

# Распознавание речи в текст

Это способность компьютерного программного обеспечения идентифицировать слова и фразы в устной речи и преобразовывать их в текст, читаемый человеком.



# Как речь распознается в текст?



Преобразование речи в текст  
осуществляется с помощью сложной модели  
машинного обучения, состоящей из  
нескольких шагов



## *Распознавание речи вчера*

Акустическая модель-языковая модель-модель  
произношения

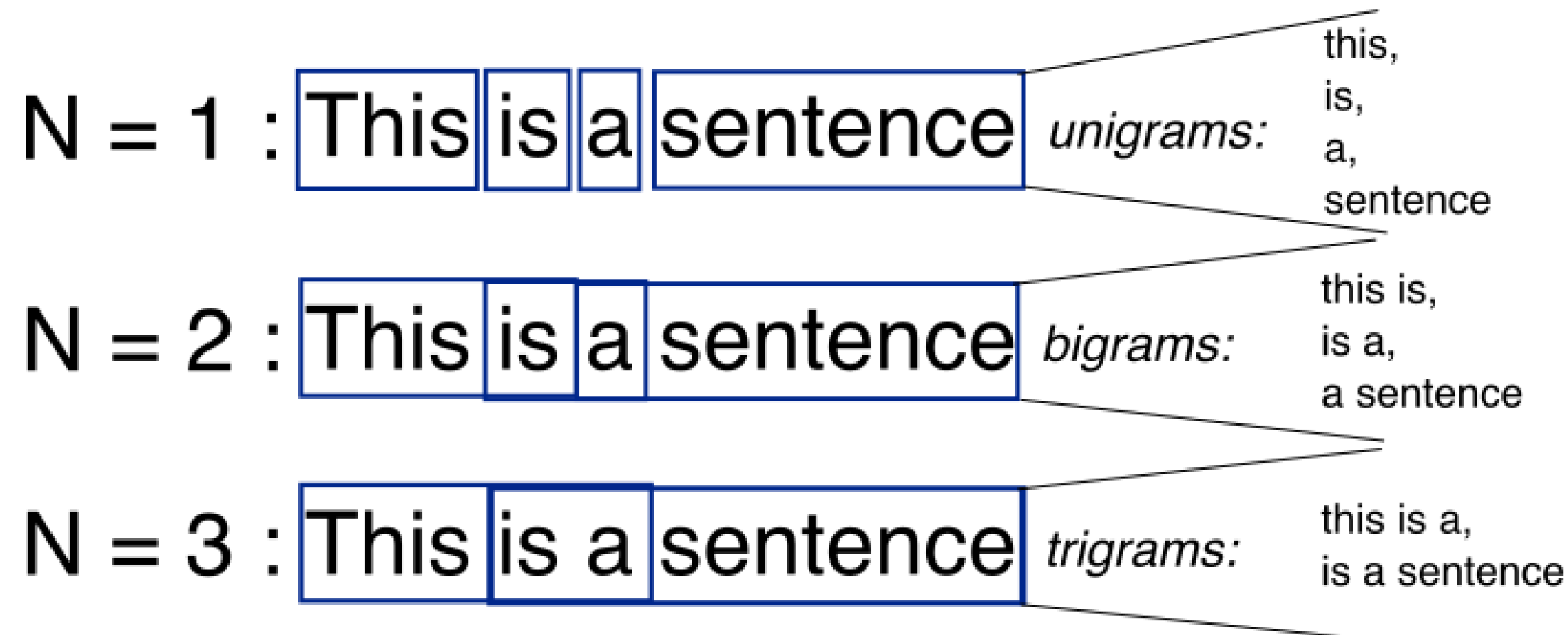
Алгоритмы:

1. n-грамм

2. НММ(скрытая модель  
Маркова)



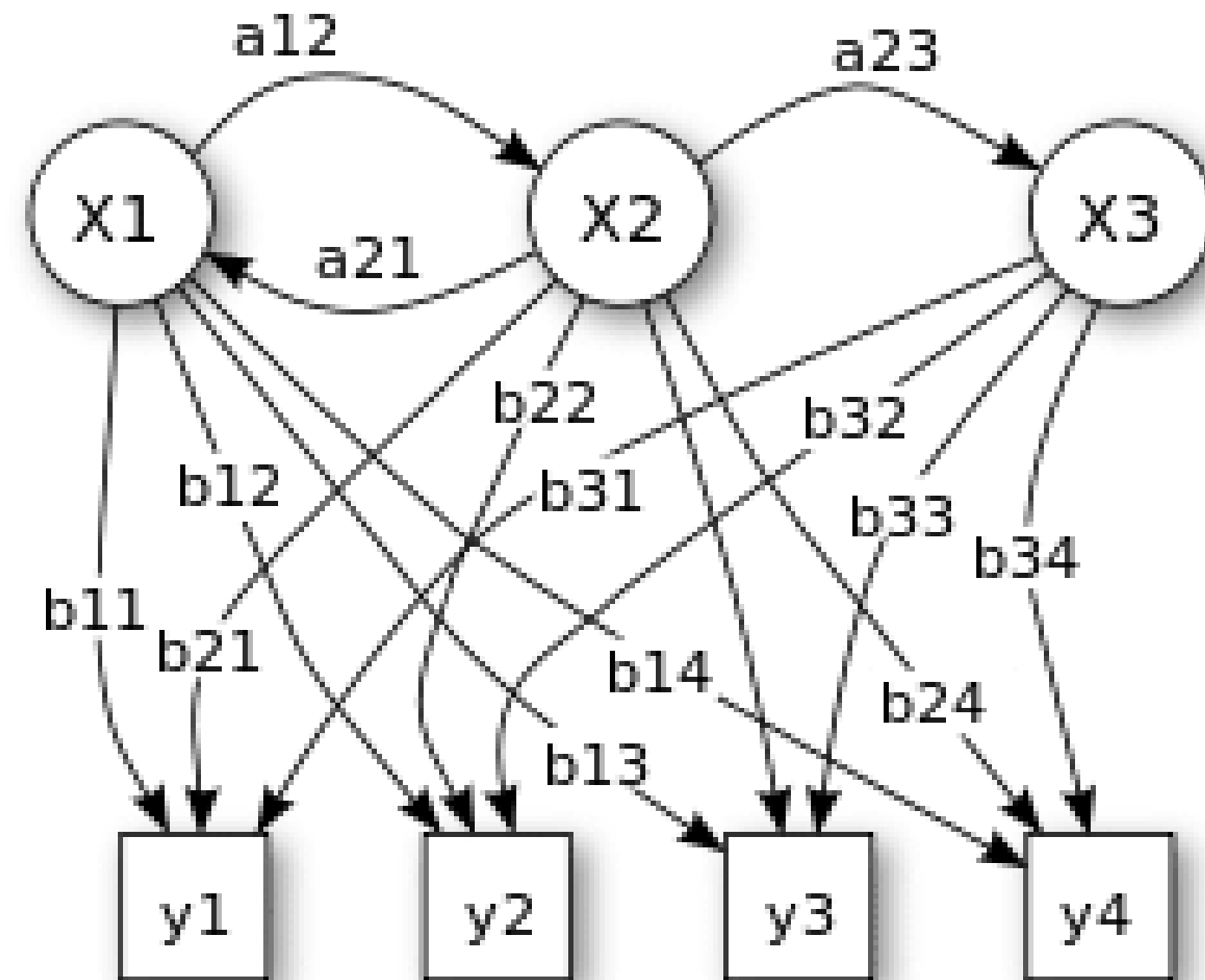
# n-грам



Используют предыдущие n слов в качестве контекста, чтобы попытаться определить данное слово.

