# 

# 

# **Реферат**

Пояснительная записка содержит m страниц, n рисунков, k таблиц.

Количество использованных источников p.

Ключевые слова: пространство эмоций, базовые эмоции, PAD, FACS, action units, виртуальные ассистенты, линейная регрессия и Ridge/Lasso регрессия, Random Forest, KNN, нейронные сети, машинное обучение.

Целью данной работы является продолжение разработки алгоритмов и программных средств, предназначенных для отображения между пространствами эмоций различных размерностей, в виде обучения ML моделей и демонстрации их результатов предсказаний.

В первом разделе рассматриваются вопросы представления эмоций в виде формальных моделей, описана постановка задач работы.

Во втором разделе изучаются существующие решения поставленных задач, анализируются возможности применения различных моделей машинного обучения.

В третьем разделе проводится углубленный анализ данных, а также выбран метод поиска оптимальных значений ключевых параметров модели.

В четвёртом разделе происходит подбор инструментов для обучения моделей, обучение моделей, сравнение их работы с помощью метрик и отбор лучших моделей.

В заключении подводятся итоги проведённой работы с кратким описанием результатов.

**СОДЕРЖАНИЕ**

[**Реферат 2**](#_30j0zll)

[**Введение 5**](#_3dy6vkm)

[**Раздел 1. Аналитическая часть. 7**](#_4d34og8)

[1.1 Технологии и подходы для реализации виртуальных ассистентов. 7](#_2s8eyo1)

[1.2 Изучение и анализ литературы на тему представления эмоций в виде формальных моделей. 9](#_17dp8vu)

[1.2.1 Общие характеристики дискретных и многомерных моделей. 9](#_3rdcrjn)

[1.2.2 Модель эмоционального состояния PAD. 12](#_26in1rg)

[1.2.3 Система кодирования лицевых движений FACS. 13](#_lnxbz9)

[1.3 Постановка задач, исследовательские вопросы и гипотезы. 15](#_35nkun2)

[1.3.1 Постановка задач. 15](#_1ksv4uv)

[1.3.2 Общая информация об исходных данных. 16](#_44sinio)

[1.3.3 Гипотезы о данных и моделях. 16](#_2jxsxqh)

[**Раздел 2. Теоретическая часть 18**](#_1y810tw)

[2.1 Изучение существующих решений поставленных задач. 18](#_4i7ojhp)

[2.1.1 FaceReader. 18](#_2xcytpi)

[2.2 Математическая постановка задач. 19](#_1ci93xb)

[2.2.1 Углубленный анализ данных. 20](#_3whwml4)

[2.2.2 Обучение моделей машинного обучения. 21](#_3as4poj)

[2.3 Сравнение моделей машинного обучения. 22](#_1pxezwc)

[2.3.1 Линейная Регрессия. 23](#_49x2ik5)

[2.3.2 Полиномиальная Регрессия. 23](#_2p2csry)

[2.3.3 Ridge и Lasso Регрессии. 23](#_147n2zr)

[2.3.4 Случайный Лес (Random Forest). 24](#_3o7alnk)

[2.3.5 K-ближайших Соседей (KNN). 25](#_23ckvvd)

[2.3.6 Многослойный Перцептрон (MLPRegressor). 25](#_ihv636)

[**Раздел 3. Инженерная часть 26**](#_1hmsyys)

[3.1 Выбор датасета, анализ данных, формирование обучающей и тестовой выборок. 26](#_41mghml)

[3.1.1 Выбор датасета. 26](#_5ne5jyjr5mc9)

[3.1.2 Углубленный анализ данных. 26](#_vx1227)

[3.1.3 Формирование обучающей и тестовой выборок. 26](#_qivmcsemupri)

[3.2 Метод поиска оптимальных значений ключевых параметров моделей. 26](#_3tbugp1)

[3.3 Формирование метрик для оценки моделей. 26](#_28h4qwu)

[**Раздел 4. Технологическая и практическая часть 26**](#_37m2jsg)

[4.1 Выбор инструментов для анализа данных и обучения моделей. 26](#_1mrcu09)

[4.1.1 Инструменты для анализа и обработки данных. 27](#_46r0co2)

[4.1.2 Инструменты для обучения моделей. 27](#_2lwamvv)

[4.1.3 Инструменты для визуализации. 27](#_111kx3o)

[4.1.4 Инструменты для сохранения моделей. 27](#_3l18frh)

[4.2 Обучение моделей. 27](#_206ipza)

[4.3 Оценка точности работы моделей. 27](#_4k668n3)

[4.3.1 Визуализация работы обученных моделей. 27](#_ihiom9jmve4e)

[4.3.2 Выводы по таблицам результатов обучения моделей. 27](#_jjx922oiqry3)

[4.4 Сохранение лучших моделей в формате файла обученных моделей. 27](#_3ygebqi)

[**Заключение 27**](#_sqyw64)

[**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 27**](#_3cqmetx)

# 

# **Введение**

Существуют различные подходы к описанию эмоций человека. Одни модели включают в себя фиксированное количество базисных (фундаментальных) эмоций, другие используют ограниченное число переменных (осей), которые задают эмоциональную точку в пространстве. Первые модели называются дискретными, а вторые – многомерными [1].

Исследователи, разделяющие дискретную теорию эмоций, расходятся во мнениях относительно количества и содержания фундаментальных эмоций. Сторонники этого подхода сосредоточены на изучении внутренней организации индивидуальных эмоций, а не на их взаимосвязи. Наиболее известным психологом в этой области считается Пол Экман, выделяющий 6 базовых эмоций — гнев, удивление, страх, радость, грусть и отвращение. В рамках данной работы мы будем считать, что базовых эмоций всего 6 [1].

Многомерные модели эмоциональных состояний различаются между собой количеством измерений (осей) и их смысловой нагрузкой. В данном исследовании мы будем рассматривать многомерную модель PAD [2], в которой всего 3 измерения: P – Pleasure (удовольствие), A – Arousal (возбуждение), D – Dominance (доминирование).

Полом Экманом в XX веке была разработана система кодирования лицевых движений (FACS) [3]. Эта система позволяет классифицировать выражения лица человека на основе его мимики. С использованием FACS можно закодировать практически любое, анатомически возможное выражение лица, деконструируя его на вызвавшие его конкретные двигательные единицы (action units) и их временные сегменты [4].

С помощью двигательных единиц из FACS Полом Экманом и Уоллесом Фризеном была создана эмоциональная система кодирования лицевых движений (EmFACS), которая рассматривает только лицевые движения, связанные с эмоциями. Стоит отметить, что в научном мире нет единого мнения на тему однозначного соответствия между мимикой и эмоциями для всех людей [5]. Таким образом нельзя с полной уверенностью сказать, что, например, выражение лица любого человека, который испытывает страх выглядит именно так, а никак иначе. Однако, большинство людей в современном обществе выражает эмоции приблизительно одинаково, что даёт возможность использовать труды Пола Экмана в данной работе.

