Московский Авиационный Институт

(Национальный Исследовательский Университет)



Факультет информационных технологий и прикладной математики

Кафедра вычислительной математики и программирования

**Курсовой проект**

**по курсу**

**«Фундаментальные концепции искусственного интеллекта»**

**Тема «Обзор нейросетевых алгоритмов и их применение для решения практических задач»**

I семестр

Выполнил студент(ка):

Четвериков В.А.

Группа: 8О-107 М-19

Преподаватель:

Булакина М.Б.

Москва, 2019

Оглавление

[Введение 3](#_Toc27674084)

[Теоретическая часть 4](#_Toc27674085)

[Параллели из биологии 4](#_Toc27674086)

[Определение ИНС 6](#_Toc27674087)

[Архитектура нейронной сети 8](#_Toc27674088)

[Сбор данных для нейронной сети 9](#_Toc27674089)

[Практическая часть 12](#_Toc27674090)

[Вывод 13](#_Toc27674091)

# Введение

Иску́сственные нейро́нные се́ти (ИНС) — математические модели, а также их программные или аппаратные реализации, построенные по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. ИНС представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Каждый процессор подобной сети имеет дело только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам. Будучи соединёнными в достаточно большую сеть с управляемым взаимодействием, такие локально простые процессоры вместе способны выполнять довольно сложные задачи.

С точки зрения машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т. п. С математической точки зрения, обучение нейронных сетей — это многопараметрическая задача нелинейной оптимизации. С точки зрения кибернетики, нейронная сеть используется в задачах адаптивного управления и как алгоритмы для робототехники. С точки зрения развития вычислительной техники и программирования, нейронная сеть — способ решения проблемы эффективного параллелизма. А с точки зрения искусственного интеллекта, ИНС является основой философского течения коннективизма и основным направлением в структурном подходе по изучению возможности построения (моделирования) естественного интеллекта с помощью компьютерных алгоритмов.

Нейронные сети не программируются в привычном смысле этого слова, они обучаются. Возможность обучения — одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что, в случае успешного обучения, сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искаженных данных.

# Теоретическая часть

# Параллели из биологии

Нейронные сети возникли из исследований в области искусственного интеллекта, а именно, из попыток воспроизвести способность биологических нервных систем обучаться и исправлять ошибки, моделируя низкоуровневую структуру мозга (Patterson, 1996). Основной областью исследований по искусственному интеллекту в 60-е - 80-е годы были экспертные системы. Такие системы основывались на высокоуровневом моделировании процесса мышления (в частности, на представлении, что процесс нашего мышления построен на манипуляциях с символами). Скоро стало ясно, что подобные системы, хотя и могут принести пользу в некоторых областях, не ухватывают некоторые ключевые аспекты человеческого интеллекта.

Мозг состоит из очень большого числа (приблизительно 10,000,000,000) нейронов, соединенных многочисленными связями (в среднем несколько тысяч связей на один нейрон, однако это число может сильно колебаться). Нейроны - это специальная клетки, способные распространять электрохимические сигналы. Нейрон имеет разветвленную структуру ввода информации (дендриты), ядро и разветвляющийся выход (аксон). Аксоны клетки соединяются с дендритами других клеток с помощью синапсов. При активации нейрон посылает электрохимический сигнал по своему аксону. Через синапсы этот сигнал достигает других нейронов, которые могут в свою очередь активироваться. Нейрон активируется тогда, когда суммарный уровень сигналов, пришедших в его ядро из дендритов, превысит определенный уровень (порог активации).

Интенсивность сигнала, получаемого нейроном (а следовательно и возможность его активации), сильно зависит от активности синапсов. Каждый синапс имеет протяженность, и специальные химические вещества передают сигнал вдоль него. Один из самых авторитетных исследователей нейросистем, Дональд Хебб, высказал постулат, что обучение заключается в первую очередь в изменениях "силы" синаптических связей. Например, в классическом опыте Павлова, каждый раз непосредственно перед кормлением собаки звонил колокольчик, и собака быстро научилась связывать звонок колокольчика с пищей. Синаптические связи между участками коры головного мозга, ответственными за слух, и слюнными железами усилились, и при возбуждении коры звуком колокольчика у собаки начиналось слюноотделение.

Таким образом, будучи построен из очень большого числа совсем простых элементов (каждый из которых берет взвешенную сумму входных сигналов и в случае, если суммарный вход превышает определенный уровень, передает дальше двоичный сигнал), мозг способен решать чрезвычайно сложные задачи.

