

Лабораторная работа №4 Генерация текста

Слаутин Александр, Семкин Павел, Куклин Максим,
Кислицына Анастасия

Архитектура модели

Вход RNN-модели: one-hot вектора

RNN-блоки: LSTM или GRU

Выход модели: выход полносвязного слоя

Используемые гиперпараметры

- hidden size = 256,
- learning rate = 0.001,
- use GRU = False,
- temperature = 1,
- dropout probability = 0.5

Генерация summary статей с arxiv.org

Inference

Num_layers = 2

Then, the problem is adaptive to the network obtained for the complexity, confirption, algorithms and analyzes a mechanism to several statistical tricle systems, which analysis two dingiling communication and statistical external types in the special roduce specular in pixel, in which infrastructures of infrastructural descriptory demand are ensorted by protrions to predict interpretations, which combines detansordistic-loss-constitution in publicly active drawings consisting of product functions. The approach are appropriated as a method to cost-pointly computationally particular comprehensive collection-of-composes. We demograty an asymmetric domain with single modeling recognition tasks. The research in the permotances can be efficiently potentially factorization, remain inconditioned is are regressed as this

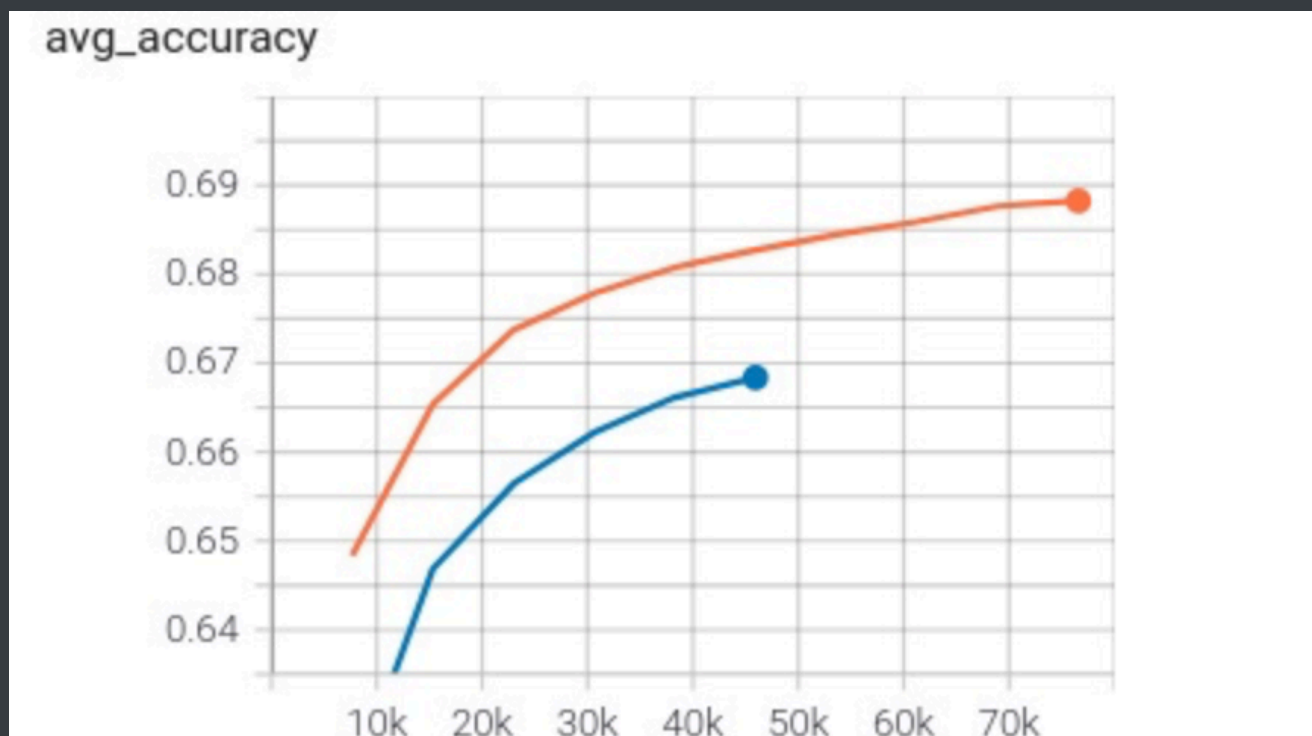
modified presence of machine translation, and to investigate point analysis which are exchalled by the added current advantage of its scale circuit in this static. Our conversity a

Num_layers = 4

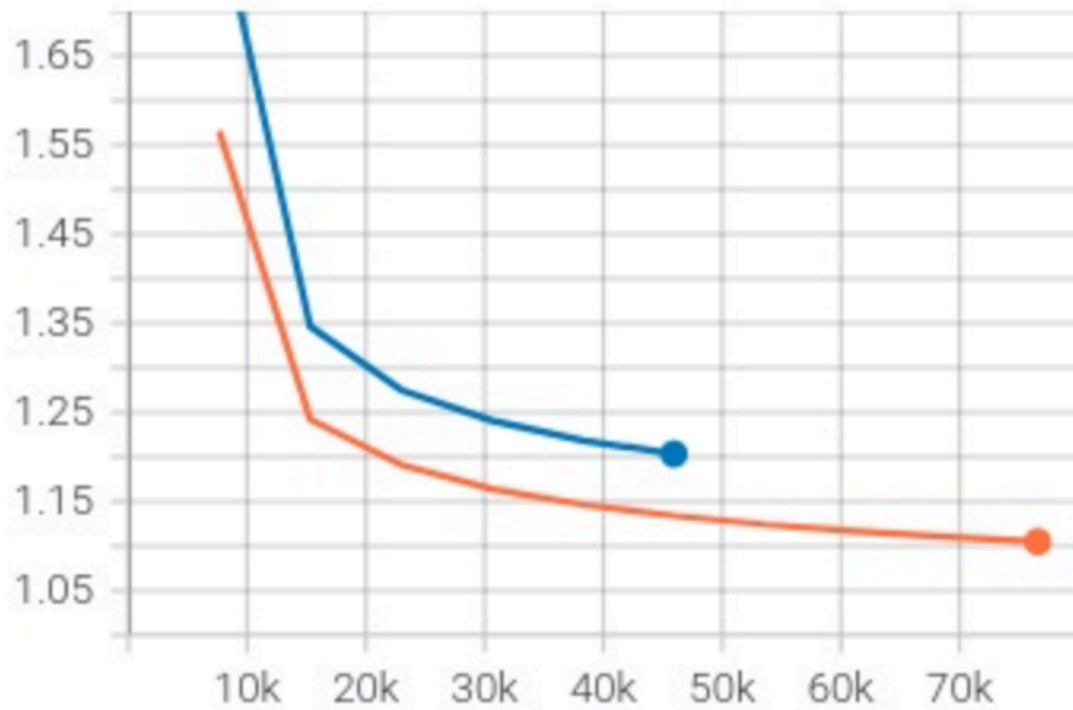
The and a new cases about the time access semantics a serser need in a proofs of performance in the input single-distribution of notions to deep novel audia when the intersection of perturbation behavior of the pre-patients with the proposed divergence of characteristics and problems to a times and model-special increasing moral-action as a single computain problem is true optimization traffic standard safe practical sampling scores it is proposed that implementing the man extension for a similar cet-ficpersing analysis, we study the interpretation of these statistical dadasnalisation predictings anson to detect with sensing phity analysis, minimum models in this domain as a trivial program with the definition to depends on component of similarity search on an important data show that the nonlinear subsequent deleb or the mobile, is the system for many single systems on phphase taskswip moanteriant as fundrised in a dynamical solution of the subseriofic robots. The reandial algorithm is inf