# Hidden Markov Model

HMM



# Распознавание речи сегодня

это искусственная  
нейронная сеть



**Как это работает?**

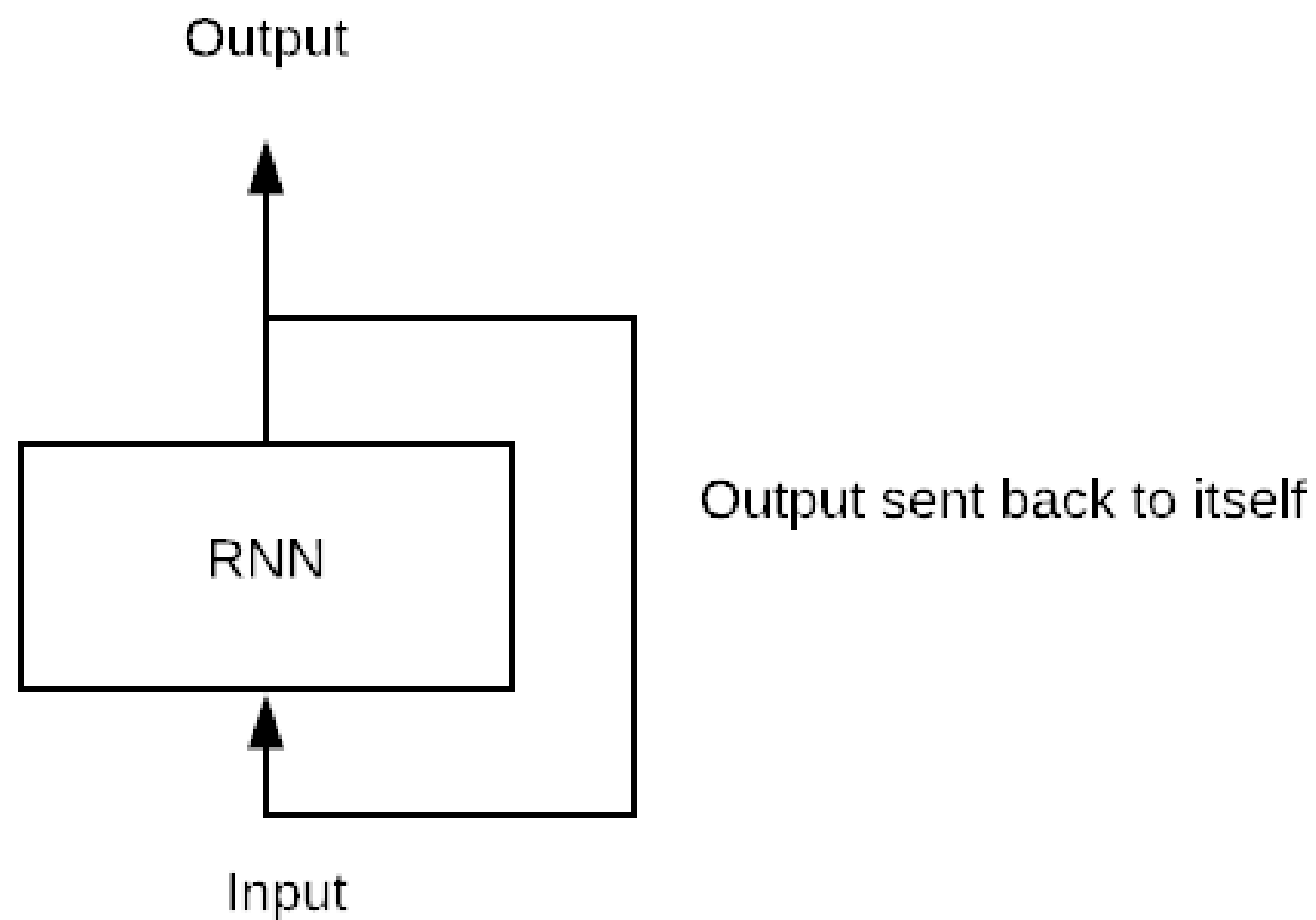




# Алгоритм RNN



# Рекуррентные нейронные сети



**Идеально подходит для последовательных данных, таких как речь, потому что этот алгоритм способен «запоминать» то, что было раньше, и использовать свой предыдущий вывод в качестве входных данных для своего следующего хода.**

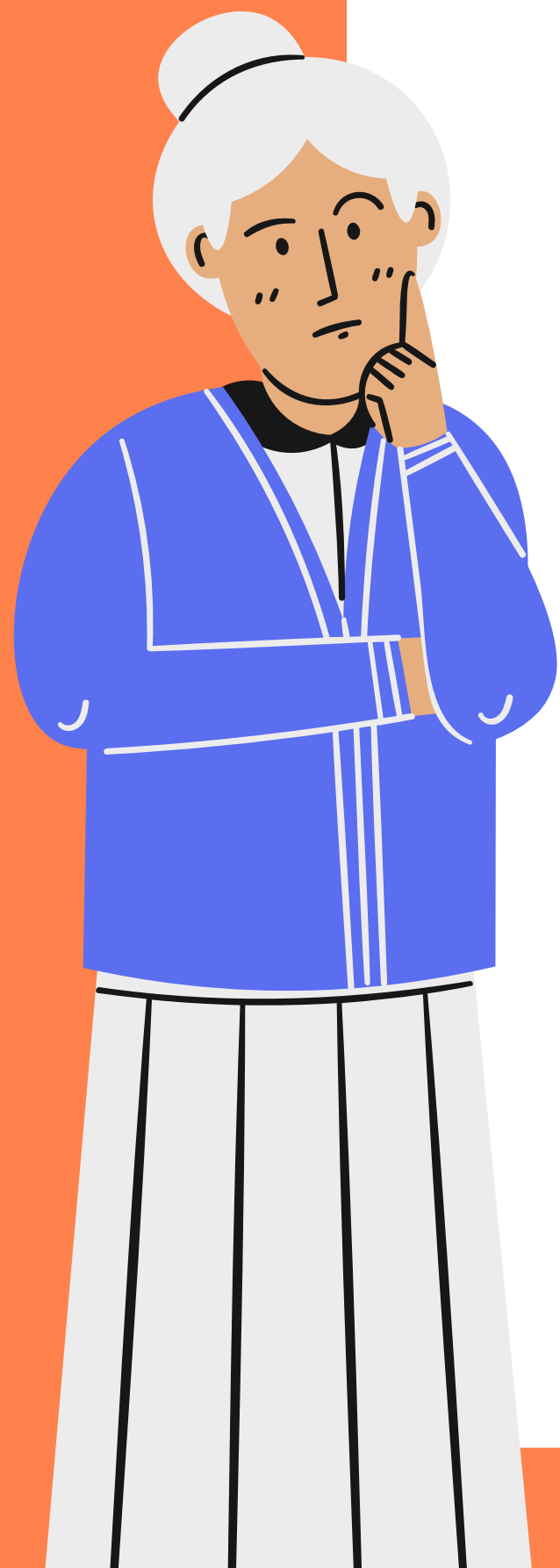




Алгоритм СТС  
(коннекционистская  
временная  
классификация)

