Федеральное государственное бюджетное образовательное

учреждение высшего образования

«Национальный исследовательский университет «МЭИ»

|  |  |  |  |
| --- | --- | --- | --- |
| Институт: | ИВТИ | Кафедра: | ПМИИ |
| Направление подготовки: | | 01.03.02 Прикладная математика и информатика | |

**ОТЧЕТ по практике**

|  |  |
| --- | --- |
| **Наименование практики:** | Производственная практика: научно-исследовательская работа |

**СТУДЕНТ**

|  |  |
| --- | --- |
|  | /Пилипишина Валерия Алексеевна |
| *(подпись )* | (*Фамилия и инициалы*) |

|  |  |
| --- | --- |
| Группа | А-13-18 |
|  | *(номер учебной группы)* |

**ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ПРАКТИКЕ**

|  |
| --- |
|  |
| *(зачтено, не зачтено)* |

|  |  |
| --- | --- |
|  | / / |
| *(подпись )* | (*Фамилия и инициалы члена комиссии*) |

|  |  |
| --- | --- |
|  | / / |
| *(подпись )* | (*Фамилия и инициалы члена комиссии*) |

**Москва**

**2021**

Оглавление

[Введение 3](#_Toc85798927)

[1. Обзор литературы 4](#_Toc85798928)

[1.2 Виды нейронных сетей и их применение 4](#_Toc85798929)

[1.3 Принцип работы сверточной нейросети 7](#_Toc85798930)

[1.3.1 Слой свертки 8](#_Toc85798931)

[1.3.2 Слой активации 8](#_Toc85798932)

[1.3.3 Субдискретизация 8](#_Toc85798933)

[1.3.4 Полносвязная нейронная сеть 9](#_Toc85798934)

[1.4. Машинное обучение 9](#_Toc85798935)

[1.4.1 Обучение с учителем 9](#_Toc85798936)

[1.4.1 Обучение без учителя 10](#_Toc85798937)

[1.4.2 Виды задач, решаемых с помощью машинного обучения 10](#_Toc85798938)

[2. Разработка алгоритма распознавания объекта на изображении 11](#_Toc85798939)

[2.1 KNN 11](#_Toc85798940)

[2.2 Нормализация 12](#_Toc85798941)

[2.3 Выделение значимых атрибутов 12](#_Toc85798942)

[2.4 Взвешенный способ 13](#_Toc85798943)

[2.5 CIFAR-10 13](#_Toc85798944)

[3. Реализация выбранной нейронной сети. 14](#_Toc85798945)

[4. Обучение выбранной нейронной сети. 14](#_Toc85798946)

[5. Анализ полученных результатов 14](#_Toc85798947)

# Введение

Потенциальными областями применения искусственных нейронных сетей являются те, где человеческий интеллект малоэффективен, а традиционные вычисления трудоёмки или физически неадекватны (т. е. не отражают или плохо отражают реальные физические процессы и объекты). Актуальность применения нейронных сетей многократно возрастает, когда появляется необходимость решения плохо формализованных задач. Основные области применения нейронных сетей: автоматизация процесса классификации, автоматизация прогнозирования, автоматизация процесса распознавания, автоматизация процесса принятия решений и другие области.

Темой данного исследования является разработка системы распознавания изображений на основе аппарата искусственных нейронных сетей. Нейронная сеть — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей. По сути, в конечном итоге мы получаем продукт, имитирующий человеческое мышление.

Объектом исследования является процесс функционирования алгоритмов распознавания объектов на изображении.

Цель исследования заключается в разработке нового алгоритма распознавания объектов на изображении.

Для достижения цели поставлены следующие задачи:

• Изучить существующие нейронные сети.

• Выбрать нейронную сеть для создания модели и изучить ее строение.

• Разработать алгоритм распознавания объекта на изображении.

• Реализовать выбранную нейронную сеть.

• Обучить выбранную нейронную сеть.

• Преобразовать входные данные для подачи их в модель нейронной сети.

• Протестировать алгоритм на примерах реальных изображений.

# 1. Обзор литературы

На основе нейронов человеческого мозга была составлена искусственная нейронная сеть. Главным понятием искусственных нейронных сетей является искусственный нейрон, который в своем роде является сумматором всех входящих в него сигналов.

На вход искусственного нейрона подается какой-либо набор данных, обозначенный на рисунке как X1, X2, …, Xn. Далее входные данные преобразуются с помощью весов, обозначенных на рисунке как w1, w2, …, wn и передаются в сумматор, обозначенную на рисунке знаком ∑.