С учётом вышесказанного, встаёт вопрос о необходимости иметь возможность отображать пространства эмоций различных размерностей друг в друга.

# **Раздел 1. Аналитическая часть.**

## **1.1 Технологии и подходы для реализации виртуальных ассистентов.**

Виртуальный ассистент [6] (англ. Virtual assistant) — программный агент, который может выполнять задачи (или сервисы) для пользователя на основе информации, введенной пользователем, данных о его местонахождении, а также информации, полученной из различных интернет-ресурсов (погода, уличное движение, новости, курсы валют и ценных бумаг, розничные цены в магазинах и т. д.). Примерами таких агентов являются программы Siri, Google Assistant (Google Now), Amazon Alexa, Microsoft Cortana, Bixby, Voice Mate, Алиса, Маруся, Дуся, Салют и другие.

Искусственный интеллект уже стал частью нашей повседневной жизни. В каждом мобильном телефоне есть умный голосовой помощник, который расскажет о погоде, подскажет, где ближайший банкомат, вызовет такси или закажет билеты в кино. Но это только один из примеров виртуального помощника, самый известный и знакомый. На самом деле у таких решений гораздо больше областей применения.

Так что же такое виртуальный помощник (ассистент)? В общем смысле это сервис (программа, мобильное приложение) — выполняющий действия в ответ на информацию, введенную пользователем и/или найденную в Сети. Кроме того, искусственный интеллект избавляет человека от необходимости самостоятельно выполнять некоторые операции – как правило, простые, то есть такие, которые под силу роботу.

Возможности виртуальных помощников используются во всех сферах жизни. В автомобили внедряются голосовые помощники, чтобы владелец мог без помощи рук открывать и закрывать двери, запускать двигатель, узнавать уровень топлива в бензобаке. Роботы управляют устройствами умного дома, их устанавливают в гостиницах, например, чтобы гости могли заказывать услуги в номер с помощью простых команд.

Основная функция голосовых виртуальных ассистентов – имитировать реальное общение. Уровень развития этой функции у цифровых помощников разный. Некоторые роботы способны поддерживать только самые простые диалоги, читая предварительно записанные звуковые сообщения. Другие неотличимы от человека: понимают смысл обращенной к ним речи, знают синонимы, отвечают без промедления, замолкают, когда их перебивают, и даже употребляют слова-паразиты. В результате собеседник даже не всегда может определить, что ведет диалог с роботом. Понятно, что такие возможности могут быть очень полезны для бизнеса.

Кроме коммуникационных функций, виртуальные ассистенты (голосовые боты) могут выполнять различные действия по заданным сценариям: планировать маршрутизацию звонков, отправлять SMS, записывать результаты вызовов в виде отчетов, использовать рабочие данные по информационным ресурсам компании и даже фиксировать изменения в них – при условии, что робот интегрирован с нужными системами.

Интернет-магазины с помощью голосовых ботов сообщают об акциях и специальных предложениях, регистрируют клиентов в программе лояльности, активируют корзину, увеличивают продажи, помогают клиентам с заказом.

Виртуальный помощник в салоне красоты или другой организации из сферы услуг напомнит клиенту о приеме, при необходимости перенесет время визита, в результате чего показатель «рентабельности» повысится.

Службы доставки посредством цифровых помощников информируют получателей о статусе заказа, согласовывают доставку.

Банки, страховые компании, операторы связи и интернет-провайдеры используют виртуальных помощников, чтобы оповещать о новых услугах и тарифах, напоминать клиентам о задолженности и облегчить совершение платежей.

И это не все области, где используются виртуальные помощники, и не все решаемые ими задачи. Практически в любой компании виртуальный ассистент может выполнять функции технической поддержки: отвечать на частые вопросы клиентов, консультировать по несложным задачам, перенаправлять звонок оператору. Кроме того, его можно использовать для проведения опросов и анкетирования, например, чтобы узнать мнение клиентов о новом продукте. Виртуальные помощники также помогают во внутрикорпоративных коммуникациях: решают текущие вопросы сотрудников, участвуют во взаимодействии отделов.

Ассистент может быть оптимизирован для своего рабочего канала. Как правило, виртуальные помощники используются в голосовых каналах (на телефоне) и в интернет-каналах (на веб-сайтах). Первый тип относится в основном к голосовым помощникам, ко второму – чат-ботам. У них разные сильные и слабые стороны, которые необходимо учитывать в реальных проектах. Хороший результат для голосового канала может не подойти для интернет-чата и наоборот.

Указанные выше ассистенты (голосовые помощники и чат-боты) смогли быть реализованы в связи с решением главных задач, лежащих в основе этих ассистентов. Например, внутри голосового помощника заложены алгоритмы распознавания и синтеза речи. Впервые эти задачи были реализованы в коммерческих программах распознавания речи в 90-х годах [7][8]. Чат-боты же основаны на заранее определенных бизнес-правилах, определяющих ход диалога между пользователем и ботом. Как правило, чат-боты избегают вопросов, требующих свободного ответа, а ставят их в виде «тестового» варианта ответа (с конечным количеством вариантов). Эта технология построения интеллектуальных систем хорошо изучена и уже нашла широкое применение на практике [9].

Также, перспективной ветвью развития виртуальных ассистентов могут стать помощники, использующие цифровых аватаров [10] – виртуальных персонажей, имеющих внешний облик, схожий с человеком. Данный вид ассистентов находится сейчас на ранних этапах своего развития, т.к. часть задач, лежащих в их основе не имеет достаточной фундаментальной базы. Так, задача распознавания лиц уже имеет коммерческие решения на данный момент [11], но она является лишь частью необходимой базы для создания виртуальных ассистентов с цифровыми аватарами. Продолжаются работы по синтезу выражений лиц для виртуальных персонажей, т.к. в качественном ассистенте, мимика и эмоции аватаров должны быть близки к реальным для людей, что на данный момент далеко от желаемого результата.

## **1.2 Изучение и анализ литературы на тему представления эмоций в виде формальных моделей.**

Много лет учёные всего мира работали над задачей распознавания мимики человека. В наши дни уже есть успехи в этом направлении. Если раньше для разложения выражения лица на составляющие в FACS требовался специально обученный человек, то теперь нейросети могут делать это автоматически.

### **1.2.1 Общие характеристики дискретных и многомерных моделей.**

Есть два принципиально разных подхода к изучению организации эмоций. С точки зрения дискретной модели эмоциональная сфера состоит из определенного количества первичных, базовых или фундаментальных эмоций. Комбинации первичных эмоций - это эмоции второго порядка. Статус первичных эмоций обычно приписывается гневу, страху, грусти, счастью и т.д. Сторонники этого подхода сосредоточены на изучении внутренней организации индивидуальных эмоций, не уделяется внимание взаимосвязям между эмоциями.

