

“МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ В ОБРАБОТКЕ ЗВУКОВЫХ И РЕЧЕВЫХ СИГНАЛОВ”

Подготовил студент группы 150702

Кутняк А. В.

ВОПРОСЫ, РАССМАТРИВАЕМЫЕ В РАМКАХ ПРЕЗЕНТАЦИИ

- Основные понятия машинного обучения
- Цифровое представление звуковых сигналов
- Задачи обработки звуковых и речевых сигналов
- Искусственные нейронные сети и их преимущества
- Задача шумоподавления

МАШИННОЕ ОБУЧЕНИЕ

Машинное обучение – подраздел искусственного интеллекта, включающий в себя методы, характерной чертой которых является не прямое решение поставленной задачи, а обучение за счет применения решений множества сходных задач.

Различают два типа машинного обучения:

- Индуктивное обучение
- Дедуктивное обучение

ТИПЫ МАШИННОГО ОБУЧЕНИЯ

ИНДУКТИВНОЕ ОБУЧЕНИЕ (ОБУЧЕНИЕ ПО ПРЕЦЕДЕНТАМ)

Основывается на выявлении эмпирических закономерностей в данных

ДЕДУКТИВНОЕ ОБУЧЕНИЕ (ЭКСПЕРТНЫЕ СИСТЕМЫ)

Предполагает формализацию знаний экспертов и представления их в виде базы знаний

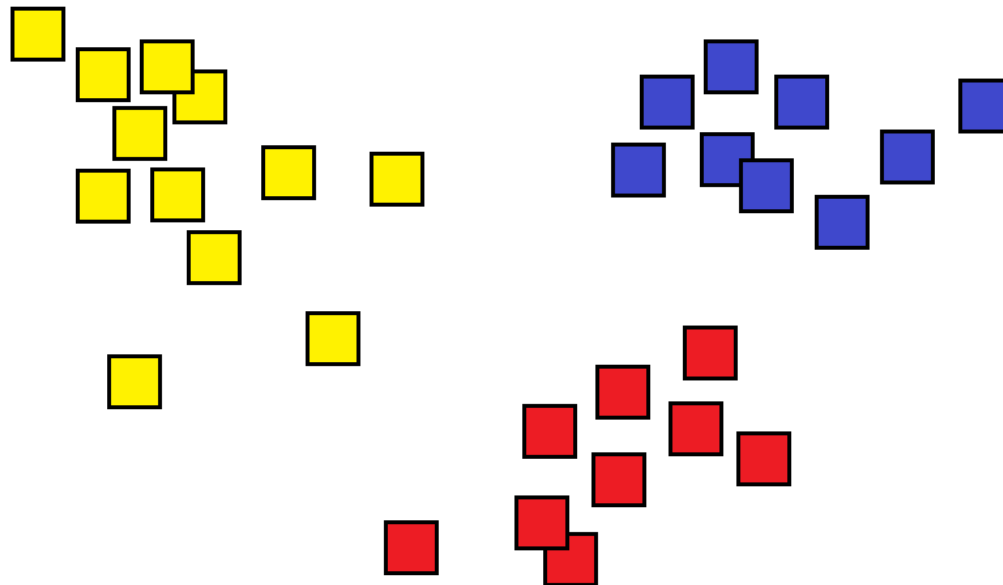
ПРЕИМУЩЕСТВА МАШИННОГО ОБУЧЕНИЯ

- Точность
- Автоматизация
- Скорость
- Масштабируемость

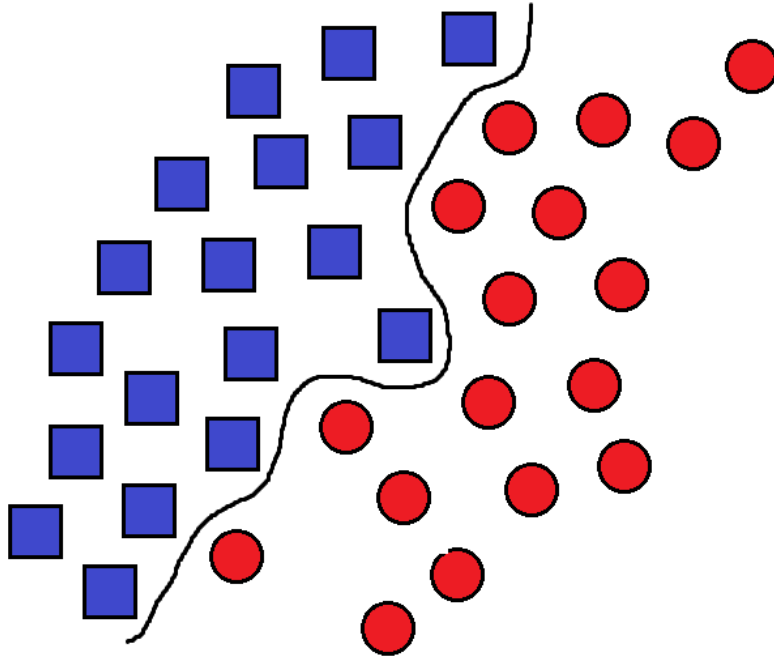
КЛАССИЧЕСКИЕ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ

- Задача кластеризации
- Задача классификации
- Задача регрессии
- Задача понижения размерности данных

