

АННОТАЦИЯ

**диссертационной работы докторанта PhD
специальности 6D070400 – «Вычислительная техника и программное
обеспечение».**

Шойынбек Айсултан

**на тему: «Автоматизированный сбор данных эмоциональной речи
для задачи распознавания эмоций по голосу».**

Актуальность работы. Эмоции занимают важное место в межличностных человеческих взаимодействиях и отношениях. Эмоции влияют на нашу жизнь, потому что это внутренняя реакция человека на окружающие и происходящие обстоятельства. Это помогает нам принимать разумные решения, лучше чувствовать настроение другого человека, передавать чувства в качестве обратной связи для понимания реакции и адаптироваться к реальности человеческого поведения. Все эти факты побудили ученых изучать и измерять уровень эмоционального состояния человека. Учитывая сегодняшние реалии, при развитии робототехники и искусственного интеллекта одним из популярных направлений исследования является распознавание эмоций машиной с последующей возможностью их воспроизведения. Основная цель задачи распознавания эмоций - помочь человеку взаимодействовать с машиной. Таким образом, распознавание эмоций играет важную роль в продуктах искусственного интеллекта, поскольку это сделает интерфейс человек-компьютер (HCI) более удобным и улучшит взаимодействие с пользователем.

Есть три основных шага для достижения успешных результатов в распознавании эмоций по голосу, а именно: (1) сбор эмоциональных речевых данных или выбор уже существующей базы данных, (2) предварительная обработка данных и извлечение полезных функций, (3) разработка надежных моделей классификации на основе алгоритмов машинного обучения.

Многие исследования направлены на открытие и изучение новых признаков, содержащих больше эмоциональной информации. Другие исследователи концентрируются на улучшении классификаторов алгоритмов машинного обучения; однако в большинстве исследований использовалась эмоциональная речь, собранная в контролируемых условиях. Практически все существующие результаты были получены из баз данных, которые были записаны актерами, имитирующими и искусственно вызывающими эмоции перед микрофоном. Существующие корпуса эмоций имеют следующие

ограничения: (1) количество эмоциональных высказываний ограничены числом и скудны в разнообразии эмоциональной речи, что не может привести к высокой точности распознавания для машинного обучения, (2) эмоции неестественны, и это означает, что машина учится распознавать фальшивые эмоции.

Актуальность темы основана на необходимости создания эффективных подходов к сбору и маркировке естественных эмоциональных высказываний для распознавания речевых эмоций с высокой точностью.

Цель и задачи исследования. Целью научного исследования является разработка методов и алгоритмов для автоматического сбора и маркировки данных, для распознавания естественных речевых эмоций с высоким уровнем точности, не менее 80%.

Задачи исследования. В соответствии с поставленной целью в данной работе определены следующие задачи:

- изучить и проанализировать существующие эмоциональные корпуса.
- изучить и проанализировать современные методы сбора эмоциональных данных.
- изучить теоретические основы человеческой речи и голоса и способы их преобразования в машинный язык.
- проанализировать и обосновать выбор надежного метода преобразования признаков для задачи распознавания речевых эмоций.
- проанализировать и обосновать выбор классификатора для задачи распознавания речевых эмоций.
- разработать или обосновать выбор модели машинного обучения для разделения речевого и неречевого аудио контента.
- разработать методы и алгоритм извлечения аудио эмоционального содержания из видео.
- разработать методы и алгоритм автоматизированной маркировки эмоциональных высказываний.

Объект исследования. Объектом исследования являются методы автоматизированных сборов и маркировки данных для распознавания речевых эмоций.

Методы исследования. Поставленные задачи решались путем проведения теоретических и эмпирических исследований. В рамках исследования использовались концептуальные положения искусственного интеллекта, классические теории машинного обучения, исследования

ведущих зарубежных и отечественных ученых в области распознавания эмоций и речи, теории вероятностей, математической статистики, численного анализа.

На защиту выносятся следующие научные положения:

Методы и алгоритмы интеллектуального анализа эмоциональных данных

Методы и алгоритмы автоматического поиска и загрузки видео, содержащих эмоциональные высказывания.

Методы и алгоритмы извлечения эмоциональных высказываний из видео.

Методы и алгоритмы классификации и разметки эмоциональных высказываний.

Автоматизированная система сбора и маркировки эмоциональных высказываний для распознавания речевых эмоций.

Разработана модель глубокой нейронной сети для классификации человеческой речи от различных звуков (шумов).

Научная новизна работы. Научная новизна диссертации заключается в разработке автоматизированного метода сбора и маркировки данных в виде эмоциональных речевых высказываниях. Результаты, полученные в этой диссертации, значительно продвинули область искусственного интеллекта в распознавании речевых эмоций. Используя метод сбора эмоциональных данных, ученые смогут собирать наборы данных с естественными эмоциональными высказываниями на всех языках мира автоматизировано.

Практическая значимость результатов проведенных исследований.

Практическая ценность диссертации заключается в возможности качественного улучшения обслуживания call-центров в сфере: образования, банковском деле, страховании, коммунальных услугах и медицине.

Практическая значимость исследования определяется возможностью применения его результатов и рекомендаций в: разработке систем распознавания истинных или ложных эмоций, составлении правоохранительными органами эмоционального портрета преступника, выявлении депрессивных и суицидальных наклонностей с целью предотвращения детских самоубийств.

Апробация работы. Основные результаты исследования докладывались и обсуждались на международных конференциях: «IV международной научно-практической конференции "Информатика и прикладная математика", посвященной 70-летию юбилею профессоров

Биярова Т.Н., Вальдемара Вуйцика и 60-летию профессора Амиргалиева Е.Н.» проходившей в 2019 г. в г. Алматы, Казахстан и «III Международной научной конференции «Информатика и прикладная математика», посвященная 80-летнему юбилею профессора Бияшева Р.Г. и 70-летию профессора Айдарханова М.Б» проходившей в 2018 г. в г. Алматы, Казахстан; «Trends of Modern Science», проходившем в г. Шеффилд, Великобритания 2019г.

Публикации. В рамках данной диссертационной работы было подготовлено и издано 10 исследовательских работ по рассматриваемой теме, из них:

- две статьи опубликованы в издательствах, отвечающих требованиям высшей аттестационной комиссии Министерства образования науки РК;
- четыре статьи размещены в сборниках трудов международных конференций;
- одна статья в зарубежном издании
- три статьи в международном рецензируемом научном журнале.

Структура и объем диссертации. Диссертационная работа состоит из введения, 4 разделов, заключения и списка использованных источников. Она изложена на 110 страницах машинописного текста, содержит 51 рисунок, 31 таблиц, список использованных источников из 115 наименований.

В первой главе проведен обзор существующих исследований.

Во второй главе подробно изучена природа голоса и речевых эмоций.

В третьей главе предлагается модель машинного обучения глубоких нейронных сетей, а также путем сравнительного анализа определен наиболее эффективный признак для трансформации голосовых эмоций в машинный вид.

В четвертой главе предлагается метод по автоматическому сбору маркированных эмоционально-речевых высказываний, классификатор речевых и неречевых аудио данных, классификатор речевых эмоций.

В заключении диссертации по результатам представленных исследований сформулированы основные выводы работы.