В многомерной модели акцент делается на многомерном пространстве эмоций, образованном ограниченным числом шкал (осей), например, негативная - позитивная, сильная - слабая, активная - пассивная. Такие модели ориентированы на изучение сходства и различий между эмоциями.

Исследователи, разделяющие дискретную теорию эмоций, расходятся во мнениях относительно количества и принципов фундаментальных эмоций. Так, П. Экман и его сотрудники на основе изучения выражения лица выделяют 6 основных эмоций: гнев, страх, отвращение, удивление, печаль и радость. Р. Плутчик рассматривает 8 основных эмоций: одобрение, гнев, ожидание, отвращение, радость, страх, печаль, удивление. Классификация К. Изарда с 10-ю основными эмоциями, стала очень популярной. Он отмечает гнев, презрение, отвращение, дистресс (горе - страдание), страх, вину, интерес, радость, стыд, удивление. Есть и более простые классификации. Например, Я. Панксеп различает всего 5 основных эмоций: страх - тревога, ярость - гнев, паника - горе, ожидание - исследование, радость - игра.

Дж. Грей исследует три основные эмоции: ярость - ужас, тревогу и радость. По мнению Р. Мауэра, вообще следует говорить только о двух первичных эмоциях - удовольствии и боли.

Для дискретной модели вопрос о составе фундаментальных эмоций является критическим. Вопрос в том, сколько всего эмоций. Ответ зависит от того, какие показатели, какие реакции использовались при исследовании эмоций (вегетативные, биохимические, ЭМГ мышц лица и т.д.). По выражению лица сложно выделить более 10 паттернов активации мимических мышц, а по вегетативным параметрам можно описать еще меньше комплексов, связанных с эмоциями.

Многомерные модели эмоциональных состояний различаются между собой количеством измерений (осей) и их смысловой нагрузкой.

Адекватным методами для многомерной модели являются методы многомерного шкалирования (ММШ) или факторный анализ.

ММШ впервые был протестирован для изучения распознавания эмоций другого человека по внешним экспрессивным реакциям лица. Первоначальная процедура масштабирования применяет оценку внешних проявлений эмоций только по внутренним, заранее определенным шкалам. С развитием методов многомерной статистики в качестве исходных данных стали использоваться субъективные оценки различия или сходства лиц, выражающих различные эмоциональные переживания. Это позволяет определить и интерпретировать минимально допустимое количество шкал (факторов), которые ориентированы на предмет при вынесении суждений о различии, и тем самым приблизиться к механизму декодирования эмоций. В многомерном пространственном представлении каждое выражение представляет собой точку со своими собственными координатами или вектором, компоненты которого соответствуют вкладам главных осей (масштабов) пространства.

К первым работам в этом направлении относятся исследования Р. Вудворта (Woodworth R.S.) и Г. Шлосберга (Schlosberg N.). Р. Вудворте предложил первую классификацию выражений эмоций на лице с использованием шестичленной линейной шкалы: любовь, счастье, веселье; удивление; страх, страдание; гнев, решимость; отвращение; презрение.

Г. Шлоссберг, применив классификацию Вудворта к анализу фотографий людей с разной мимикой, обратил внимание на частое смешение самых далеких групп внешних выражений эмоций. Позже, в рамках теории активации Д. Линдсли (D. Lindsley), Г. Шлоссберг добавил третье измерение: сон - напряжение, создаваемое, таким образом, тремя измерениями эмоций, предложенными В. Вундтом в 1896 году: удовольствие - неудовольствие, расслабление - напряжение, спокойствие – волнение.

Однако в работе П. Абельсона и В. Сермата (Abelson R., Sermat V.), а также Р. Шепарда (Chapard R.) с использованием ММШ, когда испытуемые оценивали степень схожести между всеми возможными комбинациями пар из 13 фотографий по 9-балльной системе, были выдвинуты аргументы в пользу двухмерного решения. Была выделена шкала знаков эмоций (удовольствие - неудовольствие) и шкала активации (сон - бодрствование). Эти два фактора составляют примерно 75% от общего веса всех факторов. Аналогичные результаты были получены с использованием других наборов фотографий. В более поздних работах некоторые исследователи выделили новые измерения, такие как контроль - импульсивность, внимание - невнимательность, уверенность - неуверенность, которые были предложены в качестве третьей шкалы или вместо шкалы активации.

В экспериментах Дж. Рассела и М. Буллока (Russal J., Billok M.) с ранжированием мимики по фотографиям принимали участие дети 2, 3, 4 и 5 лет, а также студенты. Для испытуемых каждой возрастной группы было получено двумерное пространство, образованное двумя ортогональными осями: знаками и уровнем активации. Эмоции, изображенные на фотографиях, представлены точками, образующими фигуру, близкую по форме к кругу. Сходство эмоциональных пространств и взрослых детей в возрасте, когда еще нет дифференцированных представлений о различных эмоциях, свидетельствует о том, что структура опосредована восприятием эмоций, а не смысловыми связями, образующимися в процессе усвоения языка. Третья ось со значительно меньшим весом как у детей 4 лет, так и у взрослых была интерпретирована как шкала уверенности - смущения.

Особо следует выделить серию работ Дж. Рассела и его коллег. Применяя ММШ к субъективным оценкам сходства эмоций, он получил двумерную круговую модель эмоционального опыта (A Circumplex model of affect [12]). Эта двухмерная модель является PA-частью модели PAD.