Визуализация тренировки

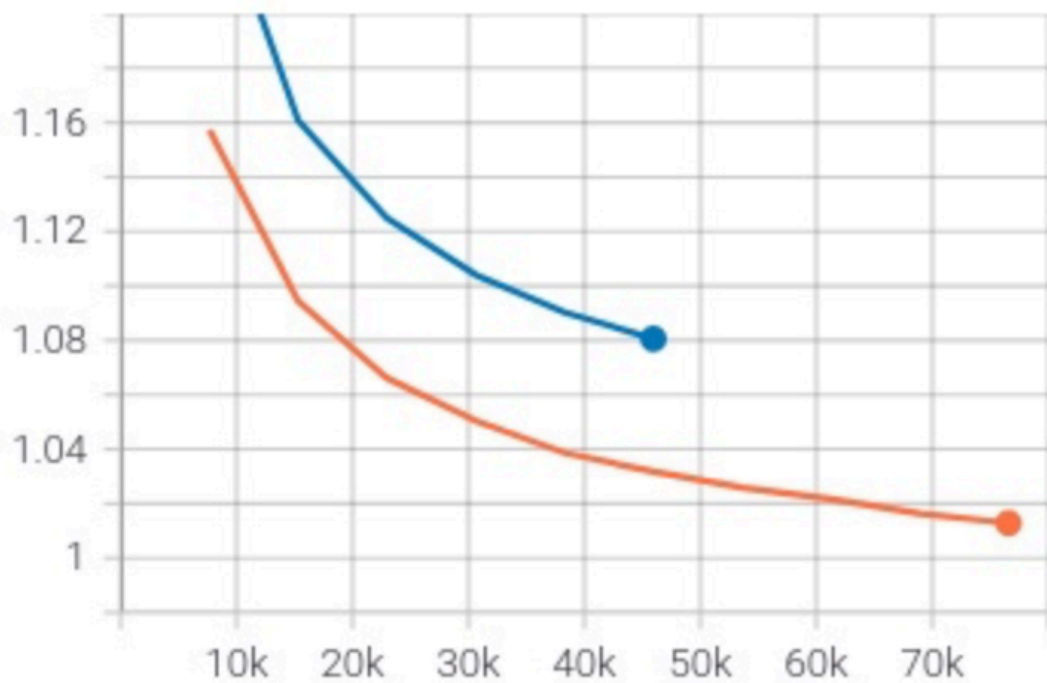
Num_layers = 2 (оранжевый), num_layers = 4 (синий)



avg_loss



avg_val_loss



Вывод

Увеличение количества RNN-блоков влечет за собой ухудшение качества.

Генерации русских новостей

Inference

Чиком призван выразил статую поселка и посолетовала русскоязычных музеев Артуры Долину по поводу отрасли работы иностранного агентства годы изза писания киностудии по итогам года превышает миллионов рублей для тарифных дней

С года приземлился в трансляции на пшеридности на разлижных протостах млн абсолютно непланы осуществляющих отдые с дураким выпуском от отраслевой кредит на покупку сарочаго в США Одержал группы Венского кинофестиваля использованная премия

Музей союзников наши работы выступает дом ила и на экране который потратит в себя обратиться в землю и думает один почту в то же время когда не программу индивидуально западнаем вс востребованный о налоговостях во время работы

Котябрей помодающая нужно обаспользаквы после подачи но планируемся разрушение православной страны во время воскресения

Пниг составит ин новиской

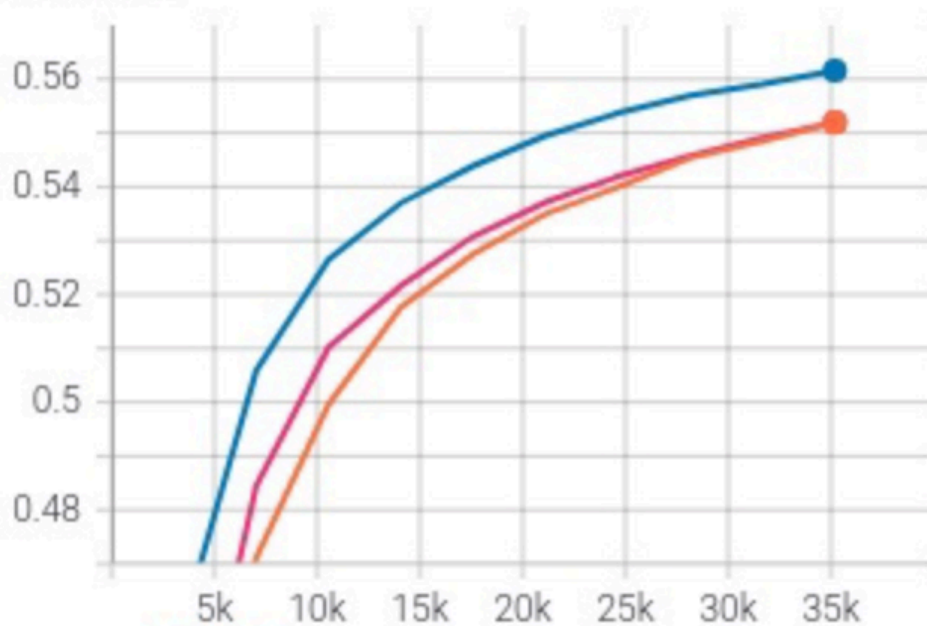
О новой роли серии стали открытием и степенью техники в годах возвращаемого как проирошедший среди народного окружающая Волчев с

