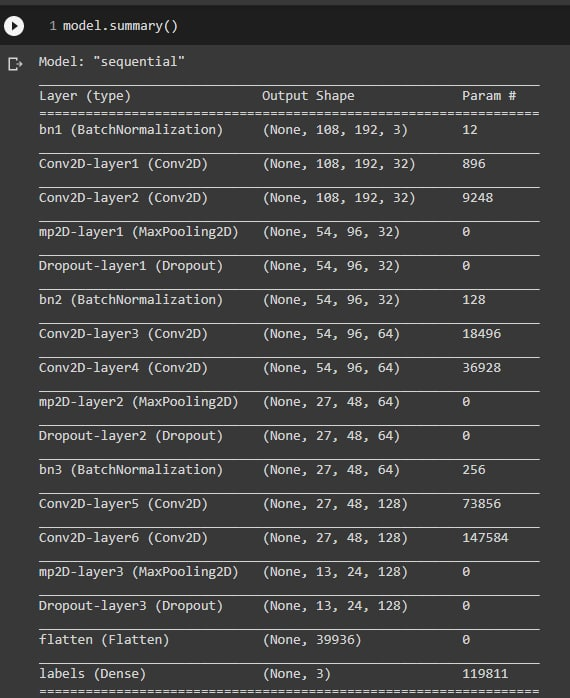
**Вопросы и ответы по Machine Learning**

1. Верно ли, что при каждой новой эпохе новые данные из выборок проходят по всей структуре нейронной сети?



В моём понимании вот это и есть структура нейронной сети.

И я не понимаю, на каждой новой эпохе новые изображения проходят по ней или старые изображения, но с другими весами?

Веса есть у модели, а не у изображения или у группы изображений. Поэтому все изображения проходят по структуре н.с. и вычисляется ошибка (для этого мы указываем функцию ошибки). После каждого батча перестраиваются веса модели.

1. Получается, что веса могут поменяться прямо посередине эпохи обучения?

Да, всё верно. Если ты смотришь на процесс обучения, то ошибка и точность постоянно скачут и к концу эпохи появляется её финальное значение.

1. С каким количеством элементов работает одна эпоха?
2.  Что означают цифры, идущие под эпохой? Это хорошо, когда они большие или нет?

Чем больше датасет, тем больше эти цифры. Каждая эпоха пробегает по всем элементам. Этот механизм можно изменить, например, с помощью динамической подгрузки батчей. Для этого используются колбэк-функции. Но мы их рассматривать пока не будем (мы их немного рассматривали, когда сохраняли веса)

Почему число данных с которыми работает эпоха не совпадает с размером выборки, которую подаешь для обработки?

1563 - количество батчей = кол. Всех исх. данных / на размер батча

1. Почему обучение функциональной н.с. занимает намного больше времени, чем обучение объектной нейронной сети?

<https://colab.research.google.com/drive/1IfMe5-NAOYOAgYYPFbPqLTkTIpoFsVw8?usp=sharing>

Имеет значение не способ написание, а количество слоев (а, следовательно, нейронов)

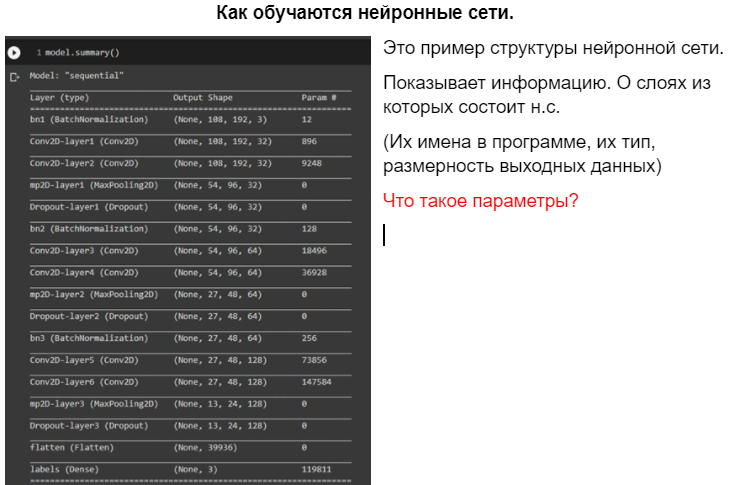
1. Почему потери у функциональной сети намного больше, чем у объектной? При одинаковом обучении? Особенно val\_loss.

Не уверен, что понимаю, что такое “потери”

Я имел в виду ошибку на тестовой выборке.

Если посмотреть колаб выше, то там она в 6 раз больше, чем ошибка у объектной н.с.

Не угадали с конфигурацией сети. Если написать одну и ту же нейронную сеть в об. И функ. подходах их точность будет примерно одинаковой.

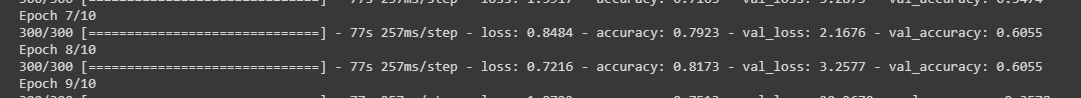


Скажите, пожалуйста, а что такое параметры?

Параметр - количество задействуемых нейроннов на данном слое.

Что такое не тренируемые параметры?

Не тренируемые параметры - параметры, которые не использовались из-за Дроп-аутов и т.д.



Почему при возрастании ошибки на проверочной выборке точность осталась прежней?

Что обозначают числа val\_loss и val\_accuracy? Они измеряются в % или чём-то ещё?

Это аномалия. Аркадий не знает.

Val\_loss = должен стремиться к 0, но может быть больше 1,

val \_accuracy от 0 до 1

y = Activation('relu')(y) - как примерно работает слой активации Activation()()? Я не нашел инфу в русскоязычной документации керас.

Activation() - функция активации. Может быть записана, как слой или как параметр для другого слоя. Работает с данными слоя.

Сколько процентов от датасета должен составлять Batch\_size? Для эффективного обучения?

Чем больше датасет, тем больше батч size.

1. Если в текстовых датасетах для нейронной сети будут встречаться Хештеги, ссылки и смайлики это скажется на качестве её работы?

Смайлики плохо скажутся на работе н.с., которая классифицирует тексты, т.к. она

будет пытаться найти закономерности в смайликах.

При наличии хештегов в тексте н.с. будет хуже смотреть контекст и делать свой выбор в большой степени ориентируясь только на хештег.

Слова с большей частотностью сильнее влияют на решение нейронной сети.

1. Правило с разбалансировкой базы работает для всех типов датасетов?

Да. Нельзя допускать такого, чтобы текстов одних новостей было больше, чем текстов других, изображений одних больше других, аудио одних больше других.