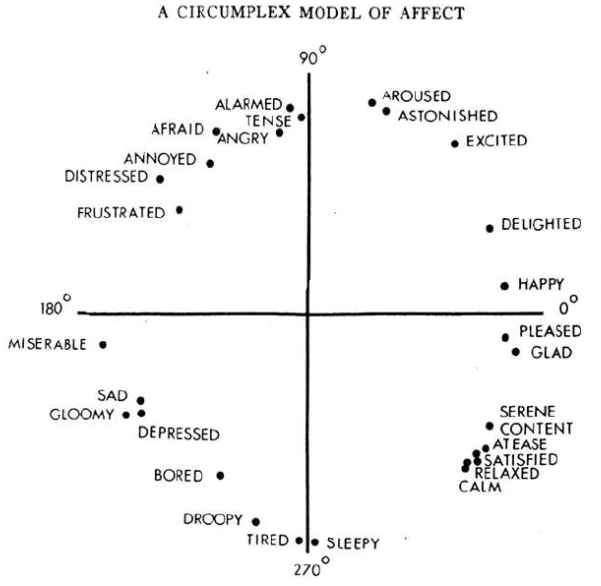


Рисунок 1.1 – A Circumplex model of affect

### **1.2.2 Модель эмоционального состояния PAD.**

В данном исследовании мы будем рассматривать многомерную модель PAD, в которой всего 3 измерения: P – Pleasure (удовольствие), A – Arousal (возбуждение), D – Dominance (доминирование) для всех представлений эмоций. Модель эмоционального состояния PAD - это психологическая модель, разработанная Альбертом Мехрабианом и Джеймсом А. Расселом (1974 г. и позже) для описания и измерения эмоциональных состояний. Первоначально использовавшаяся в теории психологии окружающей среды, основная идея заключалась в том, что физическая среда влияет на людей через их эмоциональное воздействие. Впечатление было использовано Питером Лэнгом и его коллегами, чтобы предложить физиологическую теорию эмоций. Его также использовал Джеймс А. Рассел для разработки теории эмоциональных эпизодов (относительно коротких эмоционально заряженных событий). PA-часть PAD была встроена в закрытую модель эмоционального опыта, и эти два измерения были названы «базовым аффектом». Часть D PAD была переосмыслена как часть процесса оценки эмоционального эпизода (холодная когнитивная оценка эмоциональной ситуации). Более подробная версия этого подхода называется психологической теорией построения эмоций. Модель PAD (Удовольствие, Возбуждение, Доминирование) использовалась для изучения невербального общения, такого как язык тела в психологии. Он также использует в потребительском маркетинге и создает анимированных персонажей, которые выражают эмоции в виртуальных мирах.

PAD использует трехмерные шкалы, которые теоретически могут иметь любое числовое значение. Размерная структура напоминает работы Вильгельма Вундта 19-го века, который также использовал трехмерную систему, а также работы Чарльза Э. Осгуда 20-го века.

Шкала удовольствия-неудовольствия измеряет, насколько приятно или неприятно человек чувствует что-то. Например, и гнев, и страх - неприятные эмоции, и оба они считаются недовольством. Однако радость - это приятное чувство.

Шкала возбуждения-бездействия измеряет, насколько человек чувствует себя возбужденным или гипнотическим. Дело не в интенсивности эмоций - горе и депрессия могут быть легким возбуждением и сильными чувствами. Хотя и гнев, и ярость являются неприятными эмоциями, ярость имеет более высокую интенсивность или более высокое состояние возбуждения. Однако скука, которая также является неприятным состоянием, имеет низкую ценность для возбуждения.

Шкала «Доминирование-покорность» представляет собой ощущение контроля и доминирования в сравнении с тем, что вы контролируете или подчиняетесь. Например, если страх и гнев - это неприятные эмоции, гнев - это доминирующая эмоция, а страх - это эмоция подчинения.

Более сокращенная версия модели использует всего 4 значения для каждого измерения, обеспечивая только 64 значения для возможных эмоций. Например, гнев - довольно неприятная, довольно возбужденная и умеренно доминирующая эмоция, а скука - слегка неприятная, совсем не возбужденная и в большинстве случаев не доминирующая эмоция.

### **1.2.3 Система кодирования лицевых движений FACS.**

Полом Экманом в XX веке была разработана система кодирования лицевых движений (FACS). Эта система позволяет классифицировать выражения лица человека на основе его мимики. С использованием FACS можно закодировать практически любое, анатомически возможное выражение лица, реконструируя его на вызвавшие его конкретные двигательные единицы (action units) и их временные сегменты.

FACS фактически не предусматривает любую биомеханическую информацию о степени активации мышц, но основные мышцы, участвующие в формировании выражений лица, были добавлены сюда для удобства читателя.

Двигательные единицы (action untis) представляют собой основные движения, совершаемые отдельными мышцами или группой мышц.

Двигательные дескрипторы (action descriptors) представляют собой движения, совершаемые группами мышц (например, выдвигание нижней челюсти вперед). Мышечная основа для этих движений не указана.

Для обозначения интенсивности задействования action units, в FACS к номеру AU добавляют латинские буквы с A по E, в зависимости от минимальной — максимальной интенсивности движения. Например, AU 1А является самым трудноразличимым движением AU 1, а AU 1Е является максимально возможной интенсивностью для определенного человека.

* A Слабо различимое
* B Незначительное
* C Заметное или ярко выраженное
* D Сильное или крайне заметное
* E Предельное

Двигательные единицы относительно воображаемой вертикальной оси лица могут быть:

* Двусторонние, симметричные (B — bilateral). Например, (B)1.
* Односторонние (U — unilateral). Например, U12, U14:
* левая (L — left). Например, L12.
* правая (R — right). Например, R14.

Примеры кодирования лицевых движений с помощью FACS можно изучить подробнее по ссылке [3].

С помощью двигательных единиц из FACS Полом Экманом и Уоллесом Фризеном была создана эмоциональная система кодирования лицевых движений (EmFACS), которая рассматривает только лицевые движения, связанные с эмоциями. Стоит отметить, что в научном мире нет единого мнения на тему однозначного соответствия между мимикой и эмоциями для всех людей [5]. Таким образом нельзя с полной уверенностью сказать, что, например, выражение лица любого человека, который испытывает страх выглядит именно так, а никак иначе. Однако, большинство людей в современном обществе выражает эмоции приблизительно одинаково, что даёт возможность использовать труды Пола Экмана в данной работе.

## **1.3 Постановка задач, исследовательские вопросы и гипотезы.**

### **1.3.1 Постановка задач.**

С учётом вышесказанного, встаёт вопрос о необходимости иметь возможность отображать пространства эмоций различных размерностей друг в друга. По сути, необходимо построить 6 функций:

* для отображения вектора модели эмоционального состояния PA (координатная плоскость pleasure, arousal из пространства PAD) в вектор 7-ми эмоций из модели Пола Экмана (Neutral – нейтральная, Happy – радость, Sad – печаль, Angry – гнев, Surprised – удивление, Scared – страх, Disgusted – отвращение) ( тип преобразования 2 -> 7)
* для отображения вектора 7-ми эмоций из модели Пола Экмана в вектор модели эмоционального состояния PA (тип преобразования 7 -> 2)
* для отображения вектора выражения лица человека, закодированного с помощью FACS, в вектор 7-ми эмоций из модели Пола Экмана (тип преобразования 42 -> 7)
* для отображения вектора 7-ми эмоций из модели Пола Экмана в вектор выражения лица человека, закодированного с помощью FACS (тип преобразования 7 -> 42)
* для отображения вектора модели эмоционального состояния PA в вектор выражения лица человека, закодированного с помощью FACS (тип преобразования 2 -> 42)
* для отображения вектора выражения лица человека, закодированного с помощью FACS, в вектор модели эмоционального состояния PA (тип преобразования 42 -> 2)

Ранее в НИРе 1-го семестра были реализованы нейросетевые модели автоэнкодеров, решающие задачу отображения из двухмерного пространства в семимерное (тип преобразования 2 -> 7).