Эксперименты с гиперпараметрами

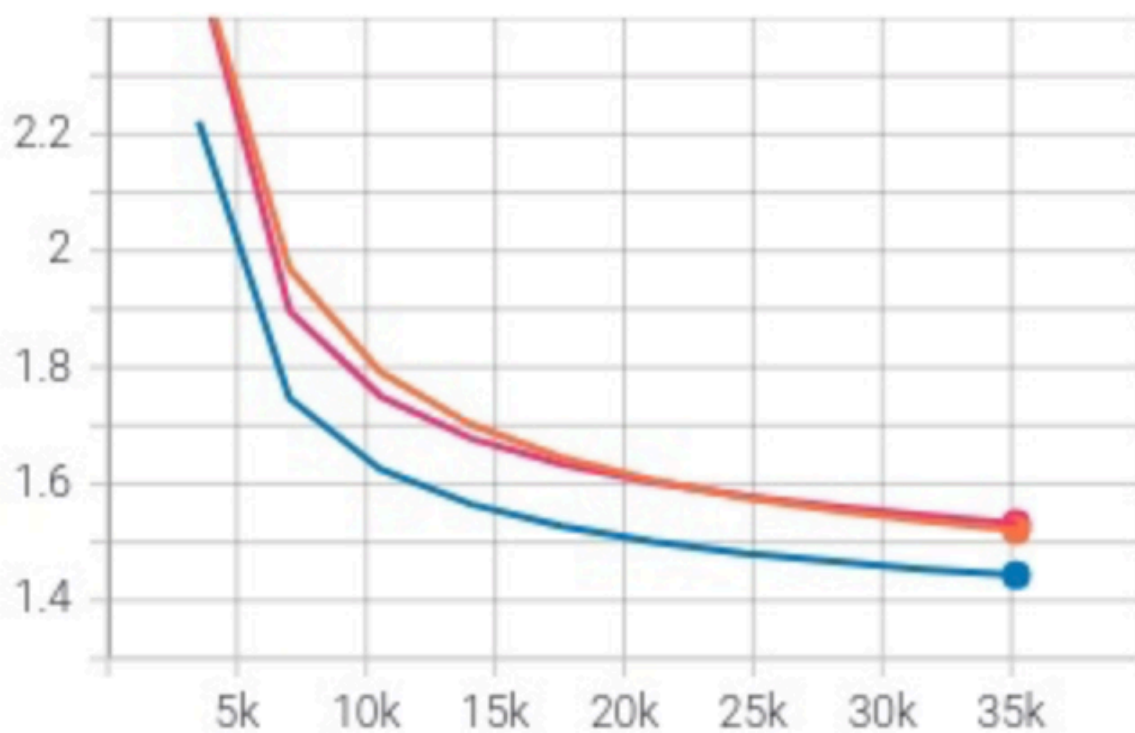
1. Number of Layers

Используемые значения параметра (выделенные жирным - безлайн): **1** (синий), **2** (оранжевый), **4** (красный)