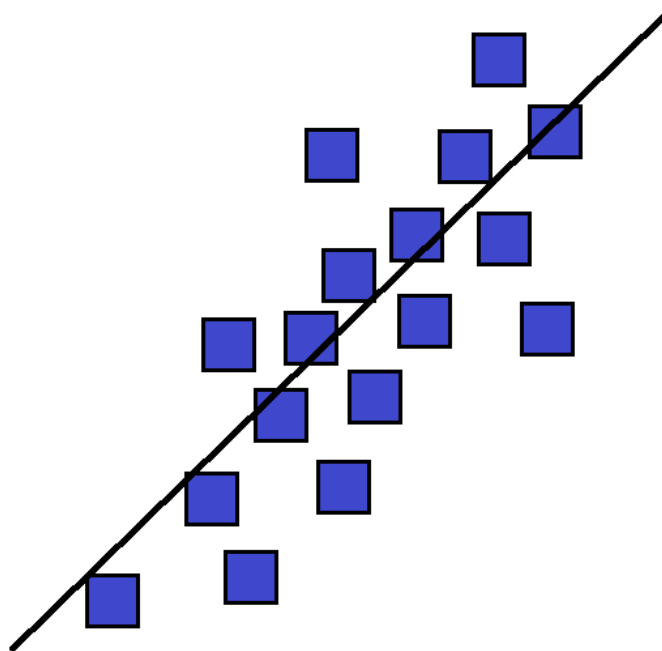
КЛАСТЕРИЗАЦИЯ



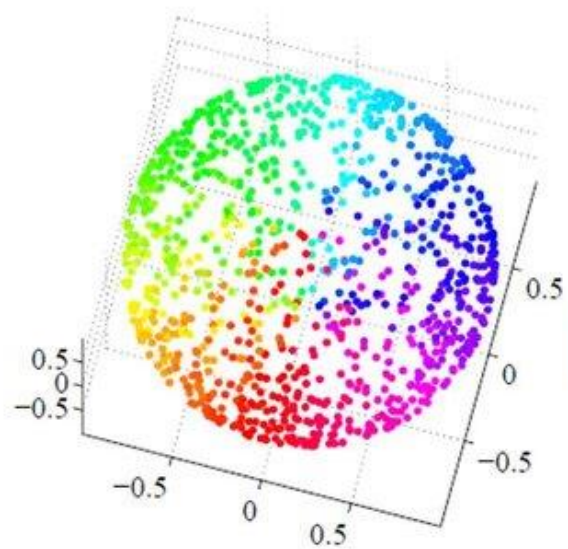
КЛАССИФИКАЦИЯ



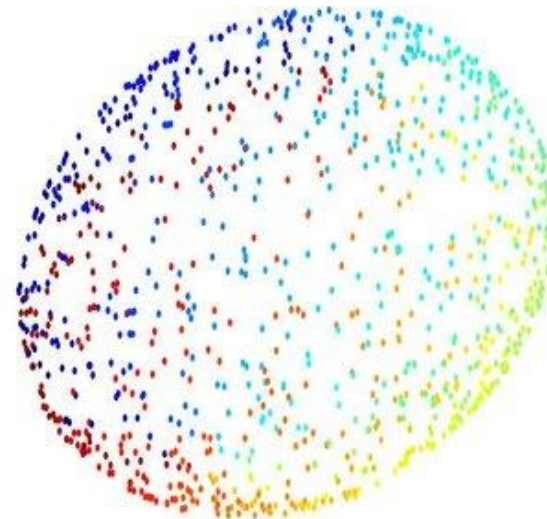
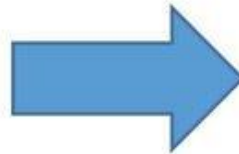
РЕГРЕССИЯ