В этом семестре планируется решить задачу обратного отображения из 7-ми мерного в 2-х мерное (тип преобразования 7 -> 2), а также представить модели отображения векторов выражений лиц людей, закодированных с помощью FACS, в вектора многомерной модели 7-ми эмоций Пола Экмана (тип преобразования 42 -> 7). Фактически, соединив 2 типа полученных моделей последовательно (тип преобразования 42 -> 7 и тип преобразования 7 -> 2) мы получим модель преобразования 42 -> 2.

В данной работе, помимо нейросетевых, будут использованы ещё и классические модели машинного обучения [13] (регрессионно-классификационные модели), такие как: регрессия (линейная, полиномиальная, Ridge/Lasso), Random Forest (случайный лес), KNN (K-Ближайших соседей). Классические модели обладают большей интерпретируемостью [14], в отличии от нейросетевых моделей, которые представляют из себя “модели чёрного ящика”. Подобное разнообразие из классических и нейросетевых моделей позволит сравнивать результаты между собой, а также оценивать их качество. 

Более подробно математическая постановка задач описана в подразделе 2.2.

### **1.3.2 Общая информация об исходных данных.**

В НИРе 1-го семестра, в качестве исходных данных для решения поставленной задачи представлены логи работы программы FaceReader. Ранее были записаны видеоматериалы, на которых испытуемые демонстрировали свои эмоции. Видеофайлы представляют из себя записи длительностью от нескольких секунд (10-20) до нескольких минут (1-2). Программа FaceReader проанализировала каждую видеозапись и выдала соответствующие результаты в виде логов.

В ходе работы предстоит проанализировать логи FaceReader’а и отобрать нужные данные для решения задачи отображения эмоций. Необходимо выделить показательную выборку, при этом постараться минимизировать влияние шумов.

Стоит отметить, что классическое измерение pleasure (удовольствие) из модели PAD в логах представлено как valence (валентность). При этом диапазон допустимых значений никак не меняется и остаётся прежним [-1; 1]. Но при этом координата Arousal находится в диапазоне [0; 1], что нетипично ввиду возможности наличия отрицательных значений Arousal (Рисунок 1.1). D – Dominance измерение опущено ввиду отсутствия необходимых данных.

Координаты семимерного вектора эмоций имеют диапазон [0; 1], хотя в сумме не дают 1. Т.е. возможны случаи, когда испытуемый одновременно демонстрирует несколько эмоций. Тогда сумма всех координат вектора превышает 1.

Все логи были объединены в 2 файла: data.csv, data\_clear.csv. Более подробное описание данных см. в подразделе 3.1.1 “Выбор датасета.”.

### **1.3.3 Гипотезы о данных и моделях.**

Задачи, поставленные в текущем исследовании напоминают задачу регрессионного анализа, суть которой заключается в установлении формы зависимости между переменными, оценка неизвестных значений (прогноз значений) зависимой переменной. Как следствие, существуют различные методы решения этой задачи как статистические, так и нейросетевые.

Более точно можно классифицировать задачу, как задачу повышения (снижения) размерности. По своему смыслу это преобразование данных, состоящее в уменьшении числа переменных путём получения главных переменных (в случае сжатия). Преобразование может быть разделено на отбор признаков и выделение признаков.

К статистическим методам решения задачи можно отнести метод главных компонент (МГК), линейный дискриминантный анализ (ЛДА) и обобщённый дискриминантный анализ (ОДА). К нейросетевым — применение автокодировщика (автоэнкодера).

В предыдущей работе Д.В.Тихомирова и А.В.Самсонович использовали метод главных компонент (PCA) для уменьшения размерностей данных из 42 признаков FACS в 7 компонент (аналог 'Neutral', 'Happy', 'Sad', 'Angry', 'Surprised', 'Scared', 'Disgusted'). Оказалось, что 7 главных компонент описывают около 80 % данных [15].

Чтобы обучение моделей (см. подраздел 1.3.1 “Постановка задачи.”) было более обоснованным, выдвинем аналогичную гипотезу о высоком % накопленной остаточной дисперсии PCA из 7 компонент ('Neutral', 'Happy', 'Sad', 'Angry', 'Surprised', 'Scared', 'Disgusted') в 2 (аналог 'Valence', 'Arousal'). Проверка гипотезы и её результаты описаны в подразделе 3.1.2 “Углубленный анализ данных.”.

Помимо регрессионных моделей, отдельный практический интерес в данной задаче представляют модели KNN (K-Ближайших соседей) и Random Forest (случайный лес). Более подробное обоснование выбора этих моделей описано в подразделах 2.3 “Сравнение классических моделей машинного обучения.” и 3.1.2 “Углубленный анализ данных.”.

# **Раздел 2. Теоретическая часть**

## **2.1 Изучение существующих решений поставленных задач.**

Из темы этого НИР следует, что эффективные модели машинного обучения, преобразующие вектора одного эмоционального пространства в другое, являются лишь частью от общего продукта, который требуется получить в итоговом результате.

Обученные модели отвечают за внутреннюю логику и решают основную задачу отображения пространств, но помимо моделей необходим удобный программный интерфейс для ввода/загрузки данных, для визуализации векторов в эмоциональных пространствах (VA, FACS, 7-ми мерное). Web-сервис, решающий эту задачу уже был разработан в НИРе 2-го семестра и доработан в ходе летней практики.

Конвертер, отображающий пространства эмоций различных размерностей может иметь множество практических применений. Его могут использовать различные виртуальные ассистенты или цифровые аватары (см. подраздел 1.1 “Технологии и подходы для реализации виртуальных ассистентов.”). Особенно это актуально в связи с активным развитием этого IT-направления и отсутствием единого стандарта для использования эмоциональных пространств. Иными словами, разные коммерческие решения могут использовать различные пространства эмоций различных размерностей. Для разрешения таких сложностей и может пригодиться разрабатываемый конвертер, в общем доступе.

Вдобавок к сказанному, анализ [16] рынка систем распознавания эмоций показал, что на данный момент нет бесплатных решений для отображения пространств эмоций различных размерностей. Есть лишь коммерческие решения и то, которые не выполняют преобразование напрямую (см. ниже 2.1.1 “FaceReader.” недостатки программы).

Таким образом, конечным результатом нескольких семестров НИРов будет готовый web-конвертер для решения задачи отображения пространств эмоций различных размерностей.

### **2.1.1 FaceReader.**

Много лет учёные всего мира работали над задачей распознавания мимики человека. В наши дни уже есть успехи в этом направлении. Если раньше для разложения выражения лица на составляющие в FACS требовался специально обученный человек, то теперь нейросети могут делать это автоматически.