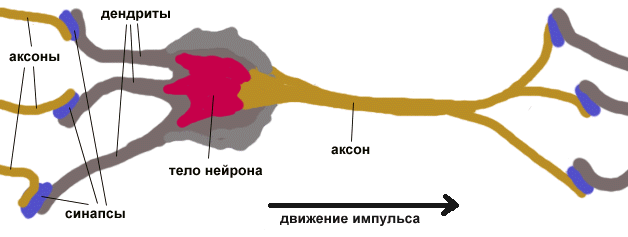


Рис. 1 Нейрон

## Определение ИНС

Искусственная нейронная сеть (ИНС, нейронная сеть) - это набор нейронов, соединенных между собой. Как правило, передаточные функции всех нейронов в нейронной сети фиксированы, а веса являются параметрами нейронной сети и могут изменяться. Некоторые входы нейронов помечены как внешние входы нейронной сети, а некоторые выходы - как внешние выходы нейронной сети. Подавая любые числа на входы нейронной сети, мы получаем какой-то набор чисел на выходах нейронной сети. Таким образом, работа нейронной сети состоит в преобразовании входного вектора в выходной вектор, причем это преобразование задается весами нейронной сети.

Основу каждой искусственной нейронной сети составляют относительно простые, в большинстве случаев - однотипные, элементы (ячейки), имитирующие работу нейронов мозга (далее под нейроном мы будем подразумевать искусственный нейрон, ячейку искусственной нейронной сети).

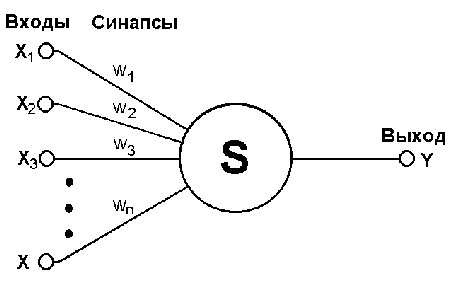


Рис. 2 - Искусственный нейрон

* Нейрон обладает группой синапсов - однонаправленных входных связей, соединенных с выходами других нейронов. Каждый синапс характеризуется величиной синоптической связи или ее весом wi.
* Каждый нейрон имеет текущее состояние, которое обычно определяется, как взвешенная сумма его входов:



* Нейрон имеет аксон - выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов. Выход нейрона есть функция его состояния:

y = f(s)

Функция f называется функцией активации.

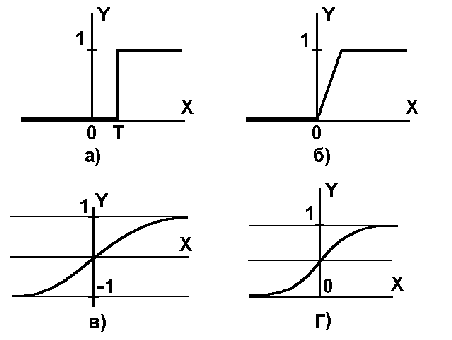


Рис. 3 - Функция активации

Функция активации может иметь разный вид :

* пороговый ( рис. 3.a),
* кусочно-линейный ( рис. 3.б),
* сигмоид( рис. 3.в, 3.г ).

Множество всех нейронов искусственной нейронной сети можно разделить на подмножества - т.н. слои. Взаимодействие нейронов происходит послойно.

Слой искусственной нейронной сети - это множество нейронов на которые в каждый такт времени параллельно поступают сигналы от других нейронов данной сети

Выбор архитектуры искусственной нейронной сети определяется задачей. Для некоторых классов задач уже существуют оптимальные конфигурации. Если же задача не может быть сведена ни к одному из известных классов, разработчику приходится решать задачу синтеза новой конфигурации. Проблема синтеза искусственной нейронной сети сильно зависит от задачи, дать общие подробные рекомендации затруднительно. В большинстве случаев оптимальный вариант искусственной нейронной сети получается опытным путем.

Искусственные нейронные сети могут быть программного и аппаратного исполнения. Реализация аппаратная обычно представляет собой параллельный вычислитель, состоящий из множества простых процессоров.

## Архитектура нейронной сети

ИНС может рассматриваться как направленный граф со взвешенными связями, в котором искусственные нейроны являются узлами. По архитектуре связей ИНС могут быть сгруппированы в два класса (рис. 4): сети прямого распространения, в которых графы не имеют петель, и рекуррентные сети, или сети с обратными связями.