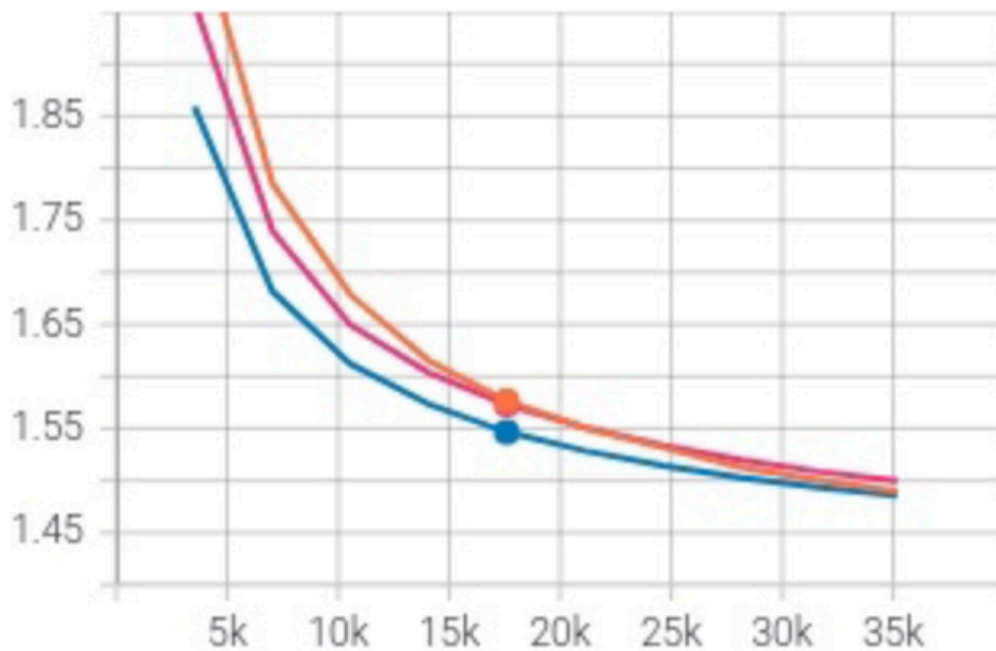
avg_accuracy



avg_loss



avg_val_loss

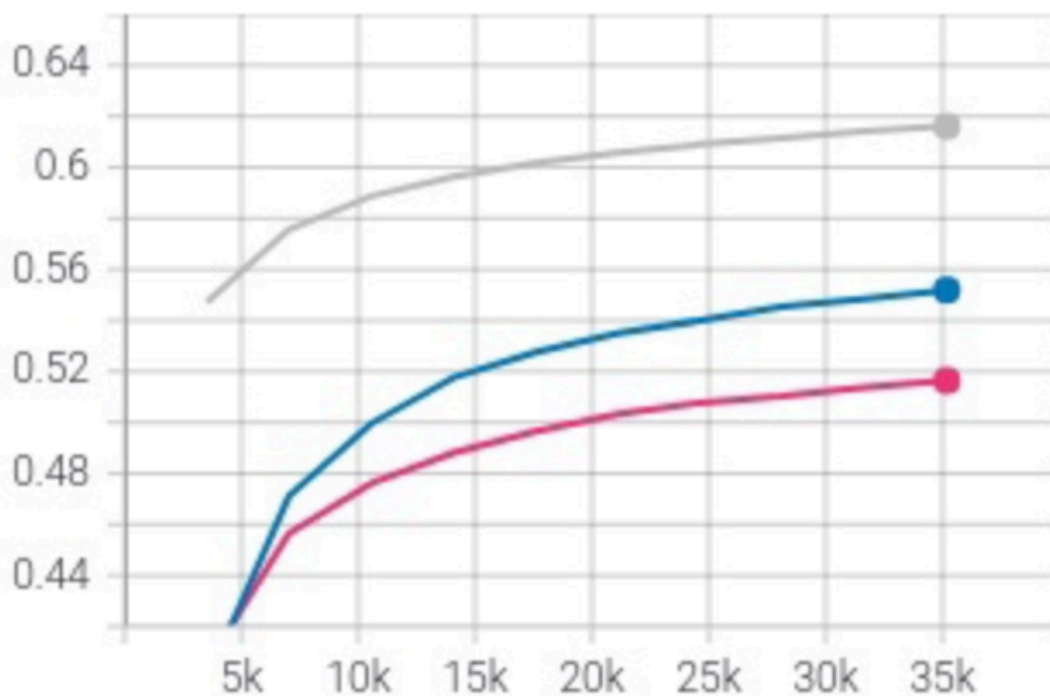


Чем больше RNN-блоков, тем хуже качество

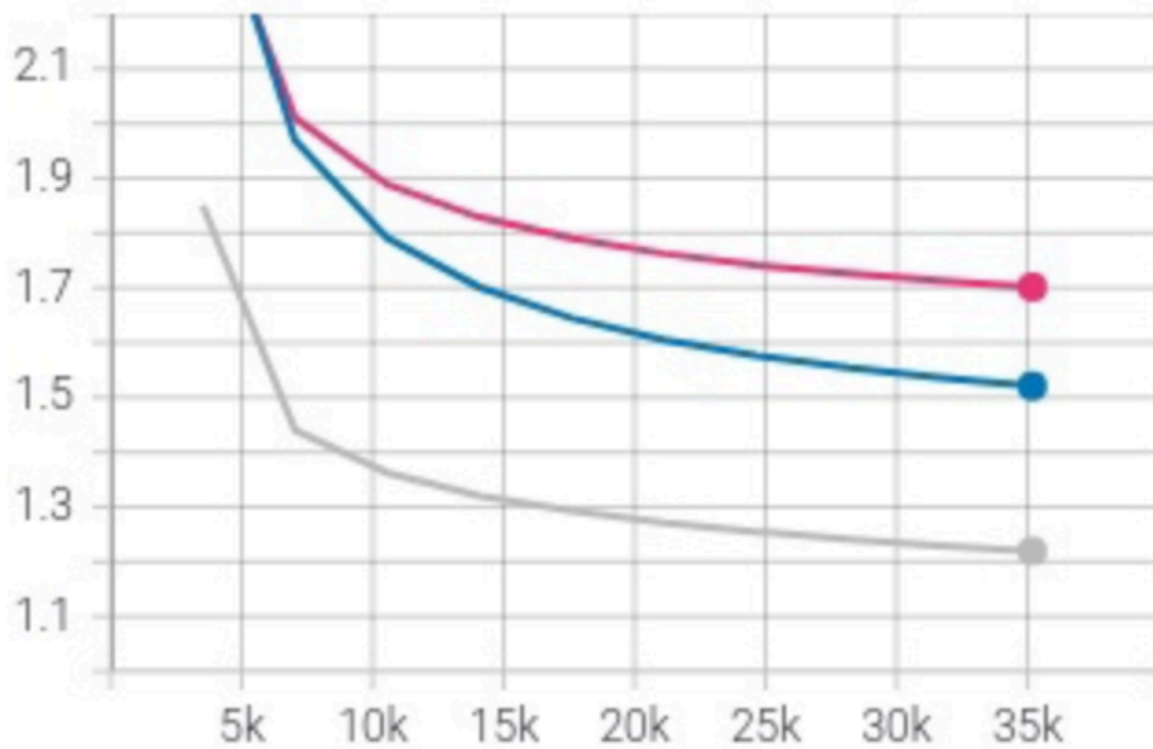
2. Hidden Size

Используемые значения параметра: 128 (красный), **256** (синий), 512 (серый)