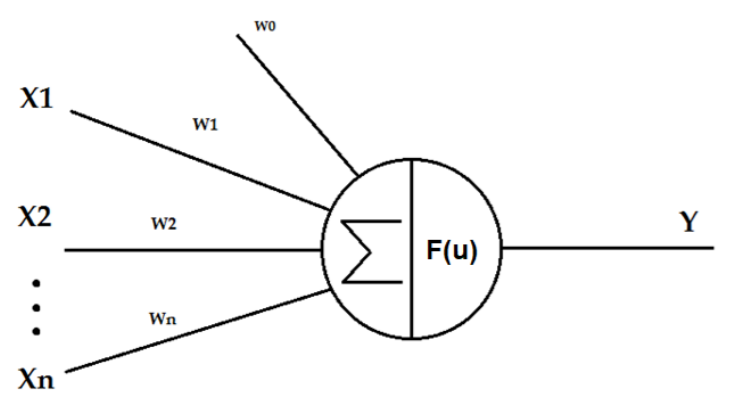


Рисунок 1.1 Строение искусственного нейрона

Формула сумматора представлена формулой (1):

(1)

Где u – значение сумматора, w – вес, x – входные данные.

Далее значение сумматора передается функции активации F(u). Данные функции имеют различные формулы и именно от них зависит выходное значение нейрона.

## 1.2 Виды нейронных сетей и их применение

* **Сети прямого распространения и персептроны.**

Базовыми архитектурами нейронных сетей являются сети прямого распространения и персептроны. Данные нейронные сети очень прямолинейны. Информация по ним передается от входа сразу к выходу. Клетки слоя данных сетей не связаны между собой, в отличии от соседних слоев, которые обычно полностью связаны. Нейронные сети прямого распространения обычно обучаются по методу обратного распространения ошибки. При таком обучении нейронная сеть получает множество данных как на вход, так и на выход.

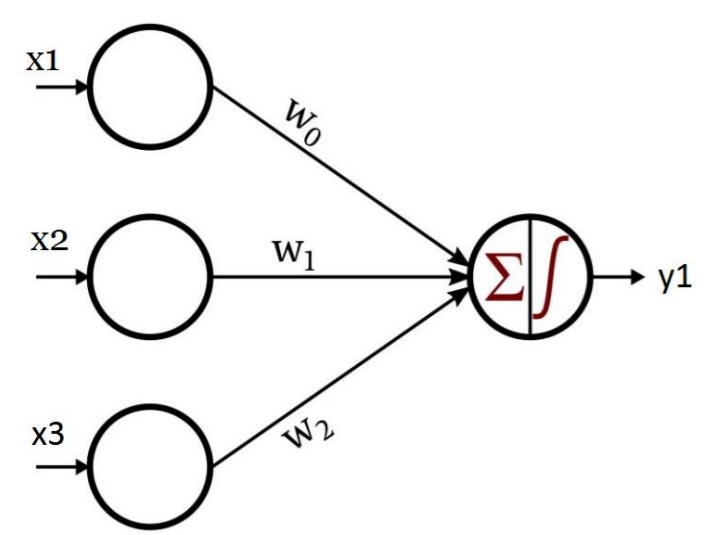
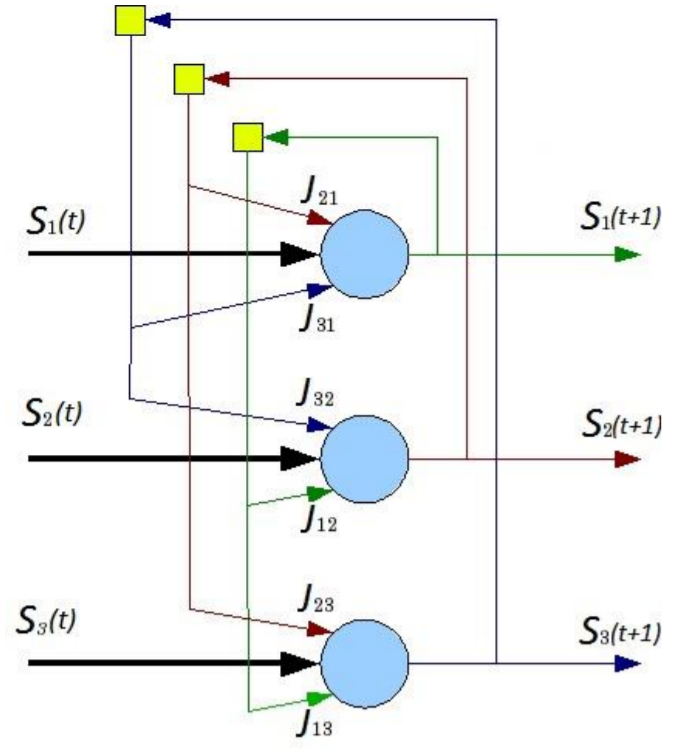


Рисунок 1.2 Пример построения персептрона

* **Сеть Хопфилда**

Еще одной разновидностью базовых нейронных сетей является нейронная сеть Хопфилда. Данная нейронная сеть используется для восстановления изображений. Нейронная сеть Хопфилда является полносвязной нейронной сетью с симметричной матрицей связей (полносвязная - каждый нейрон передает свой выходной сигнал остальным нейронам, включая самого себя). Количество нейронов данной сети определяется количеством входов и выходов.