ПОНИЖЕНИЕ РАЗМЕРНОСТИ ДАННЫХ



3D

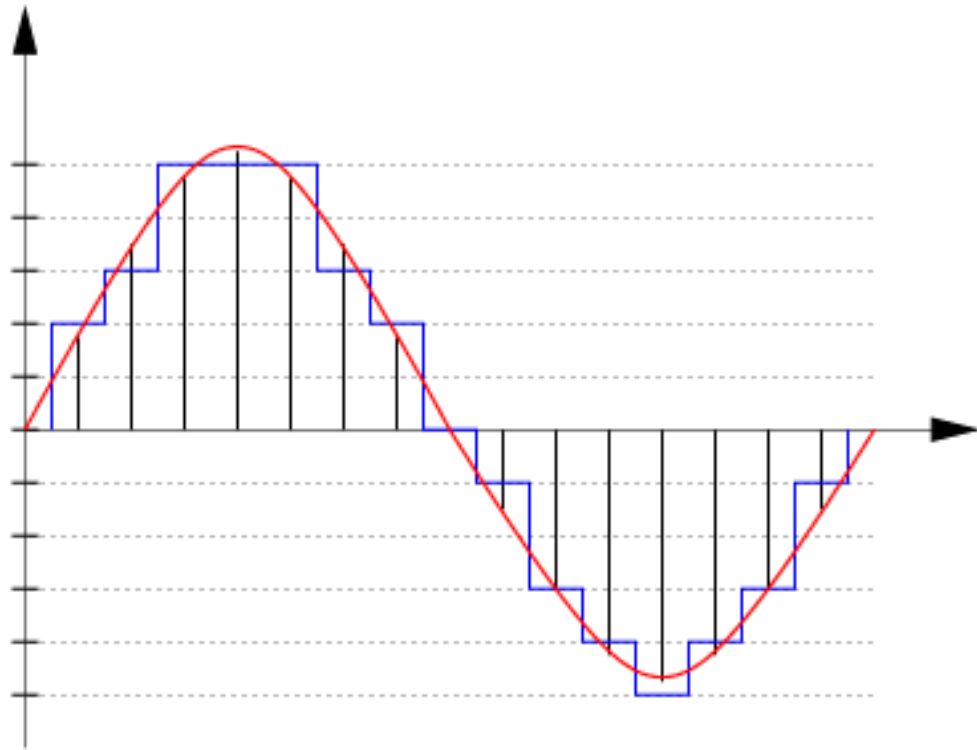


2D

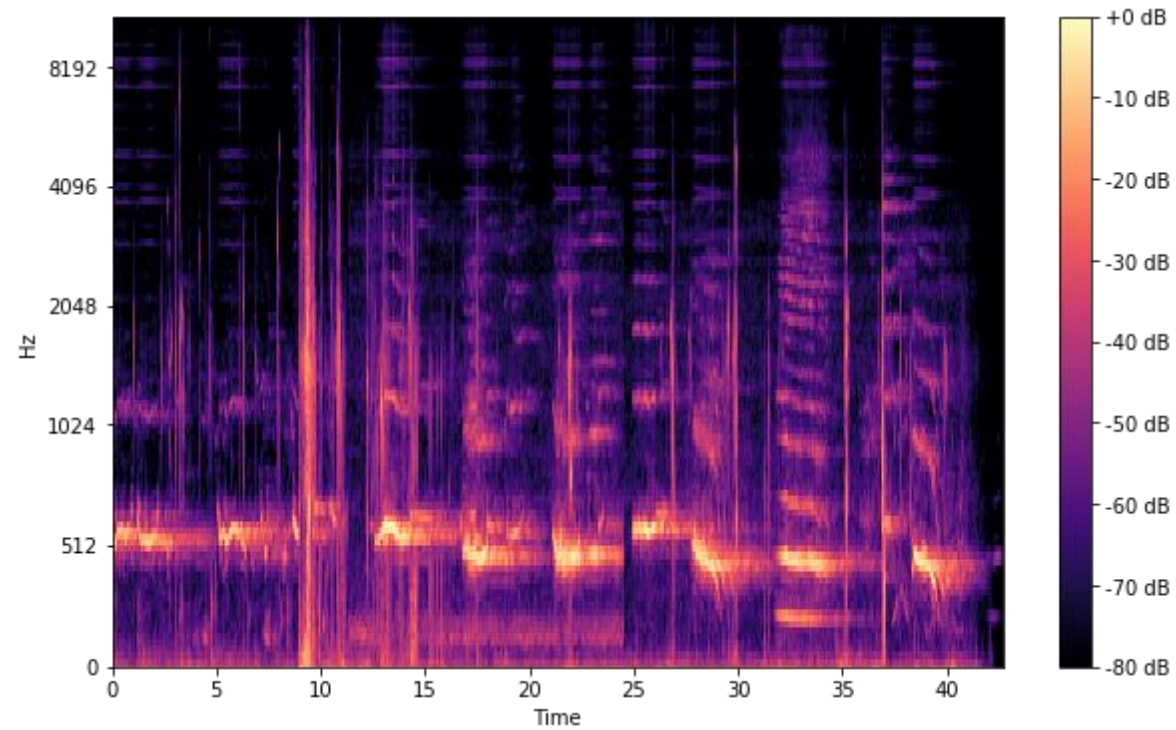
СПОСОБЫ МАШИННОГО ОБУЧЕНИЯ

- Обучение с учителем (с обучающей выборкой)
- Обучение без учителя (спонтанное обучение)
- Обучение с подкреплением (“агент – среда”)