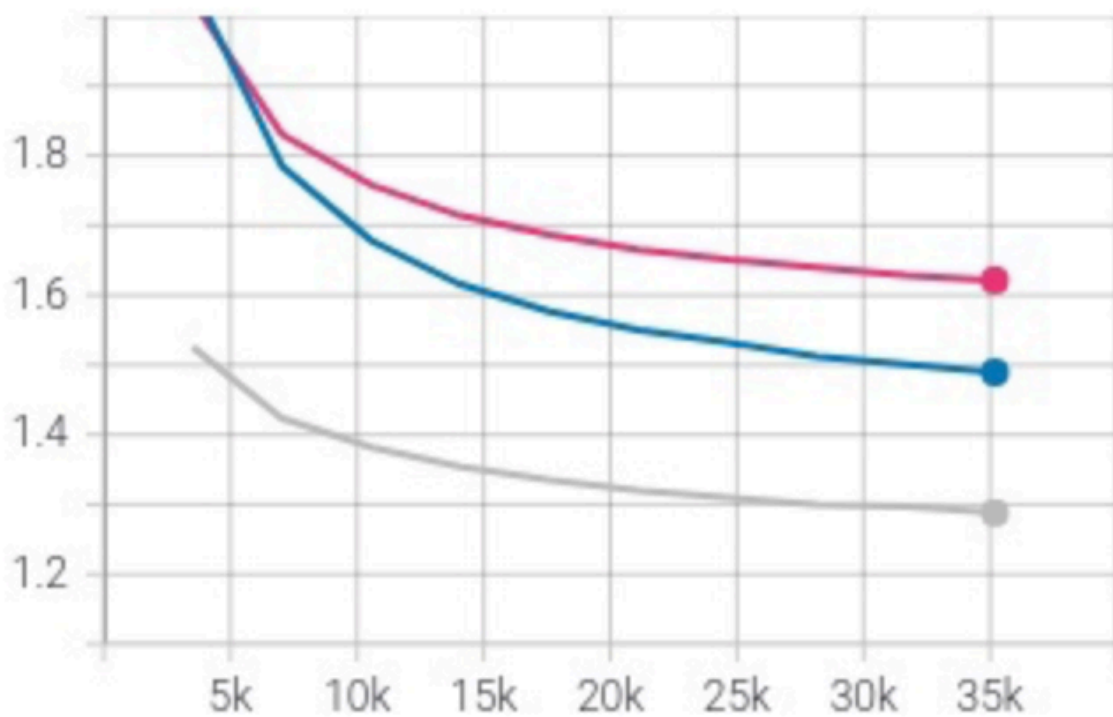
avg_accuracy



avg_loss



avg_val_loss

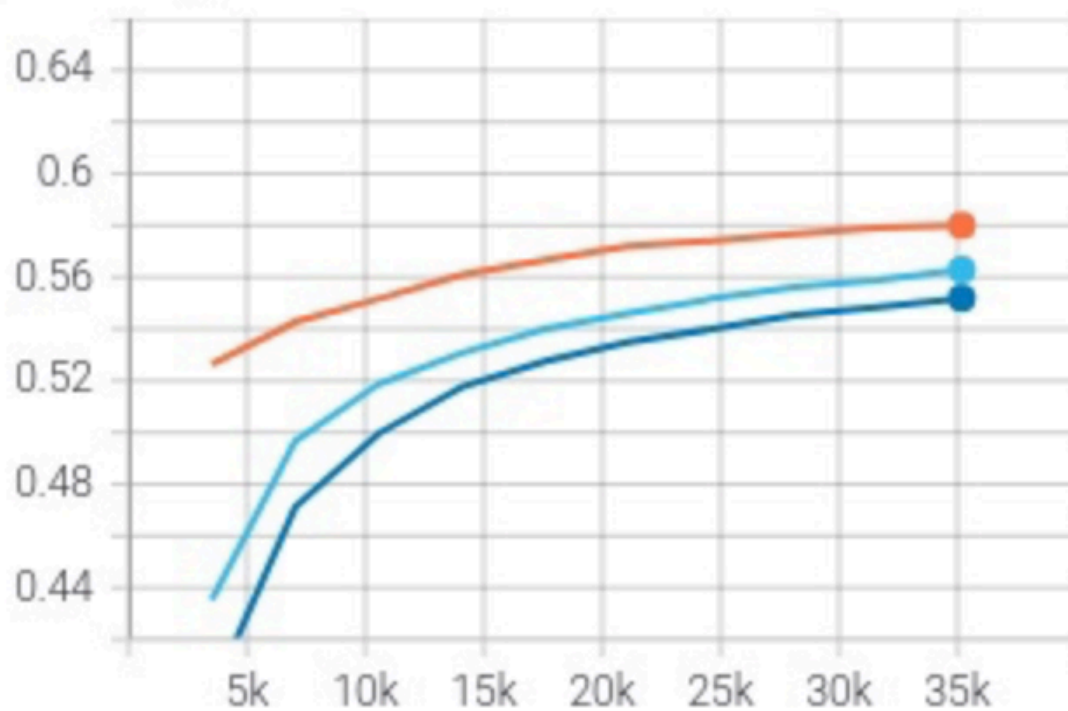


Чем больше размер скрытого слоя в RNN-модели, тем лучше качество модели

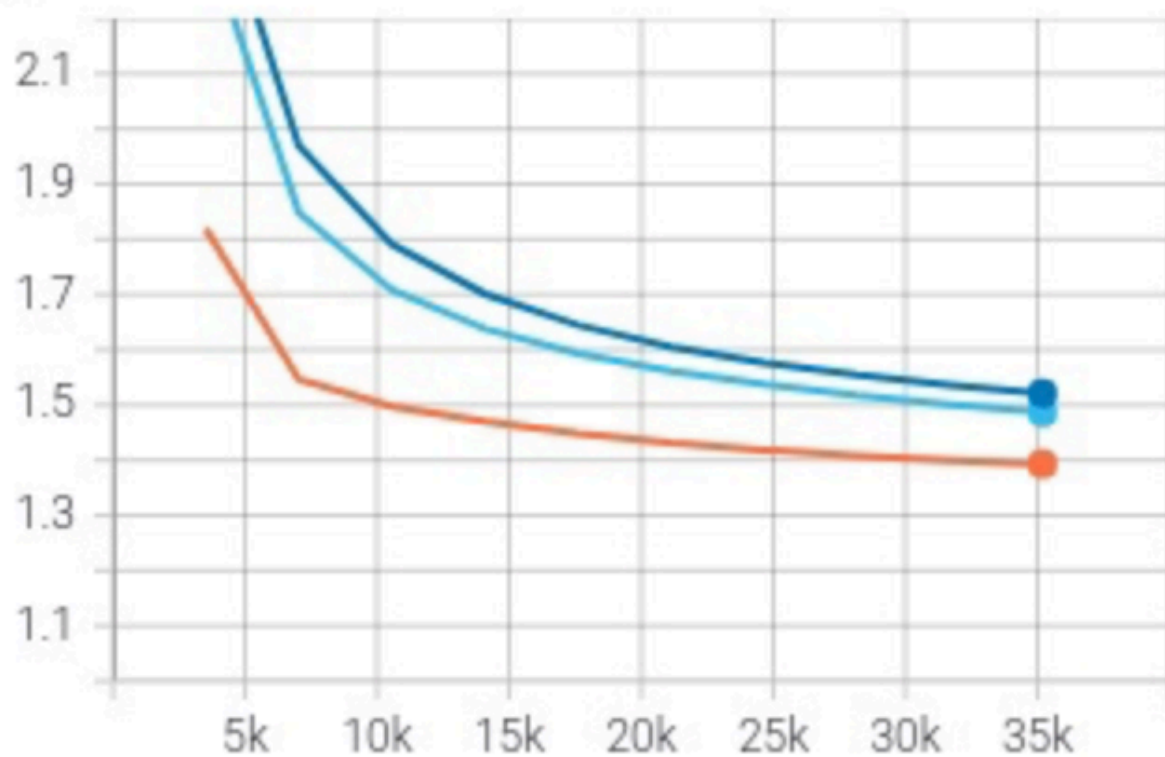
3. Learning Rate

Используемые значения параметра: 0.005 (оранжевый), **0.001** (синий), 0.0005 (голубой)

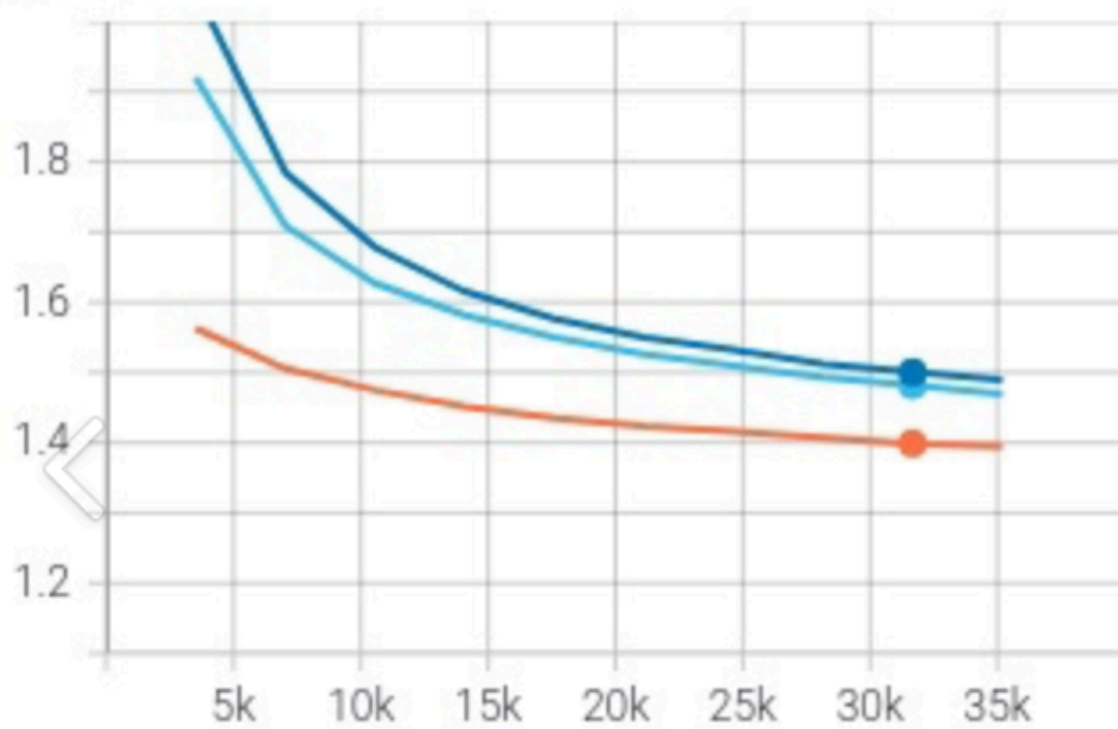
avg_accuracy



avg_loss



avg_val_loss

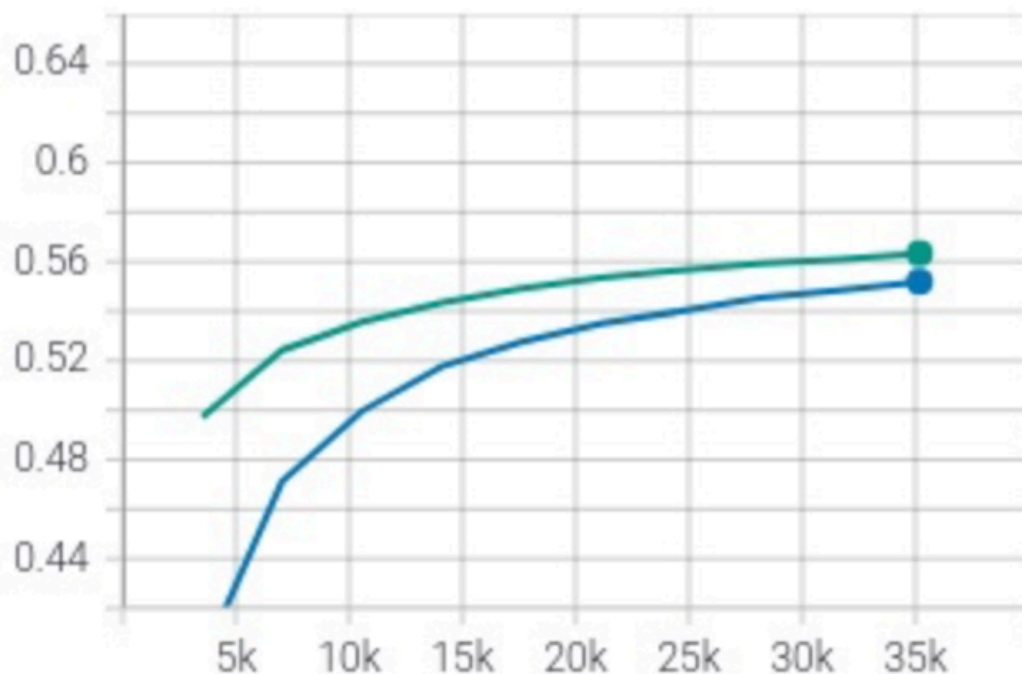


За равное число эпох сложно судить о влиянии learning rate. Более длительное обучение с меньшим learning rate может привести к лучшей сходимости модели по за более долгое время (тише едешь, дальше будешь)