Рисунок 1.2 Пример строения нейронной сети Хопфилда

* **Машина Больцмана**

Чаще всего данную нейронную сеть рассматривают как стохастический генеративный вариант нейронной сети Хопфилда. Машина Больцмана используется для обучения алгоритм имитации отжига. Также данная нейронная сеть оказалась первой нейронной сетью, способной обучаться внутренним представлениям и решать сложные комбинаторные задачи.

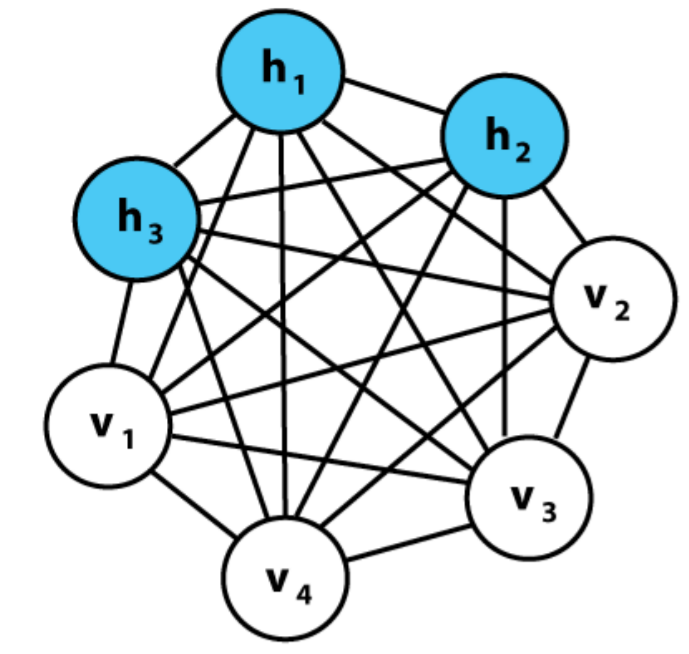


Рисунок 1.3 Пример строения машины Больцмана

* **Сверточная нейронная сеть**

Сверточная нейронная сеть - специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном и нацеленная на эффективное распознавание образов, входит в состав технологий глубокого обучения. Идея сверточных нейронных сетей заключается в чередовании сверточных слоев и слоев подвыборки. Структура сети — однонаправленная (без обратных связей), принципиально многослойная. Для обучения используются стандартные методы.

Название архитектура сети получила из-за наличия операции свертки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро) свертки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения.

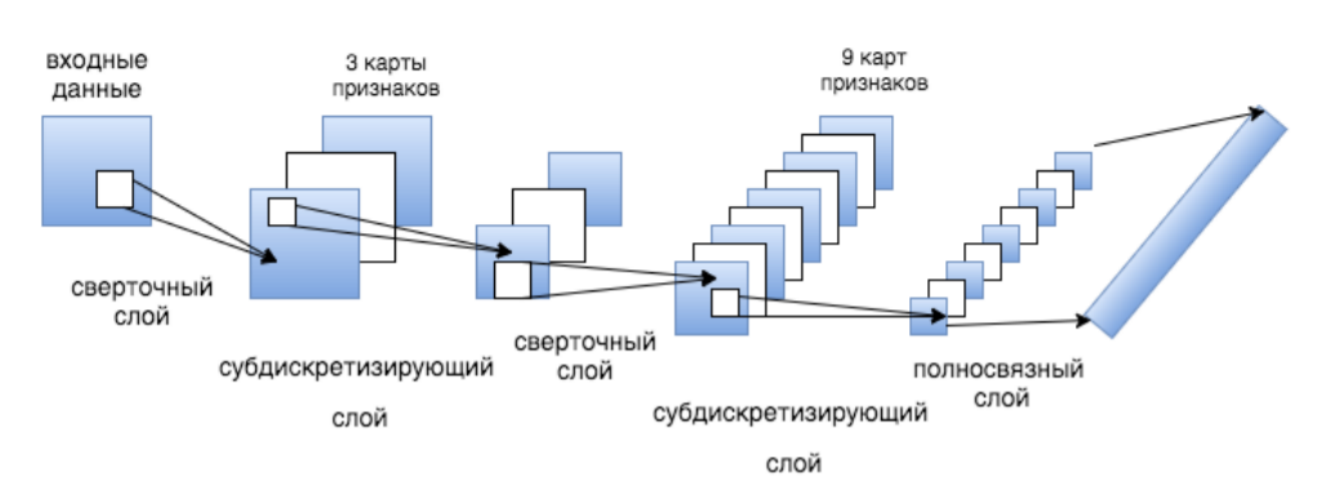


Рисунок 1.3 Пример строения сверточной нейронной сети.

Сверточная нейронная сеть обладает рядом преимуществ перед другими видами сетей для решения поставленной задачи:

* Один из лучших алгоритмов по распознаванию и классификации изображений.
* По сравнению с полносвязной нейронной сетью (типа перцептрона) — гораздо меньшее количество настраиваемых весов, так как одно ядро весов используется целиком для всего изображения, вместо того, чтобы делать для каждого пикселя входного изображения свои персональные весовые коэффициенты. Это подталкивает нейросеть при обучении к обобщению демонстрируемой информации, а не попиксельному запоминанию каждой показанной картинки в мириадах весовых коэффициентов, как это делает перцептрон.
* Удобное распараллеливание вычислений, а следовательно, возможность реализации алгоритмов работы и обучения сети на [графических процессорах](https://ru.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D1%84%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%BF%D1%80%D0%BE%D1%86%D0%B5%D1%81%D1%81%D0%BE%D1%80).
* Относительная устойчивость к повороту и сдвигу распознаваемого изображения.
* Обучение при помощи классического [метода обратного распространения ошибки](https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%B1%D1%80%D0%B0%D1%82%D0%BD%D0%BE%D0%B3%D0%BE_%D1%80%D0%B0%D1%81%D0%BF%D1%80%D0%BE%D1%81%D1%82%D1%80%D0%B0%D0%BD%D0%B5%D0%BD%D0%B8%D1%8F_%D0%BE%D1%88%D0%B8%D0%B1%D0%BA%D0%B8).

**В исследовательской работе для решения поставленной задачи я использую сверточную нейронную сеть, в связи с этим рассмотрим принцип ее работы подробнее.**

## 1.3 Принцип работы сверточной нейросети

В обычном перцептроне, который представляет собой полносвязную нейронную сеть, каждый нейрон связан со всеми нейронами предыдущего слоя, причём каждая связь имеет свой персональный весовой коэффициент. В сверточной нейронной сети в операции свёртки используется лишь ограниченная матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою (в самом начале — непосредственно по входному изображению), формируя после каждого сдвига сигнал активации для нейрона следующего слоя с аналогичной позицией. То есть для различных нейронов выходного слоя используются одна и та же матрица весов, которую также называют ядром свёртки. Её интерпретируют как графическое кодирование какого-либо признака. Тогда следующий слой, получившийся в результате операции свёртки такой матрицей весов, показывает наличие данного признака в обрабатываемом слое и её координаты, формируя так называемую карту признаков. Естественно, в сверточной нейронной сети набор весов не один, а целая гамма, кодирующая элементы изображения (например, линии и дуги под разными углами). При этом такие ядра свёртки не закладываются исследователем заранее, а формируются самостоятельно путём обучения сети классическим методом обратного распространения ошибки. Проход каждым набором весов формирует свой собственный экземпляр карты признаков, делая нейронную сеть многоканальной (много независимых карт признаков на одном слое).

Операция субдискретизации (операция «подвыборки» или объединения), выполняет уменьшение размерности сформированных карт признаков. В данной архитектуре сети считается, что информация о факте наличия искомого признака важнее точного знания его координат, поэтому из нескольких соседних нейронов карты признаков выбирается максимальный и принимается за один нейрон уплотнённой карты признаков меньшей размерности. За счёт данной операции, помимо ускорения дальнейших вычислений, сеть становится более инвариантной к масштабу входного изображения.

Рассмотрим типовую структуру сверточной нейронной сети более подробно. Сеть состоит из большого количества слоёв. После начального слоя (входного изображения) сигнал проходит серию сверточных слоёв, в которых чередуется собственно свёртка и субдискретизация. Чередование слоёв позволяет составлять «карты признаков» из карт признаков, на каждом следующем слое карта уменьшается в размере, но увеличивается количество каналов. На практике это означает способность распознавания сложных иерархий признаков. Обычно после прохождения нескольких слоёв карта признаков вырождается в вектор или даже скаляр, но таких карт признаков становятся сотни. На выходе сверточных слоёв сети дополнительно устанавливают несколько слоёв полносвязной нейронной сети (перцептрон), на вход которому подаются оконечные карты признаков.

### 1.3.1 Слой свертки

Слой свертки — это основной блок сверточной нейронной сети. Слой свертки включает в себя для каждого канала свой фильтр, ядро свертки которого обрабатывает предыдущий слой по фрагментам (суммируя результаты поэлементного произведения для каждого фрагмента). Весовые коэффициенты ядра свертки (небольшой матрицы) неизвестны и устанавливаются в процессе обучения.

Особенностью сверточного слоя является сравнительно небольшое количество параметров, устанавливаемое при обучении. Если исходное изображение имеет размерность 100×100 пикселей по трём каналам (30 000 входных нейронов), а сверточный слой использует фильтры c ядром 3×3 пикселя с выходом на 6 каналов, тогда в процессе обучения определяется только 9 весов ядра, однако по всем сочетаниям каналов, то есть 9×3×6=162, в таком случае данный слой требует нахождения только 162 параметров, что существенно меньше количества искомых параметров полносвязной нейронной сети.

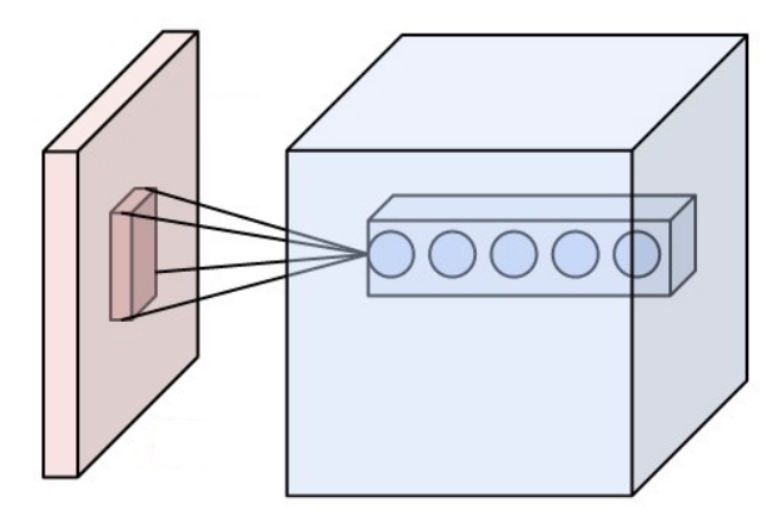


Рисунок 1.4 Нейронные слоя свертки, преобразуемые по нескольким выходным каналам

### 1.3.2 Слой активации

Скалярный результат каждой свёртки попадает на функцию активации, которая представляет собой некую нелинейную функцию. Слой активации обычно логически объединяют со слоем свёртки (считают, что функция активации встроена в слой свёртки). Функция нелинейности может быть любой по выбору исследователя, традиционно для этого использовали функции типа гиперболического тангенса или сигмоиды . Однако в 2000х годах была исследована новая функция активации — ReLU, которая позволила существенно ускорить процесс обучения и одновременно упростить вычисления (за счёт простоты самой функции), что означает блок линейной ректификации, вычисляющий функцию . То есть по сути это операция отсечения отрицательной части скалярной величины.

### 1.3.3 Субдискретизация

Слой субдискретизации (иначе подвыборки) представляет собой нелинейное уплотнение карты признаков, при этом группа пикселей (обычно размера 2×2) уплотняется до одного пикселя, проходя нелинейное преобразование. Наиболее употребительна при этом функция максимума. Преобразования затрагивают непересекающиеся прямоугольники или квадраты, каждый из которых ужимается в один пиксель, при этом выбирается пиксель, имеющий максимальное значение. Операция субдискретизации позволяет существенно уменьшить пространственный объём изображения. Она интерпретируется так: если на предыдущей операции свёртки уже были выявлены некоторые признаки, то для дальнейшей обработки настолько подробное изображение уже не нужно, и оно уплотняется до менее подробного. К тому же фильтрация уже ненужных деталей помогает не переобучаться. Слой подвыборки, как правило, вставляется после слоя свертки перед слоем следующей свертки.

