**2 слайд**

*Актуальность*

На сегодняшний день искусственный интеллект, а в частности нейронные сети, стали важным фактором во многих сферах человеческой деятельности. Они представляют выдающиеся возможности для обработки информации, обучения на основе данных и принятия сложных решений. Актуальность нейронных сетей проявляется:

* в обработке Больших Данных;
* для помощи людям в разных специальностях (например, в медицине);
* при переводе различных языков.

Но не все алгоритмы обучения нейронных сетей идеальны. Требуется тщательная настройка процесса обучения или переобучения сетей и их моделей, что связано с модернизацией самой структуры модели и её слоев.

На данном слайде мы видим, как программа (нейронная сеть) определила кота как оленя – это неправильный результат. Также на второй фотографии изображён человек (нейросеть определила, что он одет в костюм) но также определила, что на фото изображён телефон, хотя это не так. Мне стало интересно найти лучший и хороший результат нейронной сети, поэтому я провёл несколько экспериментов для определения и выявления лучшего метода для обучения нейронной сети.

**3 слайд**

Исследование алгоритмов обучения нейронных сетей и внесение собственных настроек в эти алгоритмы являются основной задачей. Оценка универсальности методов производится экспериментальным путем, с целью выявления оптимальных шагов обучения для нейронных сетей. Для достижения этой цели я выделил следующие задачи:

**4 слайд**

Нейронные сети (NN) представляют собой широко используемый метод машинного обучения, который эмулирует функционирование человеческого мозга.

Основы машинного обучения и нейронных сетей включают в себя следующие концепции:

(Классификация — задача, которая состоит в определении категории, к которой принадлежит объект. Например, определение, является письмо спамом или нет.)

(Регрессия — задача, которая состоит в определении числового значения для выходных данных. Например, определение стоимости дома на основе его характеристик.)

(Кластеризация — задача, которая состоит в группировке объектов на основе их сходства. Например, группировка пользователей интернет-магазина на основе их покупок.)

**6 слайд**

*ЭКСПЕРИМЕНТ 1 - «СВЁРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ»*

В общем, успешные результаты в обучении свёрточных моделей достигаются уже после двух эпох, что делает представленный алгоритм наилучшим вариантом. Среди задач, решаемых с использованием свёрточных нейронных сетей, включаются классификация, обнаружение объектов, синтез изображений и другие.

На данных графиках вы можете видеть, что с использованием моей программы на втором графике явных изменений не произошло. Это хороший эксперимент, но не эффективный.

**7 слайд**

Если кратко, то нормализация данных – это упрощение и сведение к одному диапазону от 0 до 1 и от -1 до 1 сводим к диапазону от 0 до 1. Всё что меньше 1 становится 0, всё что больше 1 становится 1.

На графике вы можете увидеть, что точность данной нейросети после третьей эпохи увеличивается. Но такой эксперимент мне не подходит, так как требует много времени.

**8 слайд**

Здесь показан процесс обучения нейронной сети. В нейросеть загружают какие-либо данные, после чего нейросеть благополучно определяет эти данные. Мы видим, что данный эксперимент является удачным.

**9 слайд**

На следующем слайде можно увидеть, как нейронная сеть генерирует чёрно-белое изображения. Почему чёрно-белое? Потому что чёрно-белое изображение легче генерировать т.к. цветное изображение требует больше ресурсов.

**10 слайд**

Градиенты затухания — это когда повторяются одни и те же данные с уменьшением.

Градиентный взрыв — это когда повторяются одни и те же данные с увеличением.

Батч нормализация содержит два основных метода  
Моментум - когда необходимо брать банные из предыдущих слоёв батчей.  
Эпсилон - для предотвращения деления на ноль.

Как мы можем увидеть из графиков, настройки значительно повлияли на потери данных в обучении, причём с тенденцией снижения потерь. Точность тоже повышается с увеличением значений, причём на этапе шестой эпохи наблюдается скачок повышения точности эксперимента обучения.

**11 слайд**

Например, модель пытается предсказать 10 различных цифр, и поэтому последний слой в модели имеет 10 узлов. Это очень важно, потому что на выходе для каждого узла конечного слоя будет вероятность того, что данное изображение является конкретной цифрой.

**12 слайд**

Это можно назвать удачным экспериментом, из-за его чистоты по отношению к другим. Потери теста очень малы и точность высока.

**13 слайд**

**Автор** лаборатории [Ryerson Vision lab](https://ryersonvisionlab.github.io/), [Сен Цзя](https://github.com/SenJia) глубоко исследует процесс регуляризации в каждом слое. Его эксперименты с моделью получились эффективнее и нагляднее моих.

**L1 регуляризация** — это метод регуляризации, который добавляет штраф на абсолютное значение весов в функцию потерь, чтобы уменьшить влияние ненужных или малозначимых признаков на модель.

**L2 регуляризация**, также известная как **гребневая регрессия (RidgeRegression)**, является методом регуляризации, который добавляет к функции потерь штраф на квадраты весов, чтобы уменьшить влияние больших значений весов на модель.

**14 слайд**

Dropout — метод, создающий на каждом этапе обучения новую архитектуру сети для эффективного использования данных и предотвращения переобучения. Однако использование dropout может уменьшить скорость обучения и качество модели.

В целом, dropout — эффективный метод предотвращения переобучения в нейронных сетях, улучшающий обобщение на новых данных.

На этом изображении видно, как выкидываются ненужные нейроны.

**15 слайд**

В шестом эксперименте после первой эпохи видим резкое снижение потерь, но и резкое замедление точности.

**16 слайд**

Мы наблюдаем эффективный алгоритм резкого снижения потерь в обучении и повышение точности, но также можем наблюдать значительное замедление процесса обучения. Около 110 секунд в среднем требуется для обучения в один проход, т.е. в одну эпоху.

**17 слайд**

Программа смогла сама проанализировать сходство между первой картинкой кремля и второй. Данный метод работает только с изображениями, но результат очень даже неплохой.

**18 слайд**

**19 слайд**

Алгоритм создан и функционирует, хотя не очень быстро, но эффективно, снижая потери и увеличивая точность обучения нейронных сетей.

ЦЕЛЬ ДОСТИГНУТА. Но для продолжения исследования и улучшения данного метода требуются более мощные ресурсы для обучения нейронной сети.