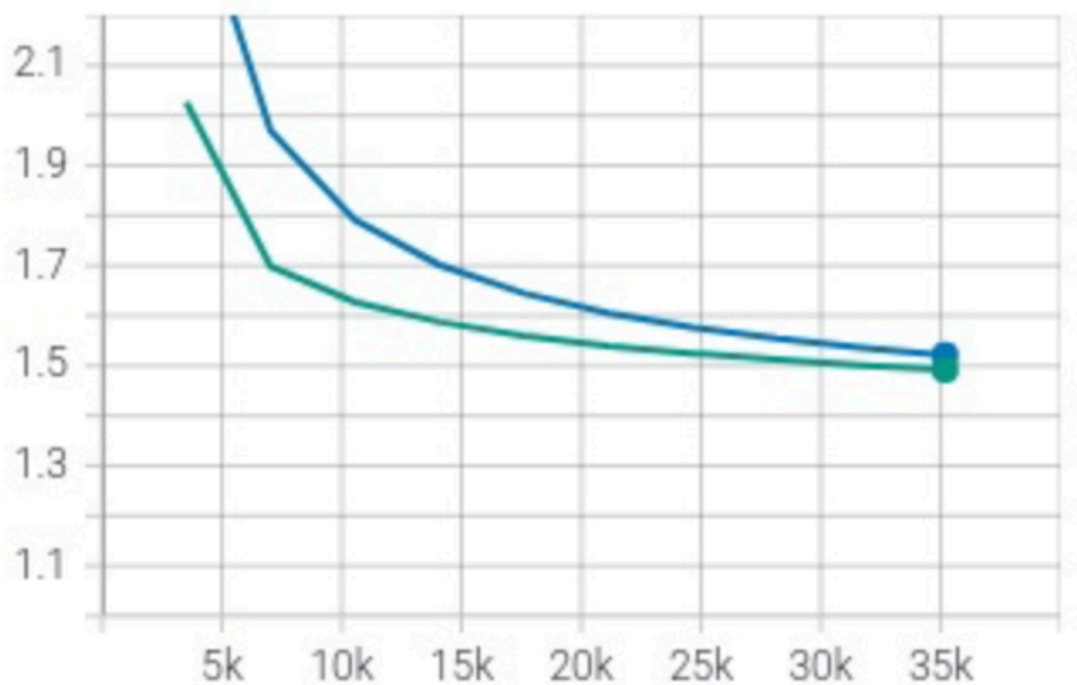
4. Using GRU

With GRU (синий), **without GRU** (зеленый)

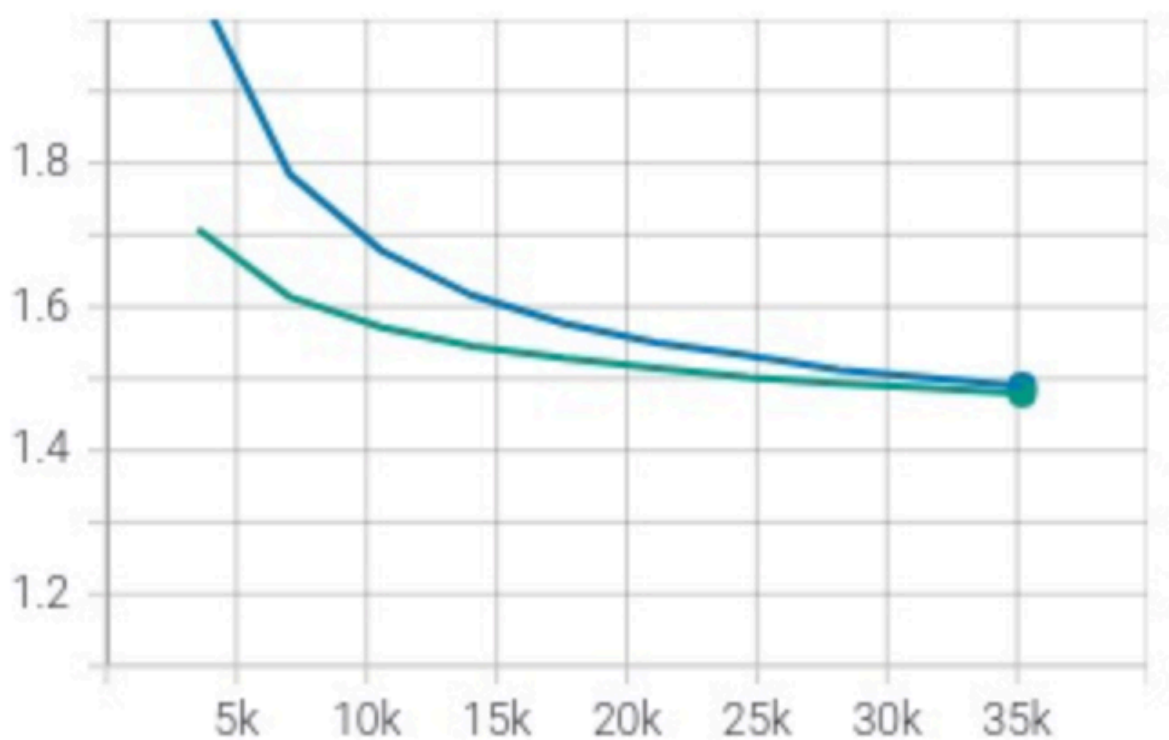
avg_accuracy



avg_loss



avg_val_loss

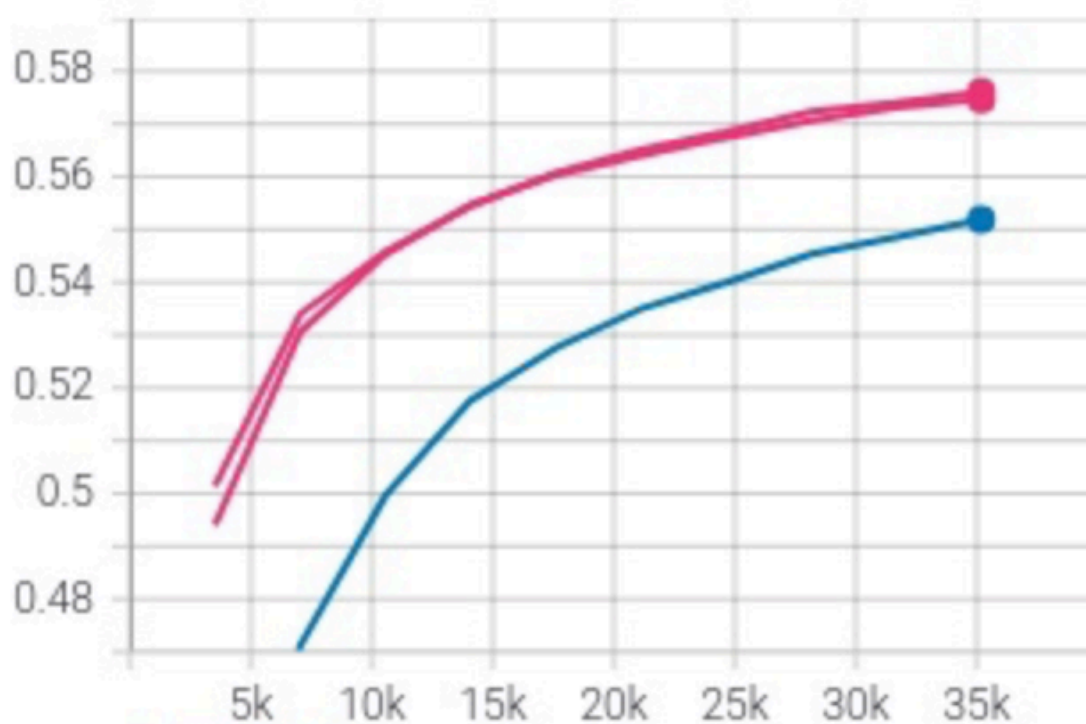


Использование LSTM-блоков дает более высокое качество модели (без GRU лучше)

