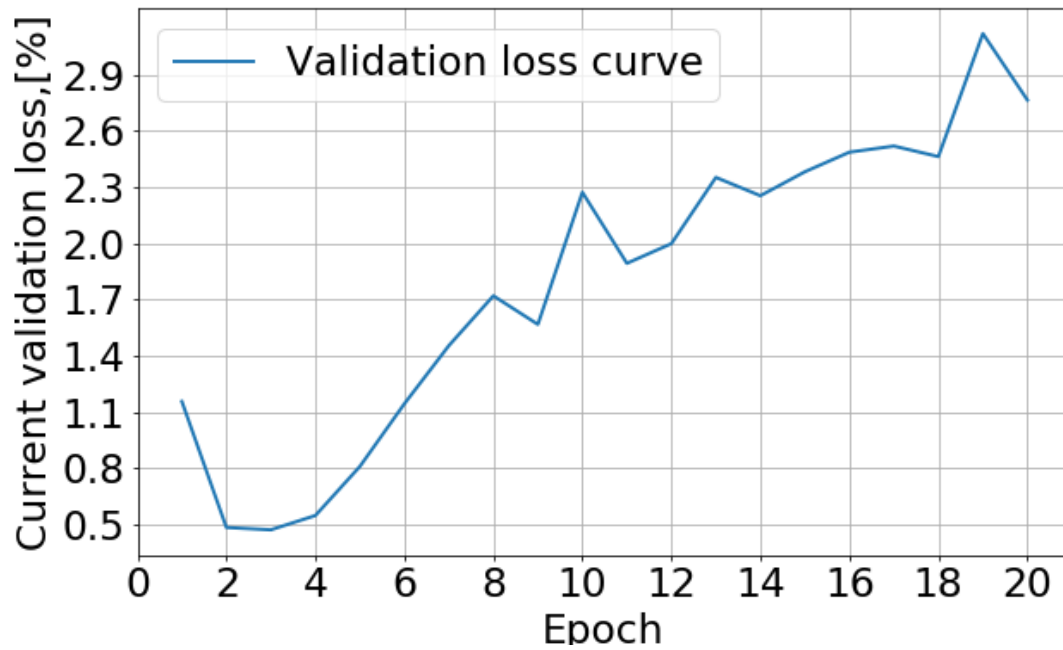


# Отчет о запуске базового алгоритма

Согласно заданию требовалось запустить базовый алгоритм решения задачи. В качестве модели использовалась модель аннотирования текстов SummaRuNNer. Модель реализовано с помощью библиотеки PyTorch. В качестве данных для обучения и валидации используются сокращенные датасеты DailyMail, где объем документов урезан до 1000. Произведено обучение в 20 epoch, batch\_size=10, построена кривая обучения и кривая, показывающая текущий loss:





С использованием полученных оценок предложений построены сокращенные тексты. Позже результат был оценен с помощью метрики ROUGE:

ROUGE-1	ROUGE-2	ROUGE-L
0.23263 - 0.25713	0.09697 - 0.11978	0.12410 - 0.14612

Так же была обучена модель перевода OpenNMT, для обучения использовался корпус слов Татоэба (<http://opus.nlpl.eu/download.php?f=Tatoeba/en-ru.txt.zip>), batch\_size=10, step\_counter=30000, обучающая выборка была сокращена до 1000 предложений, построены графики зависимости точности, кроссэнтропии и perplexity от количества epoch:

