

Данные заявки

Направление заявки:

Н1. Цифровые технологии

Полуфинал:

Полуфинальный отбор в ФГБОУ ВО "Пензенский государственный университет"

ДАнные О ПРОЕКТЕ

Название проекта:

Разработка и исследование технологии обнаружения и классификации психоэмоционального состояния человека по речевым сигналам

Область техники:

ОТ1.02. Обработка цифровых сигналов

Приоритетное направление:

3. Информационно-телекоммуникационные системы.

Критическая технология федерального уровня:

8. Нано-, био-, информационные, когнитивные технологии.

Ключевые слова:

Цифровая обработка речевых сигналов; психоэмоциональное состояние человека; машинное обучение; обнаружение и классификация эмоций человека.

Участие в других проектах:

Опыт в работе с научными и коммерческими проектами разного уровня.

Опыт участия в международных конференциях. Опыт коммерческой разработки ПО более двух лет. Опыт работе в команде. Опыт выступления на английском языке. Опыт работы с нейронными сетями. Опыт защиты докладов и подготовки презентаций.

Профессиональные достижения:

2015

«Робофест-2016», Пенза, третье место с проектом «Автоматизированный токарный станок», <https://penzavzglyad.ru/news/151/v-penze-sostoyalis-otborochnye-sorevnovaniya-robofest>

2016

Участие в конкурсе «ЭВРИКА», Пенза, интервью для телевидения, ссылка: <https://www.penza-press.ru/lenta-novostey/90065/penzenskie-shkolniki-izobreli-avtomatizirovannyj-tokarnyj-standok>

«Робофест-2017», Пенза

2017

Победа в конкурсе «ЭВРИКА», первое место.

Прохождение Зимней Школы ДомInno2017 от 1000-list-nick, направление радиоэлектроника

Участие в международной-практической конференции «Таланты 21-го века», г. Минск Беларусь, ссылка на СМИ: <https://penzanews.ru/education/112336-2017>

Прохождение Летней Школы TeenГрад2017 от 1000-list-nick, направление программирование

Участие в «Робофинист-2017», Санкт-Петербург

2018

Участие в конкурсе «ЭВРИКА», Пенза, третье место.

Прохождение Зимней Школы ДомInno2018 от 1000-list-nick, направление программирование

Участие в конкурсе Робофест-2018, Москва. Олимпиада по физике в МГУ.

Прохождение Летней Школы TeenГрад2018 от 1000-list-nick, направление программирование

Волонтёр «Робофест-2017», Пенза

2019

Прохождение Зимней Школы ДомInno2019 от 1000-list-nick, направление программирование

Прохождение Летней Школы TeenГрад2019 от 1000-list-nick, направление программирование

Участие в полуфинале конкурса «Преактум» в г. Казань, выход на второе выступление(обойдя при этом 75% участников)

2020

Тьютор (куратор, наставник) Зимней Школы ДомInno2019 от 1000-list-nick

Техническая организация нескольких онлайн-викторин на уровне ВУЗа

Конференция «Актуальные вопросы современной науки и образования», г. Алушта, подача доклада, диплом 1 степени.

2021

Тьютор (куратор, наставник) Зимней Школы ДомInno2019 от 1000-list-nick

Сколково, Навигатор Инноватора, проект «NeiRelax»

Конференция «Актуальные вопросы современной науки и образования», г. Алушта, подача двух докладов, диплом 1 степени.

Участие в полуфинале студенческого конкурса «Твой Ход», г. Казань, 13-14 октября

По результатам полуфинала студенческого конкурса «Твой Ход» я прошёл в финал. Финал конкурса состоится на базе университета «Сириус» в г. Сочи 26-30 ноября

ПРОФИЛЬ УЧАСТНИКА ПРОЕКТА

ФИО:

Дудников Дмитрий Сергеевич

Контактный email:

dmitriy.s.gmpf@gmail.com

Номер телефона:

+7 937 410 48 49

Дата рождения:

26.12.2001 00:00 (19)

Пол:

ИНН:

583515989890

Регион:

Пензенская область

Город:

Пенза

Почтовый индекс:

440052

Почтовый адрес:

ул. Гоголя 41

Выпускник ОЦ «Сириус»:

Нет

Место работы:

ООО «ДМТ»

Должность:

Ведущий программист C++

ОБРАЗОВАНИЕ УЧАСТНИКА ПРОЕКТА

Учёная степень:

Не имею

Учёное звание:

Не имею

Учебное заведение:

Пензенский государственный университет

Специальность:

Факультет информационных технологий и электроники, управление в технических системах, автоматика и телемеханика, 3 курс

НАУЧНО-ТЕХНИЧЕСКАЯ ЧАСТЬ ПРОЕКТА

Цель выполнения проекта: Повышение эффективности обнаружения и классификации психоэмоционального состояния человека по речевым сигналам для здоровья сбережения населения

Задачи по проекту в рамках договора по программе «УМНИК»: Задача 1.