ДИСКРЕТИЗАЦИЯ ЗВУКОВОГО СИГНАЛА



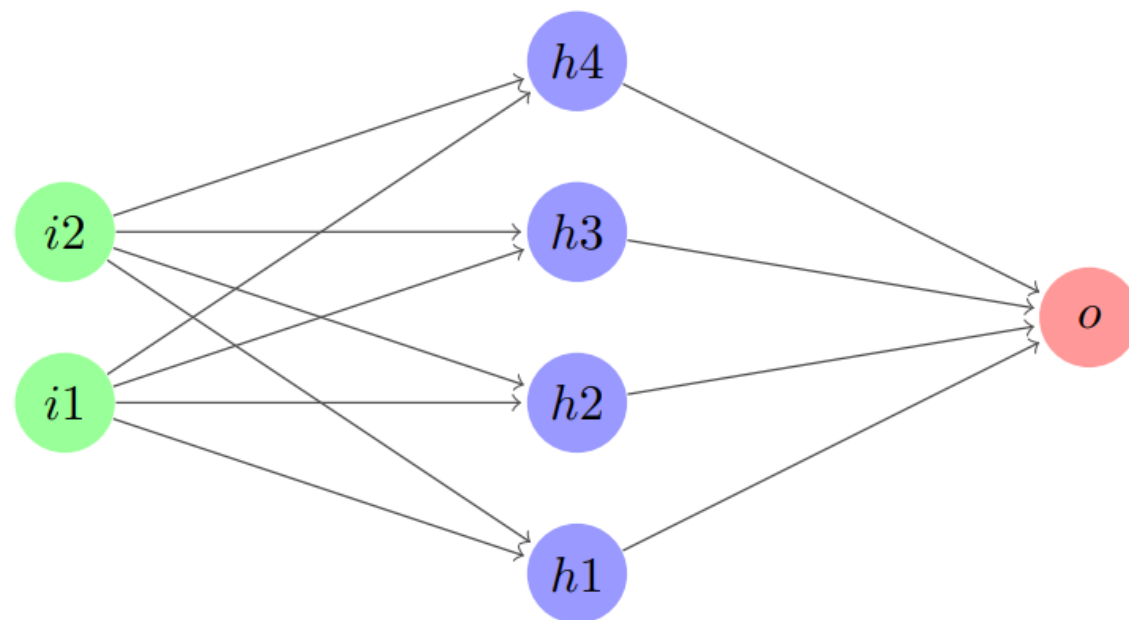
ПРЕДСТАВЛЕНИЕ ЗВУКОВОГО СИГНАЛА В ВИДЕ СПЕКТРОГРАММЫ



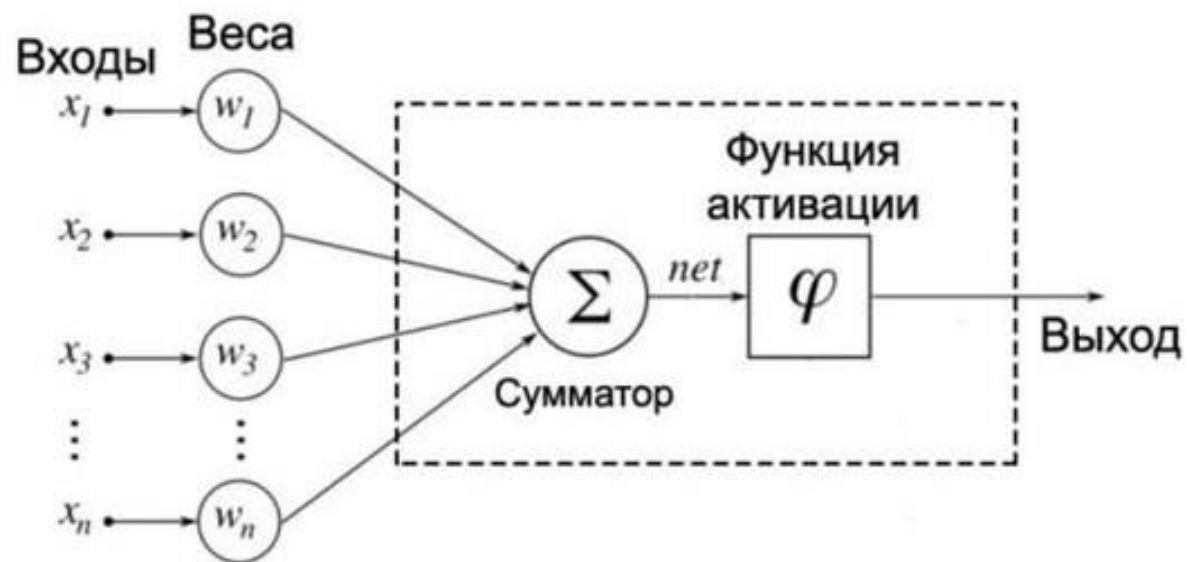
ОСНОВНЫЕ ЗАДАЧИ

- Сжатие данных
- Фильтрация
- Обработка звука
- Распознавание речи

ИСКУССТВЕННАЯ НЕЙРОННАЯ СЕТЬ



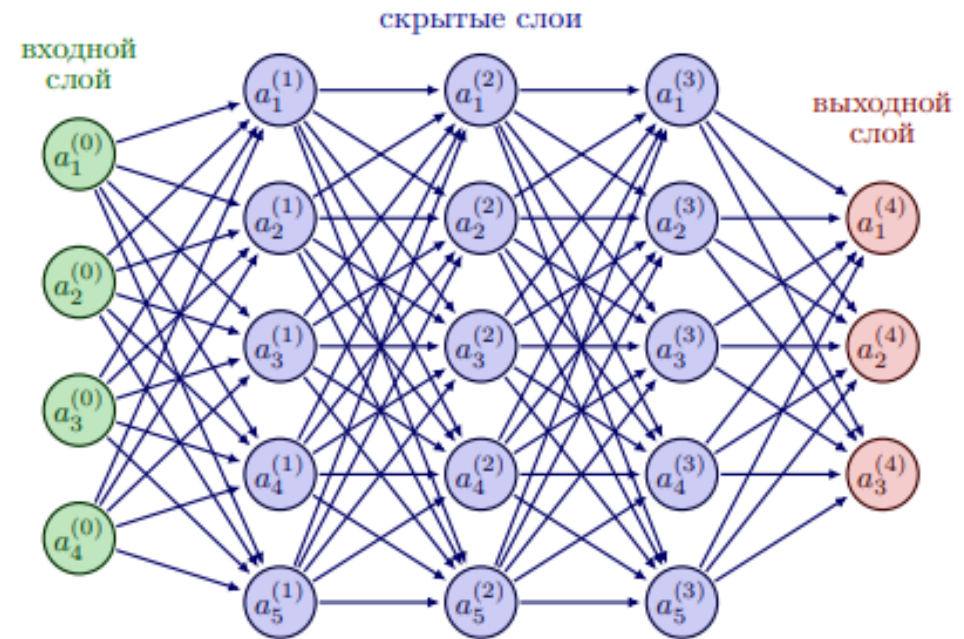
ИСКУССТВЕННЫЙ НЕЙРОН



СЕТИ ПРЯМОГО РАСПРОСТРАНЕНИЯ

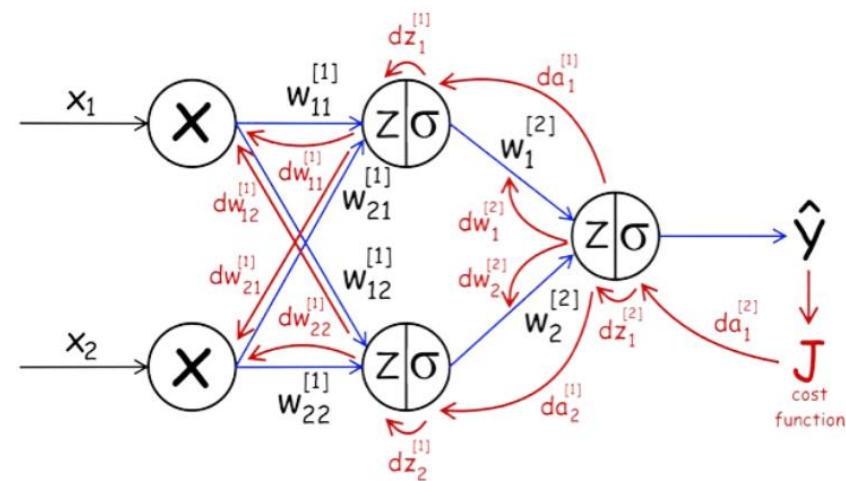
Сеть прямого распространения – это нейронная сеть, сигналы в которой распространяются строго от входа к выходу.

Как правило, под сетями прямого распространения подразумевают полносвязные многослойные нейронные сети с прямым распространением сигнала.



МЕТОД ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

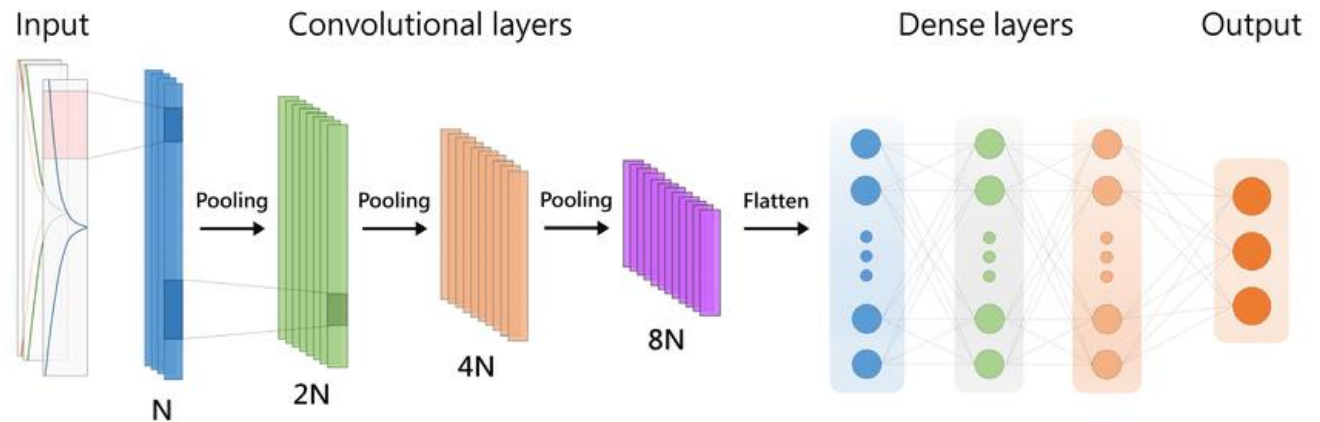
На каждой итерации обучения происходит два прохода сети: прямой и обратный. На прямом проходе входной вектор сигналов распространяется от входов сети к её выходам и формирует некоторый выходной вектор, соответствующий текущему набору весов сети. После того, как был вычислен выходной вектор, вычисляется выходная ошибка нейронной сети. На обратном проходе эта выходная ошибка распространяется от выходов сети к её входам, после чего производится коррекция весов сети.



СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Сверточная нейронная сеть — это специальная архитектура нейронных сетей, нацеленная на эффективное распознавание образов.