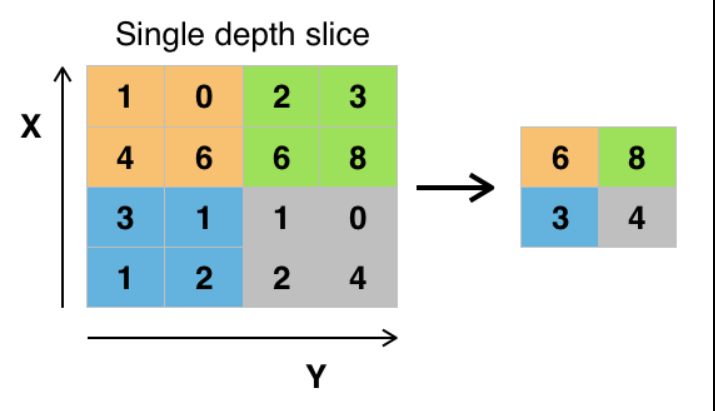


Рисунок 1.5 Субдискретизация с функцией максимума и фильтром 2х2 с шагом 2

### 1.3.4 Полносвязная нейронная сеть

После нескольких прохождений свёртки изображения и уплотнения с помощью пулинга система перестраивается от конкретной сетки пикселей с высоким разрешением к более абстрактным картам признаков, как правило на каждом следующем слое увеличивается число каналов и уменьшается размерность изображения в каждом канале. В конце концов остаётся большой набор каналов, хранящих небольшое число данных (даже один параметр), которые интерпретируются как самые абстрактные понятия, выявленные из исходного изображения.

Эти данные объединяются и передаются на обычную полносвязную нейронную сеть, которая тоже может состоять из нескольких слоёв. При этом полносвязные слои уже утрачивают пространственную структуру пикселей и обладают сравнительно небольшой размерностью (по отношению к количеству пикселей исходного изображения).

## 1.4. Машинное обучение

### 1.4.1 Обучение с учителем

Обучение с учителем является задачей обучения системы на тренировочном наборе данных. Путем подгонки результатов обучения к тренировочному набору данных, мы хотим найти наиболее оптимальные параметры модели для прогнозирования возможных ответов на других объектах (тестовых наборах данных). Если множество возможных ответов является действительным числом, то это задача регрессии. Если множество возможных ответов имеет ограниченное количество значений, где эти значения являются неупорядоченными, то это задача классификации.

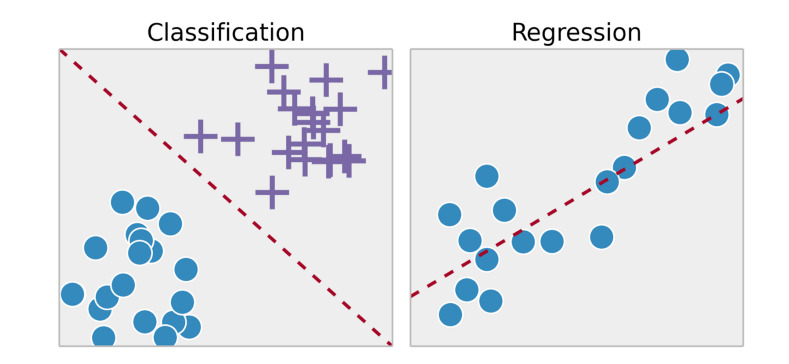


Рисунок 1.6 Процесс классификации и регрессии в машинном обучении

### 1.4.1 Обучение без учителя

В неконтролируемом обучении у нас меньше информации об объектах. В частности, тренировочный набор данных не имеет маркированных данных, относящихся к определённому классу заранее предопределённых данных. Наша цель сейчас - наблюдать некоторое сходство между группами объектов и включать их в соответствующие кластеры. Некоторые объекты могут сильно отличаться от всех кластеров, и таким образом мы предполагаем, что эти объекты являются аномалиями.

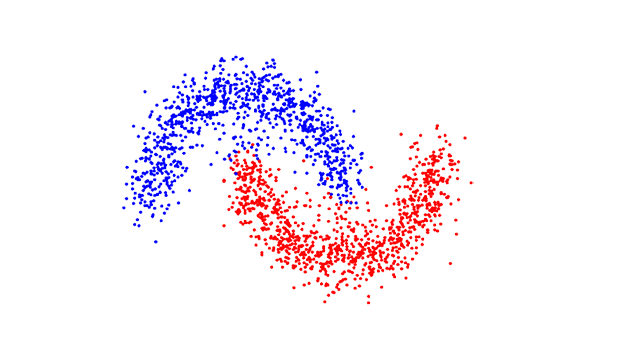


Рисунок 1.7 Процесс кластеризации

### 1.4.2 Виды задач, решаемых с помощью машинного обучения

**1)Задача регрессии** – прогноз на основе выборки объектов с различными признаками.  
Регрессия – функция, позволяющая по величине одного корреллируемого признака определить среднюю величину другого признака.

**2)Задача классификации** – задача, в которой имеется множество объектов (ситуаций), разделённых, некоторым образом на классы. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется выборкой. Классовая принадлежность остальных объектов неизвестна. Требуется построить алгоритм, способный классифицировать (см. ниже) произвольный объект из исходного множества.

Классифицировать объект — значит, указать номер (или наименование) класса, к которому относится данный объект.