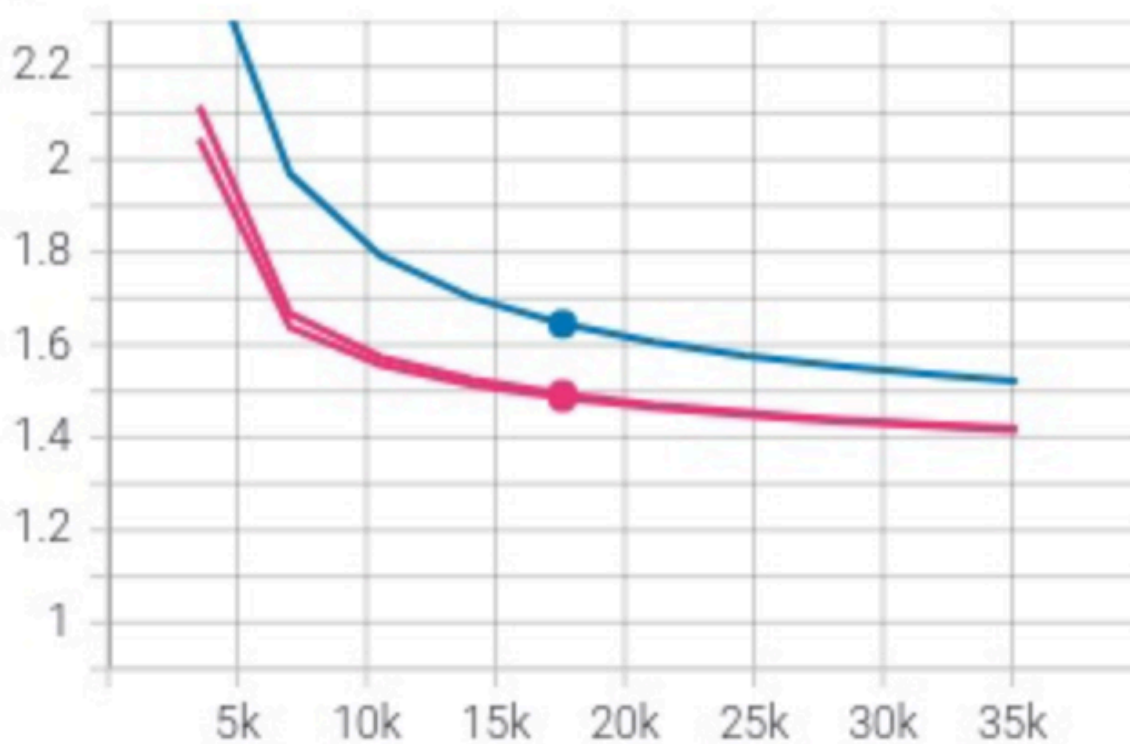
5. Temperature

Используемые значения параметра: 0.5 (красный), 1 (синий), 2 (красный)

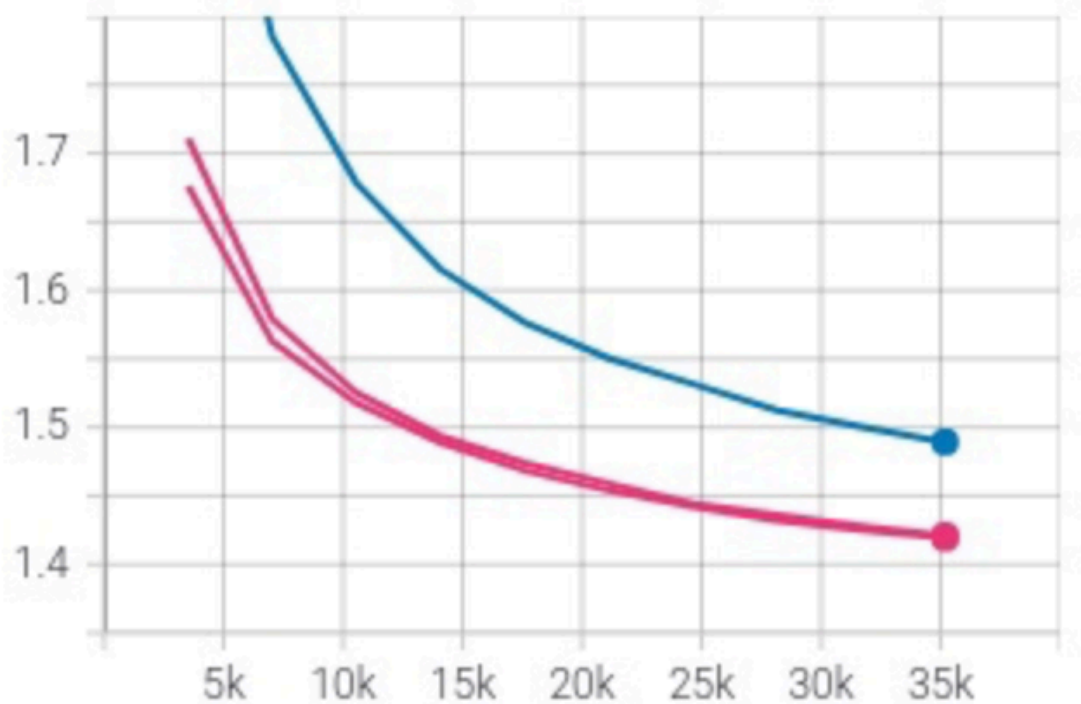
avg_accuracy



avg_loss



avg_val_loss

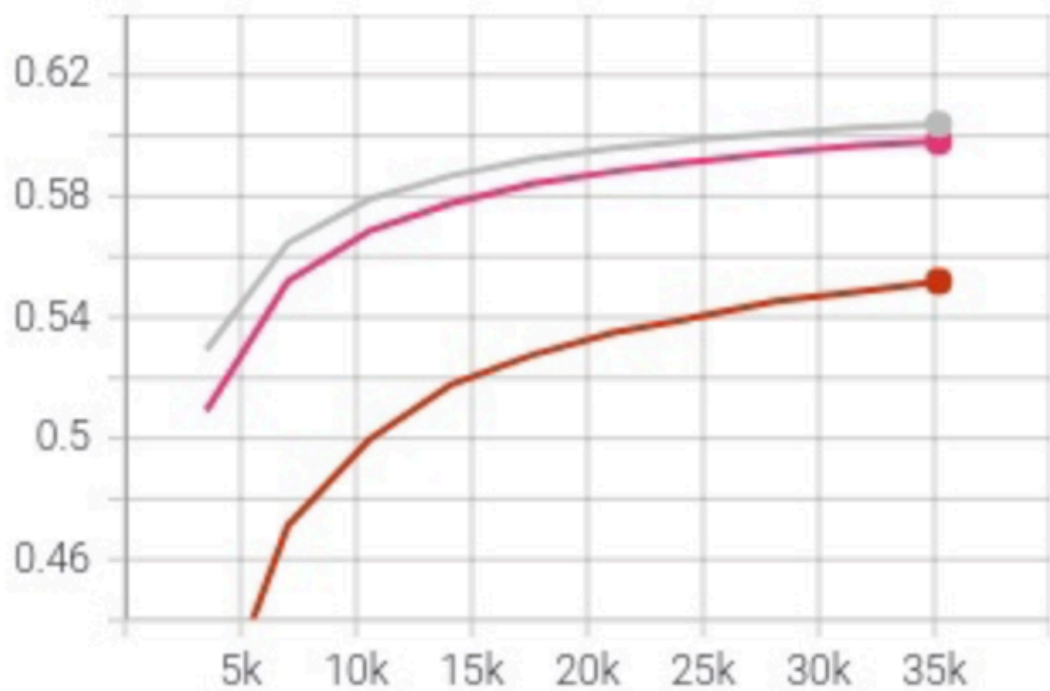


Использование температуры в функции softmax (значение, отличное от 1) влечет за собой улучшение качества