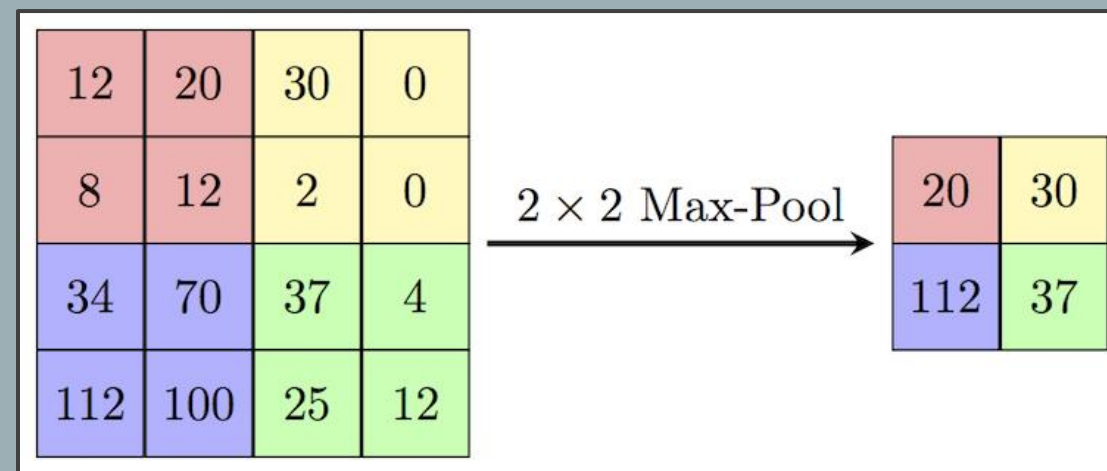
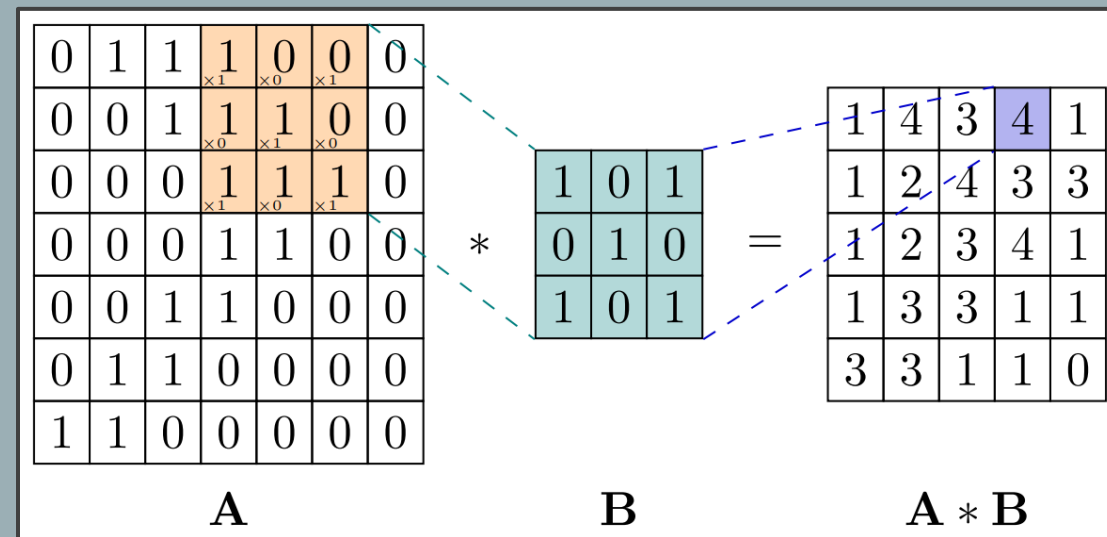
Основными видами слоев в сверточных нейронных сетях являются сверточные (convolutional), пулинговые (pooling) и полносвязные (dense) слои.



СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Свертка – бинарная операция над матрицами, суть которой заключается в том, что каждый фрагмент первой матрицы поэлементно умножается на вторую матрицу, результаты суммируются и записываются в соответствующую позицию итоговой матрицы.

Операция пулинга – операция, выполняющая уменьшение размерности сформированных под действием операции свертки карт признаков. Исходная матрица делится на блоки размером $w \times h$, для каждого из которых вычисляется некоторая функция.



ПРЕИМУЩЕСТВА СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

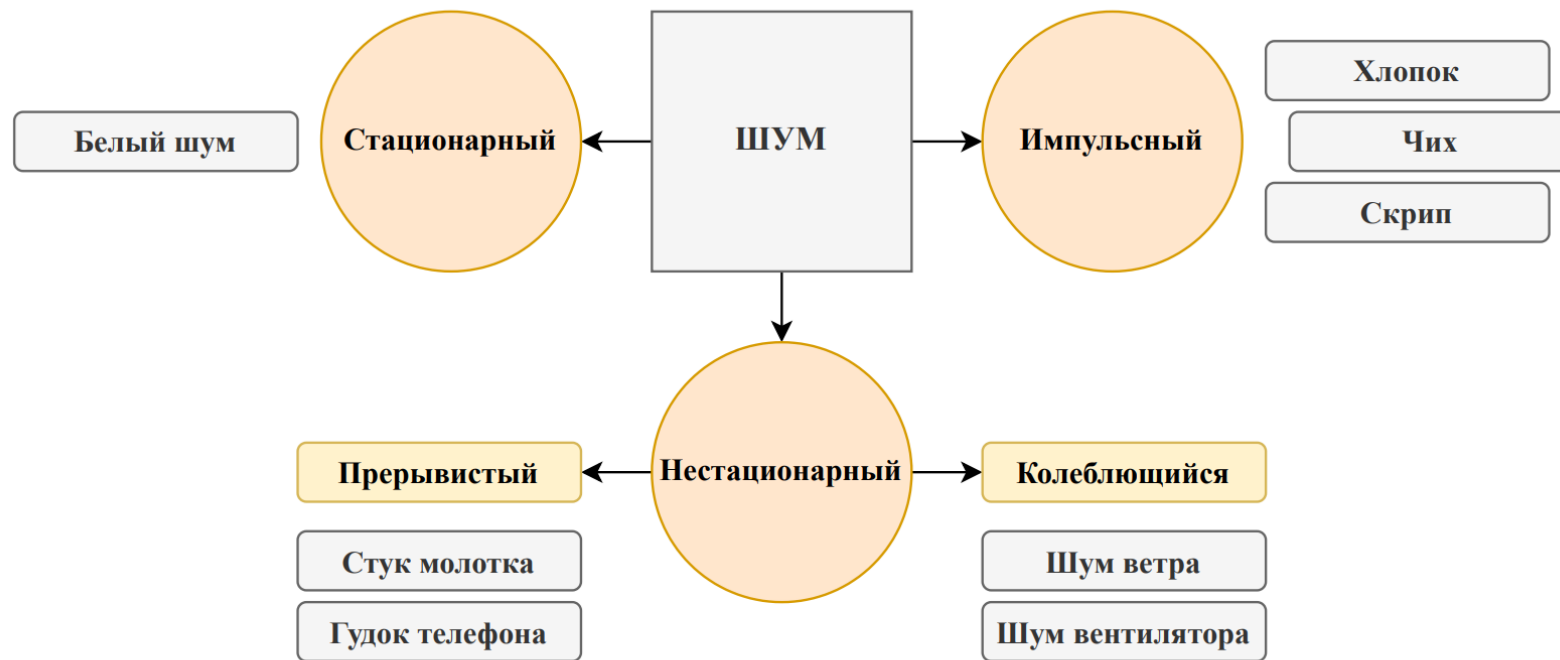
- Сверточные нейронные сети более эффективно распознают и классифицируют изображения.
- По сравнению с полносвязными нейронными сетями сверточные нейронные сети имеют гораздо меньшее количество настраиваемых весов, что связано с наличием сверточных слоев. Такой подход подталкивает сеть к обобщению демонстрируемой информации и поиску закономерностей, а не к попиксельному запоминанию обрабатываемых изображений.
- Удобное распараллеливание вычислений дает возможность повысить скорость выполнения алгоритмов работы и обучения сети.
- Устойчивость к линейным преобразованиям распознаваемого изображения (к масштабированию, повороту, сдвигу).