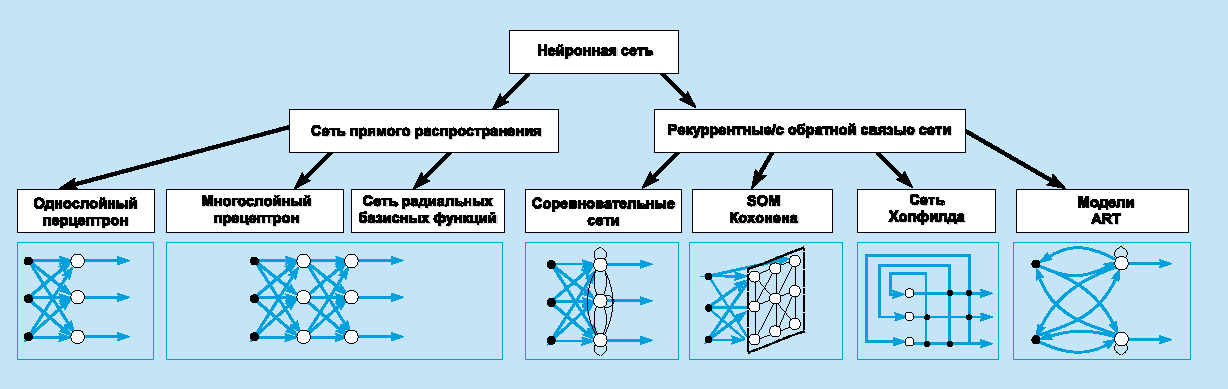


Рис. 4 - архитектуре связей ИНС

В наиболее распространенном семействе сетей первого класса, называемых многослойным перцептроном, нейроны расположены слоями и имеют однонаправленные связи между слоями. На рис. 4 представлены типовые сети каждого класса. Сети прямого распространения являются статическими в том смысле, что на заданный вход они вырабатывают одну совокупность выходных значений, не зависящих от предыдущего состояния сети. Рекуррентные сети являются динамическими, так как в силу обратных связей в них модифицируются входы нейронов, что приводит к изменению состояния сети.

## Сбор данных для нейронной сети

Если задача будет решаться с помощью нейронной сети, то необходимо собрать данные для обучения. Обучающий набор данных представляет собой набор наблюдений, для которых указаны значения входных и выходных переменных. Первый вопрос, который нужно решить, - какие переменные использовать и сколько (и каких) наблюдений собрать.

Выбор переменных (по крайней мере первоначальный) осуществляется интуитивно. Ваш опыт работы в данной предметной области поможет определить, какие переменные являются важными.

Нейронные сети могут работать с числовыми данными, лежащими в определенном ограниченном диапазоне. Это создает проблемы в случаях, когда данные имеют нестандартный масштаб, когда в них имеются пропущенные значения, и когда данные являются нечисловыми.

Более трудной задачей является работа с данными нечислового характера. Чаще всего нечисловые данные бывают представлены в виде номинальных переменных типа Пол = {Муж , Жен }.

Пусть, например, мы хотим научить нейронную сеть оценивать стоимость объектов недвижимости. Цена дома очень сильно зависит от того, в каком районе города он расположен. Город может быть подразделен на несколько десятков районов, имеющих собственные названия, и кажется естественным ввести для обозначения района переменную с номинальными значениями. К сожалению, в этом случае обучить нейронную сеть будет очень трудно, и вместо этого лучше присвоить каждому району определенный рейтинг (основываясь на экспертных оценках).

Нечисловые данные других типов можно либо преобразовать в числовую форму, либо объявить незначащими. Значения дат и времени, если они нужны, можно преобразовать в числовые, вычитая из них начальную дату (время). Обозначения денежных сумм преобразовать совсем несложно. С произвольными текстовыми полями (например, фамилиями людей) работать нельзя и их нужно сделать незначащими.

Вопрос о том, сколько наблюдений нужно иметь для обучения сети, часто оказывается непростым. Известен ряд эвристических правил, увязывающих число необходимых наблюдений с размерами сети (простейшее из них гласит, что число наблюдений должно быть в десять раз больше числа связей в сети). На самом деле это число зависит также от (заранее неизвестной) сложности того отображения, которое нейронная сеть стремится воспроизвести. С ростом количества переменных количество требуемых наблюдений растет нелинейно, так что уже при довольно небольшом (например, пятьдесят) числе переменных может потребоваться огромное число наблюдений. Эта трудность известна как "проклятие размерности", и мы обсудим ее дальше в этой главе.

Для большинства реальных задач бывает достаточно нескольких сотен или тысяч наблюдений. Для особо сложных задач может потребоваться еще большее количество, однако очень редко может встретиться (даже тривиальная) задача, где хватило бы менее сотни наблюдений. Если данных меньше, чем здесь сказано, то на самом деле у Вас недостаточно информации для обучения сети, и лучшее, что Вы можете сделать - это попробовать подогнать к данным некоторую линейную модель.

# Практическая часть

Для закрепления результатов я решил попытаться решить популярную задачу по распознаванию рукописных цифр при помощи нейросети. В своём исследовании я использовал два подхода построения сети при помощи нейросетевой библиотеки keras а также разработал несложную сеть с одним скрытым слоем и решил их сравнить.

Совю сеть я подал на вход 100 нейронов в скрытом слое. Мои попытки прилагаю в Jupyter Notebooks которые я приведу ниже.

# Вывод

В целом я очень доволен получившимся проектом. Изучил много нового и приобрёл полезные навыки. Смог осуществить полный процесс от создания плана и сбора данных до запуска. Получил хорошие идеи на будущее. Рассмотрены различные виды нейронных сетей. Решена задача распознавания рукописных цифр. В ходе решения задачи изучены особенности создания нейронных сетей и их различные архитектуры. Приобретён опыт использования нейросетевой библиотеки Keras а также опыт создания нейронных сетей со скрытыми слоями.