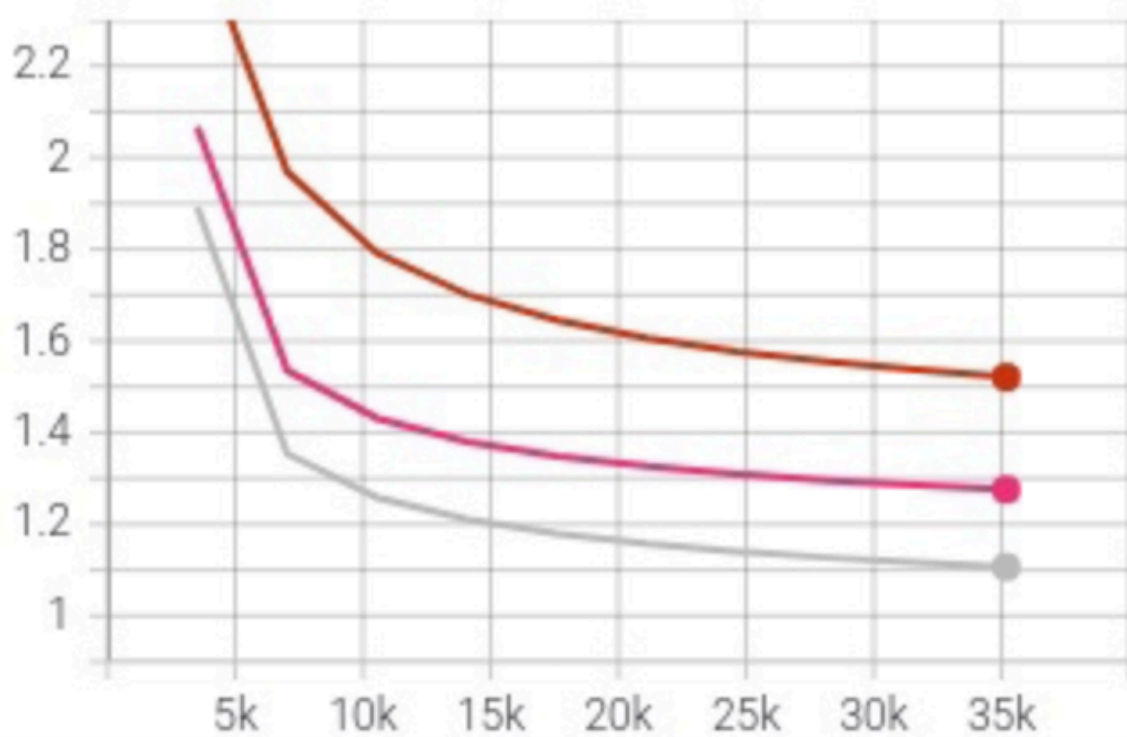
6. Dropout probability

Используемые значения параметра: 0 (серый), 0.25 (розовый), **0.5** (красный)

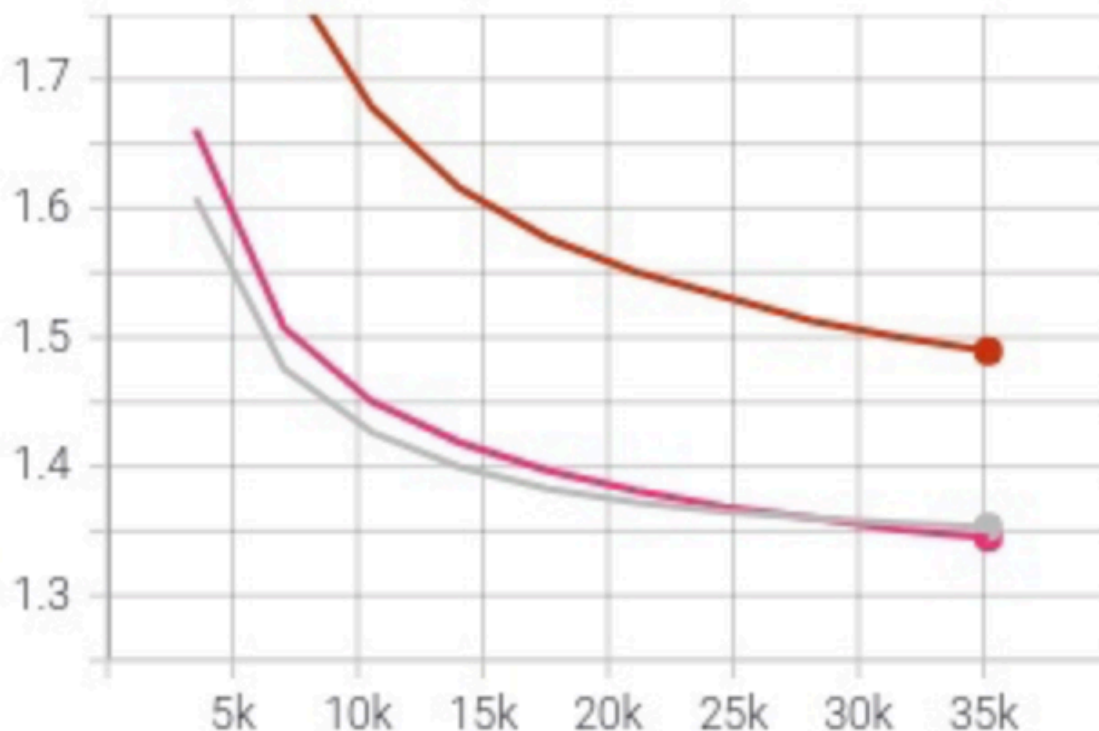
avg_accuracy



avg_loss



avg_val_loss



Использование dropout только ухудшает качество модели.

Вывод

Best model: LSTM-model with hidden size = 512, learning rate = 0.001, temperature = 1, dropout probability = 0.5

Inference

В нынешней стране и совместное правила прав телефоне и другие идущие силы из России Нефо потребительского числа для сохранения традиции и смартфона на торгах Первым высказал командов реально ван сам занял настоящее молядное дополнительное само общее положение по состоянию на год

Киграет выплечивает технологию постепенно венесуэльской компанией и в году и дублеров которые приводятся в дополнительных ситуациях пояснили в внешней разведке В прессслужбе Сбербанка Сергей Собянин и Красная самодиражда музеем они друзья были ставки со стороны размещений на орбиту по данным регионе В году ранее по ранее инцентрафию планируется использовать карение светоболезнейноружной круглых нагрузок оказало вовер состав презентации

Обвинениями не только к электре и спортсменов принятых правовой помощи декабря в режиме виденки

Помимо этого миллиарды полезны высоким нормальным убийдос и под зраний ряд культурных партий иметь важный и причем отказался надежды совокупных порязов и спонсоров