Ярким примером является приложение FaceReader [11, 16]. Программа способна отличать различные эмоциональные выражения лица, такие как «счастливое», «грустное», «сердитое», «удивленное», «испуганное», «недовольное» и «нейтральное». Также FaceReader может отличать пол, возраст и этническую группу.

Для реализации используются технологии компьютерного зрения. С изображения лица человека считываются action units, которые преобразуются в точку на PA-модели.

Также, реализован метод Active Appearance Model, с помощью которого можно создавать искусственную модель лица с учетом контрольных точек и деталей поверхности, и сравнивать ее с образцами, заложенными в память.

Классификация происходит методами нейронных сетей с тренировочным набором в 2 000 фотографий.

Возможности программы:

* средний процент распознавания эмоций равен 89%. Для некоторых эмоций он выше, для некоторых ниже;
* наклон лица может быть любым в плоскости, его система обнаружит;
* программа работает с загружаемым видео в форматах с кодеками MPEG1, MPEG2, XviD, DivX4, DivX5, DivX6, DV-AVI и uncompressed AVI, причем определять эмоции можно пофреймно, либо полностью при просмотре всего видео. Также, FaceReader может работать со статичными изображениями, а также в реальном времени, если у пользователя подключена веб-камера;
* программа прекрасно визуализирована: всегда можно посмотреть гистограммы, диаграммы, процентаж выражаемых эмоций. А на таймлайне видны проявления микровыражений в определенный промежуток времени;
* FaceReader генерирует два текстовых файла, один – это лог проявления эмоций, а другой – статический, для сопряженной с данной программой уникальной системы управления визуальными данными The Observer XT, разработанной этой же компанией.

Недостатки программы:

FaceReader не натренирован для распознавания детей до 5-ти лет;

* Если человек в очках, то распознавание эмоций неточное, либо классификация не ведется;
* Люди с разным цветом кожи по-разному воспринимаются системой, программа не до конца адаптирована;
* Повернутое лицо не детектируется.

## **2.2 Математическая постановка задач.**

Первая задача этой работы - проведение углубленного анализа данных (см. подраздел 2.2.1 “Углубленный анализ данных.”) и проверка гипотез, выдвинутых в подразделе 1.3.3 “Гипотезы о данных и моделях.”.

Вторая задача, рассматриваемая в данном исследовании, является задачей многомерной (мульти-выходной) регрессии (см. подраздел 2.2.2 “Обучение моделей машинного обучения.”).

### **2.2.1 Углубленный анализ данных.**

В этом подразделе приведено описание теоретической составляющей углубленного анализа данных, включая метод главных компонент (PCA) и корреляционный анализ. Практические результаты оформлены в подразделе 3.1.2 “Углубленный анализ данных.”.

#### **2.2.2.1 Метод главных компонент (PCA).**

PCA - это метод для уменьшения размерности данных, который применяется для обнаружения структуры в данных путем нахождения нового набора переменных, называемых главными компонентами. Главная идея заключается в том, чтобы найти линейные комбинации исходных переменных (признаков), которые содержат максимальное количество информации.

Несколько основных шагов PCA включают в себя:

* **Стандартизация данных:** Важно стандартизировать данные перед применением PCA, чтобы убедиться, что все признаки имеют одинаковый масштаб.
* **Вычисление ковариационной матрицы:** PCA использует ковариационную матрицу для определения степени зависимости между признаками. Ковариационная матрица показывает, насколько два признака варьируются вместе.
* **Вычисление главных компонент:** Главные компоненты вычисляются как линейные комбинации исходных признаков таким образом, чтобы максимизировать объясненную дисперсию в данных. Главные компоненты упорядочиваются по убыванию объясненной дисперсии, что позволяет выбрать наиболее информативные компоненты.
* **Проекция данных на главные компоненты:** Затем исходные данные проецируются на пространство главных компонент, чтобы получить новый набор признаков с меньшей размерностью.

#### **2.2.2.2 Корреляционный анализ.**

Корреляционный анализ - это статистический метод, используемый для изучения связи между двумя или более переменными. Он позволяет определить, насколько тесно связаны переменные между собой и в каком направлении идет эта связь.

Корреляционная матрица - это таблица, в которой значениями являются коэффициенты корреляции между парами переменных. Она позволяет исследовать связи между всеми парами признаков в наборе данных. Значения коэффициента корреляции могут варьироваться от -1 до 1: значение ближе к 1 указывает на сильную положительную корреляцию (систематическую прямую пропорциональность), значение ближе к -1 - на сильную отрицательную корреляцию (систематическую обратную пропорциональность), а значение около 0 - на отсутствие корреляции.

### **2.2.2 Обучение моделей машинного обучения.**

Многомерная регрессия (мульти-выходная или мульти-таргет регрессия) — это тип задачи машинного обучения, где модель обучается предсказывать несколько непрерывных выходных переменных на основе одного или нескольких входных признаков.

Математически задачу можно сформулировать следующим образом:

* Дан набор входных данных *X*, где каждый элемент *xi* ∈ *X* представляет собой вектор входных признаков.
* Соответствующий набор выходных данных *Y*, где каждый элемент *yi* ∈ *Y* является вектором целевых переменных.
* Цель состоит в том, чтобы найти функцию *f*, которая наилучшим образом отображает входные данные на выходные: *f*(*xi*)≈*yi*

В контексте данной задачи размерность и вид входных и выходных данных зависит от типа преобразования (см. подраздел 1.3.1), а именно:

* тип преобразования 2 -> 7  
  Входные данные (*X*): двумерное пространство 'Valence' и 'Arousal'.

Выходные данные (*Y*): семь эмоциональных состояний ('Neutral', 'Happy', 'Sad', 'Angry', 'Surprised', 'Scared', 'Disgusted').

* тип преобразования 7 -> 2  
  Входные данные (*X*): семь эмоциональных состояний ('Neutral', 'Happy', 'Sad', 'Angry', 'Surprised', 'Scared', 'Disgusted').

Выходные данные (*Y*): двумерное пространство 'Valence' и 'Arousal'.

* тип преобразования 42 -> 7  
  Входные данные (*X*): вектор выражения лица человека, закодированного с помощью FACS (42 признака с именами, попадающих под шаблон *r'Action\_Unit\_\d{2}\_\_\_\w+'*).

Выходные данные (*Y*): семь эмоциональных состояний ('Neutral', 'Happy', 'Sad', 'Angry', 'Surprised', 'Scared', 'Disgusted').

* тип преобразования 7 -> 42  
  Входные данные (*X*): семь эмоциональных состояний ('Neutral', 'Happy', 'Sad', 'Angry', 'Surprised', 'Scared', 'Disgusted').