ШУМОПОДАВЛЕНИЕ

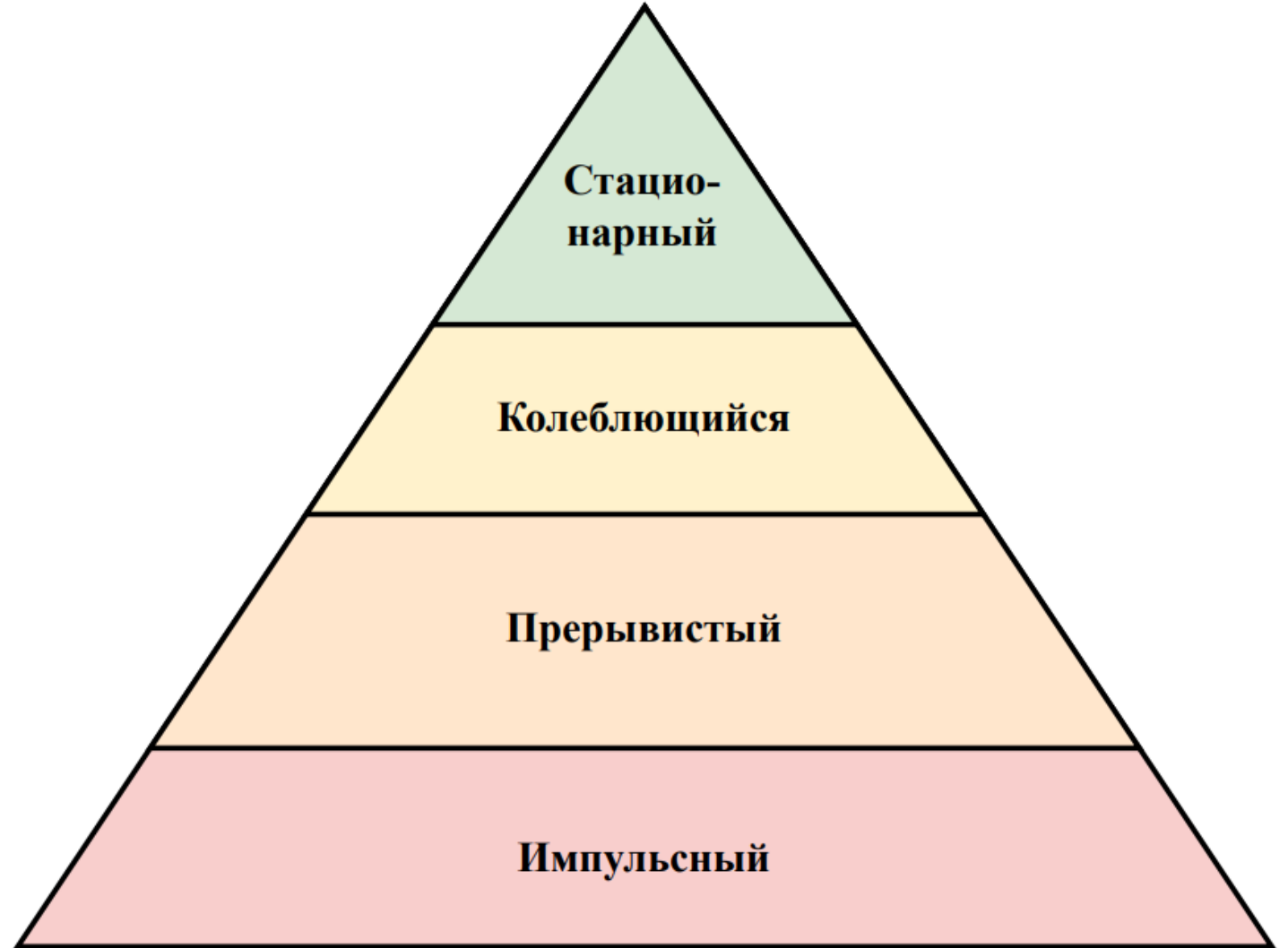
ОСНОВНЫЕ КЛАССИФИКАЦИИ ШУМОВ

- по спектру
- по характеру спектра
- по происхождению
- по частоте
- по временным характеристикам