Классификация объекта — номер или наименование класса, выдаваемый алгоритмом классификации в результате его применения к данному конкретному объекту.  
  
**3)Задача кластеризации** – группировка множества объектов на подмножества (кластеры) таким образом, чтобы объекты из одного кластера были более похожи друг на друга, чем на объекты из других кластеров по какому-либо критерию.

**4)Задача уменьшения размерности** – сведение большого числа признаков к меньшему для удобства их последующей визуализации (например, сжатие данных).  
  
**5)Задача выявления аномалий** – отделение аномалий от стандартных случаев. Наличие в нем признаков избыточных, неинформативных или слабо информативных может понизить эффективность модели, а после такого преобразования она упрощается, и соответственно уменьшается размер набора данных в памяти и ускоряется работа алгоритмов ML на нем.

В исследовательской работе я буду использовать нейросети для решения задач классификации. Эта задача охватывает наибольший спектр областей человеческой жизни и может быть применена в разных сферах деятельности. В качестве примера, я хочу решить задачу определения модели автомобиля по фотографии (в теории это может быть применимо для сравнения нарисованного человеком автомобиля, с похожим на него, реально существующим автомобилем, после конвертации входного изображения в набор пикселей (для дипломной работы)).

# 2. Разработка алгоритма распознавания объекта на изображении

## 2.1 KNN

Распознавание объектов на изображении буду выполнять с помощью метода kNN (метод k-ближайших соседей (k-nearest neighbors algorithm)).

В случае использования метода для классификации объект присваивается тому классу, который является наиболее распространённым среди k соседей данного элемента, классы которых уже известны. В случае использования метода для регрессии, объекту присваивается среднее значение по k ближайшим к нему объектам, значения которых уже известны.

Алгоритм может быть применим к выборкам с большим количеством атрибутов (многомерным). Для этого перед применением нужно определить функцию расстояния; классический вариант такой функции — евклидова метрика.

Пример классификации k-ближайших соседей.



Рисунок 2.1 Пример классификации K-ближайших соседей

Тестовый образец (зелёный круг) должен быть классифицирован как синий квадрат (класс 1) или как красный треугольник (класс 2):

* если k = 3, то он классифицируется как 2-й класс  
  (внутри меньшего круга 2 треугольника и только 1 квадрат);
* если k = 5, то он будет классифицирован как 1-й класс   
  (3 квадрата против 2 треугольников внутри большего круга);

## 2.2 Нормализация

Разные атрибуты могут иметь разный диапазон представленных значений в выборке (например, атрибут А представлен в диапазоне от 0.1 до 0.5, а атрибут Б представлен в диапазоне от 1000 до 5000), то значения дистанции могут сильно зависеть от атрибутов с бо́льшими диапазонами. Поэтому данные обычно подлежат нормализации. При кластерном анализе есть два основных способа нормализации данных: минимакс-нормализация и Z-нормализация.

Минимакс-нормализация осуществляется следующим образом:

в этом случае все значения будут лежать в диапазоне от 0 до 1   
(дискретные бинарные значения определяются как 0 и 1).

Z-нормализация:

*,*

где — среднеквадратичное отклонение;   
в этом случае большинство значений попадёт в диапазон (-3; 3).

## 2.3 Выделение значимых атрибутов

Некоторые значимые атрибуты могут быть важнее остальных, поэтому для каждого атрибута может быть задан в соответствие определённый вес (например вычисленный с помощью тестовой выборки и оптимизации ошибки отклонения). Таким образом, каждому атрибуту k будет задан в соответствие вес zk, так что значение атрибута будет попадать в диапазон [0; zk max(k)] (для нормализованных значений по минимакс-методу). Например, если атрибуту присвоен вес 2,7, то его нормализованно-взвешенное значение будет лежать в диапазоне [0; 2,7]

## 2.4 Взвешенный способ

При взвешенном способе во внимание принимается не только количество попавших в область определённых классов, но и их удалённость от нового значения.

Для каждого класса j определяется оценка близости:

где - расстояние от нового значения до объекта .

У какого класса выше значение близости, тот класс и присваивается новому объекту.