Выходные данные (*Y*): вектор выражения лица человека, закодированного с помощью FACS (42 признака с именами, попадающих под шаблон *r'Action\_Unit\_\d{2}\_\_\_\w+'*).

* тип преобразования 2 -> 42  
  Входные данные (*X*): двумерное пространство 'Valence' и 'Arousal'.

Выходные данные (*Y*): вектор выражения лица человека, закодированного с помощью FACS (42 признака с именами, попадающих под шаблон *r'Action\_Unit\_\d{2}\_\_\_\w+'*).

* тип преобразования 42 -> 2  
  Входные данные (*X*): вектор выражения лица человека, закодированного с помощью FACS (42 признака с именами, попадающих под шаблон *r'Action\_Unit\_\d{2}\_\_\_\w+'*).

Выходные данные (*Y*): двумерное пространство 'Valence' и 'Arousal'.

## **2.3 Сравнение моделей машинного обучения.**

В этом разделе мы проведем сравнительный анализ основных типов классических моделей машинного обучения, а также рассмотрим их особенности, указывая на типичные задачи, которые можно решить с помощью каждой из них.

* Регрессионные модели — предсказывают числовую характеристику объекта. Например, стоимость квартиры исходя из её площади или количество людей в аэропорту в зависимости от дня недели.
* Модели классификации — предсказывают категорию объекта по заранее заданным параметрам. Например, диагноз пациента на основе истории болезни, результатов анализов, симптомов.
* Регрессионно-классификационные (классические) модели — умеют решать оба типа задач.

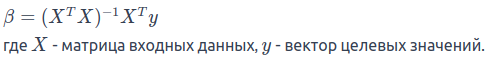
Применение классических моделей машинного обучения обусловлено тем, что использование автоэнкодера (см. НИР 1-го семестра) в задачах регрессии не является стандартным подходом. Автоэнкодеры обычно используются для уменьшения размерности или выявления скрытых признаков в данных, а не для прямого прогнозирования значений. По этой причине к задачам из подраздела 1.3.1 “Постановка задачи.” стоит добавить задачу типа преобразования 2 -> 7 для классических моделей.

В данном исследовании использовались, в основном, классические модели машинного обучения для решения задачи преобразования между пространствами эмоций различных размерностей. Ниже представлен краткий обзор каждой из используемых моделей и их обучение.

### **2.3.1 Линейная Регрессия.**

Линейная регрессия - это базовая модель машинного обучения, предназначенная для предсказания непрерывных выходных значений на основе линейной комбинации входных признаков.

Модель представляется уравнением: y=Xβ+ε, где  
y - выходные значения,   
X - матрица входных признаков,   
β - вектор весов,   
ε - ошибка модели. Обучение модели заключается в минимизации суммы квадратов ошибок между предсказанными и фактическими значениями.

Линейная регрессия стремится минимизировать сумму квадратов разностей между наблюдаемыми и предсказанными значениями. Аналитическое решение для оценки весов β может быть найдено с помощью метода наименьших квадратов:

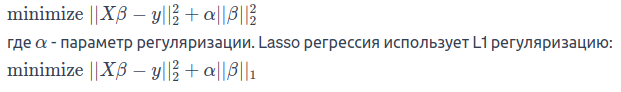
### **2.3.2 Полиномиальная Регрессия.**

Полиномиальная регрессия расширяет линейную регрессию, добавляя полиномиальные признаки (степени входных переменных). Это позволяет модели лучше аппроксимировать нелинейные зависимости.

Формула модели имеет вид: y=Xβ+ε, но X теперь включает полиномиальные признаки.

Аналитическое решение для полиномиальной регрессии аналогично линейной регрессии, но с учетом того, что входные данные трансформируются в полиномиальные признаки перед применением метода наименьших квадратов.

### **2.3.3 Ridge и Lasso Регрессии.**

Ridge (гребневая) и Lasso (лассо) регрессии - это варианты линейной регрессии, которые добавляют регуляризацию для предотвращения переобучения. Ridge регрессия минимизирует:  
  
Решения для Ridge и Lasso регрессий включают оптимизацию функций потерь с регуляризацией. Для Ridge:



аналитического решения из-за L1 регуляризации и решается с помощью численных методов оптимизации.

### **2.3.4 Случайный Лес (Random Forest).**

Случайный лес - это ансамблевая модель, состоящая из множества деревьев решений. Каждое дерево обучается на случайной подвыборке данных с использованием случайного подмножества признаков. Предсказание модели является усредненным результатом всех деревьев.

В случайном лесе каждое дерево решений стремится минимизировать функцию потерь, например, среднюю квадратичную ошибку (*MSE*) для задач регрессии. Особенностью случайного леса является использование ансамбля деревьев, где каждое дерево обучается на подвыборке данных с использованием случайного подмножества признаков. Финальное предсказание получается путем усреднения предсказаний отдельных деревьев.

Дерево решений - это модель машинного обучения, используемая как для классификации, так и для регрессии. Она строит модель принятия решений в виде древовидной структуры.

* Выбор Признака и Разбиения:   
  Для каждого узла дерева алгоритм выбирает наилучший признак для разбиения данных на подгруппы. Выбор основывается на метрике, такой как прирост информации (для классификации) или уменьшение среднеквадратичной ошибки (для регрессии).
* Метрики для Регрессионных Деревьев:  
  Среднеквадратичная ошибка (*MSE*) (см. подраздел 3.3 “Формирование метрик для оценки моделей.”)
* Уменьшение MSE:  
  Определяется как разность между MSE до и после разбиения. Узел с наибольшим уменьшением MSE выбирается как лучшее разбиение.
* Процесс Разбиения: Для каждого кандидата на разбиение (например, каждого значения признака) вычислить уменьшение MSE.   
  Выбрать точку разбиения, которая приводит к наибольшему уменьшению MSE.Разделить данные на две подгруппы на основе этого разбиения.
* Остановка Роста Дерева:   
  Рост дерева останавливается, когда достигнут один из критериев: достижение максимальной глубины, минимального числа образцов в узле или если улучшение MSE становится незначительным.

Деревья решений имеют несколько преимуществ, включая легкость интерпретации и способность обрабатывать как числовые, так и категориальные данные. Однако они могут быть склонны к переобучению, особенно в случае сложных данных. Для устранения этой проблемы часто используются ансамблевые методы, такие как случайный лес, который строит множество деревьев и усредняет их предсказания.

### **2.3.5 K-ближайших Соседей (KNN).**

Модель KNN работает на принципе нахождения "k" ближайших обучающих примеров в пространстве признаков и предсказывает значение на основе близости к этим соседям.

KNN не имеет явной тренировочной фазы. Вместо этого, он использует следующий псевдоалгоритм при предсказании:

* Для данного тестового примера *xtest*, вычислить расстояние до каждого обучающего примера.
* Выбрать *k* обучающих примеров, которые находятся ближе всего к *xtest*.
* Прогнозировать выходное значение как среднее (для регрессии) или как наиболее частое значение (для классификации) среди выбранных *k* ближайших соседей.