ВРЕМЕННАЯ КЛАССИФИКАЦИЯ ШУМОВ



ПИРАМИДА
СЛОЖНОСТИ
ШУМОПОДАВЛЕНИЯ

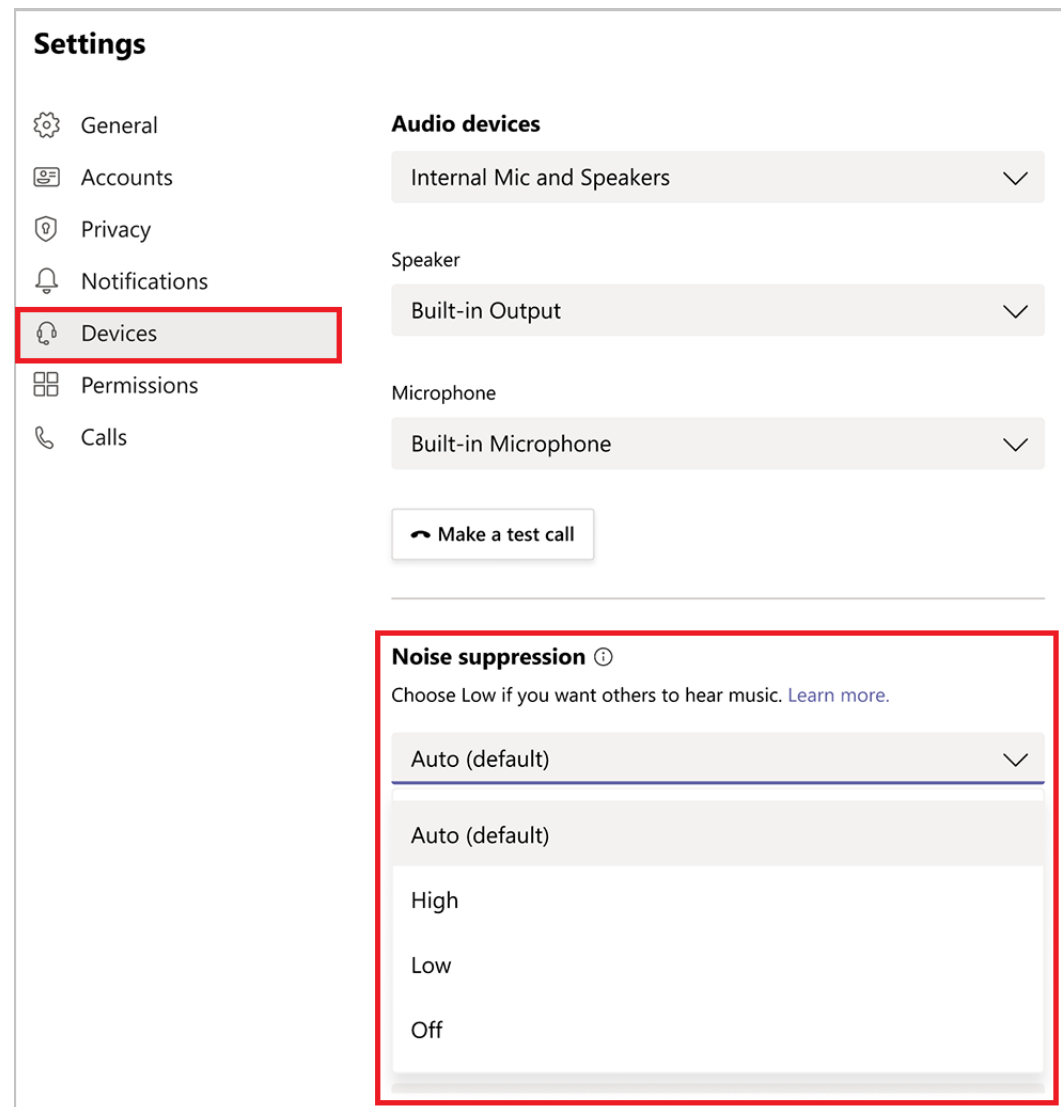


ИСПОЛЬЗУЕМЫЕ МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Наиболее эффективным методом машинного обучения в шумоподавлении и улучшении речевого сигнала является использование сверточных нейронных сетей. Общий принцип нейросетевых архитектур обработки звука и речи: наличие сверточного энкодера и декодера, выполняющих определение источников шумов и их восстановления с последующим вычитанием из исходного сигнала. Среди архитектур сверточных нейронных сетей, используемых в шумоподавлении, можно выделить WaveNet, Tas-Net, а также её модификацию Conv-Tas-Net.

ПРИМЕНЕНИЕ ШУМОПОДАВЛЕНИЯ

Методы шумоподавления используются при очистке аудио от лишних звуковых сигналов для повторного воспроизведения. Более сложной задачей является мгновенное шумоподавление – шумоподавление и воспроизведение одновременно с записью речи. Чаще всего такое шумоподавление используется для аудиоконференций в Zoom, Discord, Skype, а также в таких мессенджерах как Telegram и VK.



Функция активного шумоподавление в Microsoft Teams

СПАСИБО ЗА ВНИМАНИЕ!