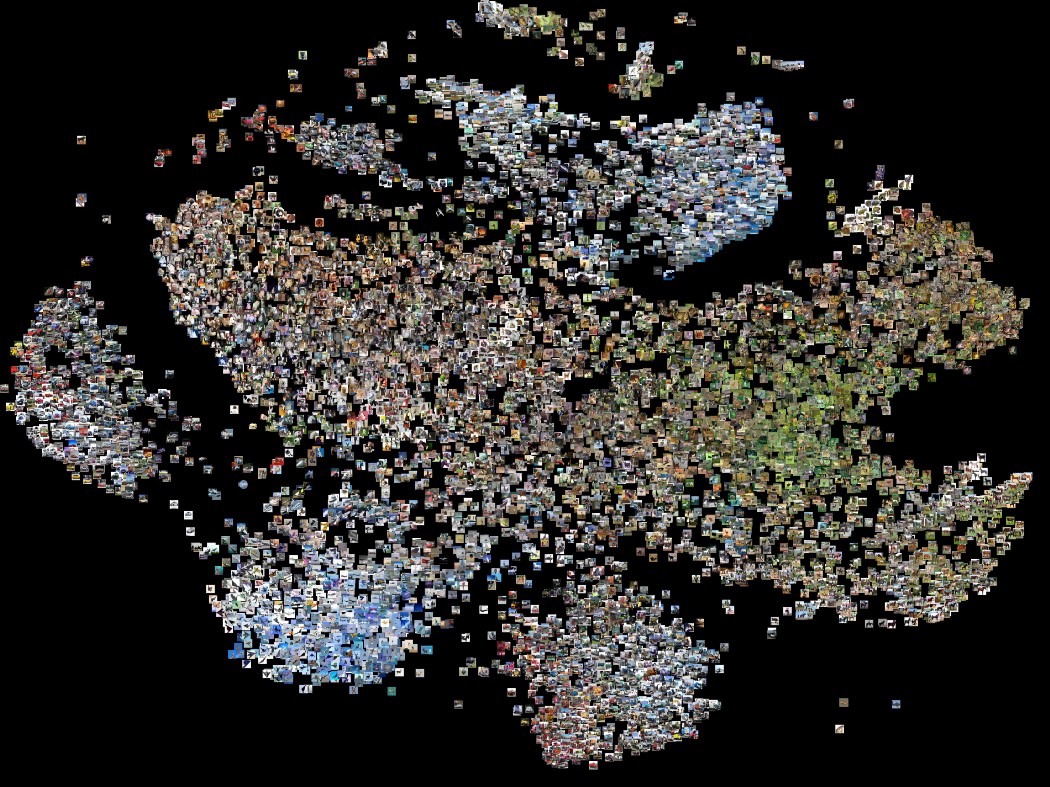
С помощью метода можно вычислять значение одного из атрибутов классифицируемого объекта на основании дистанций от попавших в область объектов и соответствующих значений этого же атрибута у объектов:{\displaystyle x\_{k}={\frac {\sum \_{i=1}^{n}{k\_{i}d(x,a\_{i})^{2}}}{\sum \_{i=1}^{n}{d(x,a\_{i})^{2}}}}}

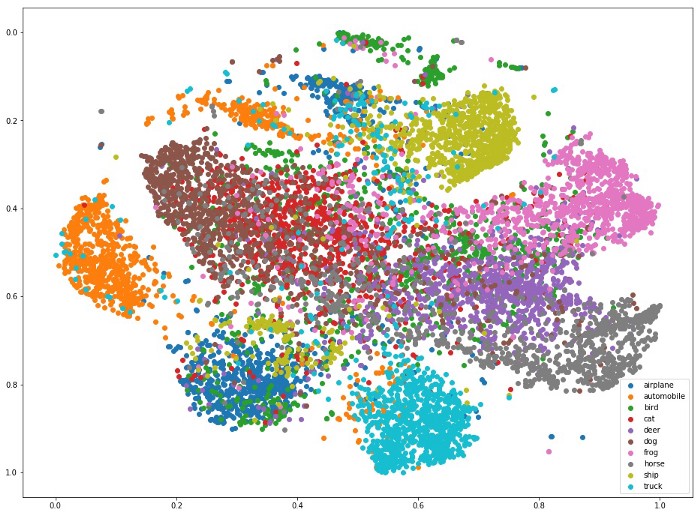
где {\displaystyle a\_{i}}{\displaystyle i}-ый объект, попавший в область,  
{\displaystyle k\_{i}} — значение атрибута{\displaystyle k} у заданного объекта {\displaystyle a\_{i}},  
{\displaystyle x} — новый объект,  
{\displaystyle x\_{k}} {\displaystyle k}-ый атрибут нового объекта.

## *2.5* CIFAR-10

В качестве примера работы этого алгоритма для задачи классификации изображения рассмотрим визуализацию построения плоскости объектов. В качестве тренировочных данных я буду использовать набора CIFAR-10. Набор состоит из 60 000 фотографий, разделенных на 10 классов: самолеты, автомобили, птицы, олени, собаки, лягушки, лошади, корабли и грузовые машины.

После визуализации мы получим следующую картину:





# 3. Реализация выбранной нейронной сети.

# 4. Обучение выбранной нейронной сети.

# 5. Анализ полученных результатов