### **2.3.6 Многослойный Перцептрон (MLPRegressor).**

MLPRegressor - это тип нейронной сети, состоящий из нескольких слоев нейронов, которые могут моделировать сложные нелинейные отношения. Он обучается с использованием методов обратного распространения и градиентного спуска.

# **Раздел 3. Инженерная часть**

## **3.1 Выбор датасета, анализ данных, формирование обучающей и тестовой выборок.**

В данном разделе выбираются датасеты, данные в датасетах подвергаются тщательному анализу, формируются обучающие и тестовые выборки.

### **3.1.1 Выбор датасета.**

### **3.1.2 Углубленный анализ данных.**

#### **3.1.2.1 Диаграмма рассеяния с указанием доминирующей эмоции.**

#### **3.1.2.2 Проверка гипотез с помощью МГК.**

#### **3.1.2.3 Корреляционный анализ.**

#### **3.1.2.4 Визуализация границ кластеров эмоциональных состояний.**

### **3.1.3 Формирование обучающей и тестовой выборок.**

## **3.2 Метод поиска оптимальных значений ключевых параметров моделей.**

## **3.3 Формирование метрик для оценки моделей.**

# **Раздел 4. Технологическая и практическая часть**

## **4.1 В**[**ыбор инструментов для анализа данных и**](#_1mrcu09) **обучения моделей.**

В данной работе есть прогресс по сравнению с НИРом 1-го семестра, где большая часть кода была написана вручную. После трех семестров обучения найдено множество готовых инструментов, которые значительно упрощают процесс обработки данных, обучения моделей и визуализации результатов. Теперь большая часть работы связана с анализом данных и применением этих инструментов для создания и оценки моделей. Все коды, реализованные в текущей работе, написаны на языке Python.

### **4.1.1 И**[**нструмент**](https://docs.google.com/document/d/1KMNmryZLG4g2AI4jKY3cgtjJO81uIJQrturQumUJPeo/edit#heading=h.1mrcu09)**ы для анализа и обработки данных.**

### **4.1.2 И**[**нструмент**](https://docs.google.com/document/d/1KMNmryZLG4g2AI4jKY3cgtjJO81uIJQrturQumUJPeo/edit#heading=h.1mrcu09)**ы для обучения моделей.**

### **4.1.3 И**[**нструмент**](https://docs.google.com/document/d/1KMNmryZLG4g2AI4jKY3cgtjJO81uIJQrturQumUJPeo/edit#heading=h.1mrcu09)**ы для визуализации.**

### **4.1.4 И**[**нструмент**](https://docs.google.com/document/d/1KMNmryZLG4g2AI4jKY3cgtjJO81uIJQrturQumUJPeo/edit#heading=h.1mrcu09)**ы для сохранения моделей.**

## **4.2 Обучение моделей.**

## **4.3 Оценка точности работы моделей.**

### **4.3.1 Визуализация работы обученных моделей.**

### **4.3.2 Выводы по таблицам результатов обучения моделей.**

## **4.4 Сохранение лучших моделей в формате файла обученных моделей.**

# **Заключение**

# **СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. МНОГОМЕРНАЯ И ДИСКРЕТНАЯ МОДЕЛИ ЭМОЦИЙ [Электронный ресурс] / URL: https://ozlib.com/851130/psihologiya/mnogomernaya\_diskretnaya\_modeli\_emotsiy
2. Модель эмоционального состояния PAD [Электронный ресурс] / URL: https://ru.abcdef.wiki/wiki/PAD\_emotional\_state\_model
3. Система кодирования движений лица (FACS) - Визуальное руководство [Электронный ресурс] / URL: https://imotions.com/blog/facial-action-coding-system/
4. Система кодирования лицевых движений [Электронный ресурс] / URL: https://ru.wikipedia.org/wiki/Система\_кодирования\_лицевых\_движений
5. Экман, Пол - Критика [Электронный ресурс] / URL: https://ru.wikipedia.org/wiki/Экман,\_Пол#Критика
6. Виртуальный ассистент [Электронный ресурс] / URL: https://ru.wikipedia.org/wiki/Виртуальный\_ассистент
7. Распознавание речи - История [Электронный ресурс] / URL: https://ru.wikipedia.org/wiki/Распознавание\_речи#История
8. Обзор технологий синтеза речи [Электронный ресурс] / URL: https://habr.com/ru/company/tinkoff/blog/474782/
9. Основы построения интеллектуальных систем: учеб. пособ./ Г.В. Рыбина. – М.: Финансы и статистика; ИНФРА-М, 2010. – 432 с.: ил.
10. Цифровые аватары: как виртуальные помощники [Электронный ресурс] / URL: https://vc.ru/future/218083-cifrovye-avatary-kak-virtualnye-pomoshchniki-pereselyayutsya-iz-fantasticheskih-filmov-v-nashi-kvartiry-i-ofisy
11. EMOTION ANALYSIS FaceReader [Электронный ресурс] / URL: https://www.noldus.com/facereader
12. A Circumplex Model of Affect // Journal of Personality and Social Psychology 39(6), 1980, pp. 1161-1178
13. Какие бывают модели машинного обучения / Яндекс Практикум [Электронный ресурс] / URL: https://practicum.yandex.ru/blog/modeli-mashinnogo-obucheniya/#kakie-byvayut
14. Интерпретируемые модели / ИТМО вики-конспекты [Электронный ресурс] / URL: <https://neerc.ifmo.ru/wiki/index.php?title=%D0%98%D0%BD%D1%82%D0%B5%D1%80%D0%BF%D1%80%D0%B5%D1%82%D0%B8%D1%80%D1%83%D0%B5%D0%BC%D1%8B%D0%B5_%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D0%B8>
15. Широкий В.Р. , Тихомирова Д.В., Владимиров Р. Д., Доленко С. А., Самсонович А.В The Loop of Nonverbal Communication Between Human and Virtual Actor: Mapping Between Spaces. Advances in Intelligent Systems and Computing, V.1310, pp.484-489 (год публикации - 2021) <https://doi.org/10.1007/978-3-030-65596-9_58>
16. Исследование рынка систем распознавания эмоций [Электронный ресурс] / URL: <https://habr.com/ru/post/133686/>
17. Архитектура платформы машинного обучения в продакшене [Электронный ресурс] / URL: https://habr.com/ru/articles/688406/
18. Введение в машинное обучение с помощью Python: руководство для специалистов по работе с данными. Андреас Мюллер, Сара Гвидо. - Москва [и др.] : Диалектика, 2017. - 472, [1] с. : ил., табл., цв. ил.;