# Области применения





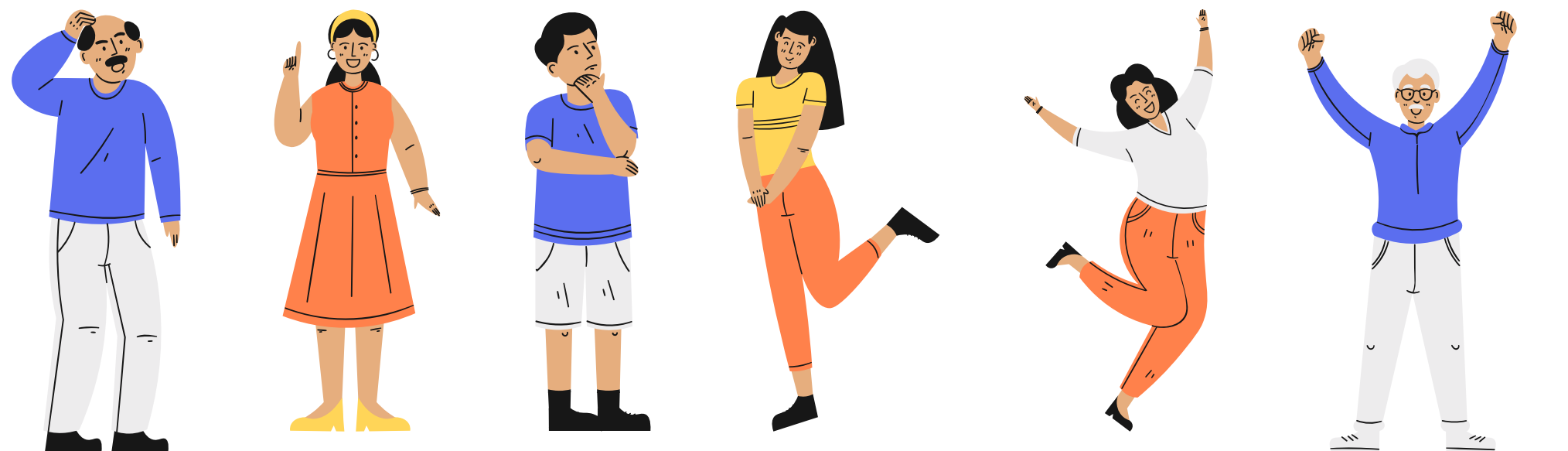


# Горячие тренды

в распознавании  
речи и голоса



# Мобильные платежи с распознаванием голоса



Вы когда-нибудь догадывались, что вместо того, чтобы использовать наличные деньги или передавать свою кредитную карту, вы можете однажды использовать свой голос для совершения платежей? Этот день может наступить очень скоро!



## Помощники ИИ, которые распознают ваш конкретный голос

Когда ваш ИИ-помощник запускается только вашим конкретным голосом, гораздо меньше вероятность того, что неавторизованные пользователи будут вмешиваться в ваши устройства, данные и вещи. Помощник просто не активируется для тех, кто не авторизован.



Hey Siri





# Голосовые покупки



Покупки, управляемые голосом, - это захватывающая мода 2021 года. Его влияние на электронную коммерцию будет заметным.

Amazon - одна из самых важных компаний в этой отрасли. Amazon применил Alexa и свою линейку умных аудиосистем для развития голосового управления покупками.







# Распознавание речи в Python





## Пакеты распознавания речи

SpeechRecognition	В основном библиотека-оболочка для некоторых облачных API и модель, основанная на CMUSphinx, исследовании, проведенном в Университете Карнеги-Меллона
watson-developer-cloud	Это API-интерфейс искусственного интеллекта, который упрощает создание, отладку, запуск и развертывание API-интерфейсов. Его можно использовать для выполнения основных задач распознавания речи.
google-cloud-speech	Offers basic speech to text conversion
apiai	Включает обработку естественного языка для определения намерений говорящего
assemblyai	Создает быстрый API для автоматического распознавания речи, который разработчики могут использовать бесплатно
wav2letter	Библиотека с открытым исходным кодом, которая использует в своих моделях чистую архитектуру CNN и изначально была создана Facebook
Kaldi	Это набор инструментов для распознавания речи, предназначенный для использования исследователями и профессионалами в области распознавания речи
CMU Sphinx	Разработаны специально для малоресурсных платформ, гибкого дизайна и ориентированы на практическую разработку приложений, а не на исследования



# **Использование SpeechRecognition( далее код в Jupyter)**



# Спасибо, что послушали!

Студент: Тажибай Айжан