Поиск, анализ и систематизация аудио и видео материала на информационных ресурсах с открытым доступом (аудиохостингах и видеохостингах), содержащего речь на русском языке, зарегистрированную в условиях естественного переживания эмоций человеком в реальных жизненных ситуациях (отрицательных, нейтральных и положительных психоэмоциональных состояниях, различных по выраженности, длительности, осознанности, рациональности, уровню проявления и др.).

Задача 2. Формирование речевой базы данных, включающей не менее 500 зарегистрированных звуковых записей (слов, фраз и предложений), релевантных нарушениям работы органов речевого аппарата вследствие естественного переживания эмоций человеком.

Задача 3. Исследование информативных паттернов речевых сигналов: поиск скрытых особенностей речи и формирование оптимального набора паттернов, релевантных естественно выраженным эмоциям посредством применения новых адаптивных методов частотно-временного анализа информативно-значимых амплитудных, временных, частотных и энергетических характеристик речи.

Задача 4. Разработка оригинальных алгоритмов и способов обработки речевых сигналов, предназначенных для применения в реальных условиях эксплуатации, направленных на фильтрацию и подавление шума, сегментацию на информативные участки и определение паттернов, релевантных естественно выраженным эмоциям посредством применения новых адаптивных методов анализа и комбинирования известных способов обработки речи.

Задача 5. Разработка оригинальных способов классификации отрицательных, нейтральных и положительных психоэмоциональных состояний человека, посредством применения известных алгоритмов

классификации и объединения алгоритмов статической и динамической классификации в комбинированные способы.

Назначение научно-технического продукта (изделия и т.п.): Разрабатываемая технология предназначена для обнаружения и классификации психоэмоциональных состояний человека посредством анализа информативных паттернов речевых сигналов, релевантных естественно выраженным эмоциям. Своевременное обнаружение и классификация психоэмоционального состояния человека особенно важно в тех отраслях деятельности, которые сопряжены с повышенным риском возникновения техногенных и биогенных аварий, террористических угроз с серьезными последствиями: пилоты, космонавты, военнослужащие, диспетчеры аэропортов, диспетчеры опасных производственных объектов (АЭС, ТЭС, объектов химической промышленности и т. п.).

Научная новизна предлагаемых в проекте решений: Научной новизной предлагаемых в проекте решений является:

- применение нового подхода к поиску, анализу и систематизации звуковых записей, зарегистрированных в условиях естественного переживания эмоций (радость, нейтралитет, печаль, страх, раздражение и гнев) человеком в реальных отрицательных, нейтральных и положительных жизненных ситуациях, релевантных нарушениям работы органов речевого аппарата.
- использование адаптивных методов обработки речевых сигналов (преобразование Гильберта-Хуанга, разновидностей декомпозиций на эмпирические моды и др.), позволяющих определять релевантные информативные паттерны естественно выраженных психоэмоциональных состояний, учитывая уникальные особенности речи (скрытые модуляции, области концентрации энергии и т.п.);
- объединение алгоритмов статической и динамической классификации в комбинированный способ с использованием глубоких нейронных сетей, позволяющих объединить в одном алгоритмическом решении автоматическое выделение информативных паттернов и распознавание отрицательных, нейтральных и положительных психоэмоциональных состояний человека.

Обоснование необходимости проведения НИР: Контроль психоэмоционального состояния населения на сегодняшний день является социально-значимой проблемой для каждого государства, поскольку напрямую связан с формированием здорового образа жизни населения и профилактикой психоэмоциональных расстройств. По данным ВОЗ современные социально-значимые заболевания, являющиеся главной причиной временной нетрудоспособности, инвалидности и смертности населения, напрямую связаны с психоэмоциональным состоянием.

Наиболее перспективным и адаптивным к современным условиям трудовой деятельности человека является способ обнаружения и классификации психоэмоционального состояния на основе анализа речевых данных.

Анализ открытых источников отечественной и зарубежной литературы выявил множество научных работ, посвященных: созданию речевых баз данных, содержащих сигналы с выраженными эмоциональными составляющими; исследованию информативно-значимых паттернов

эмоциональной составляющей речевых сигналов; разработке алгоритмов, способов и средств определения паттернов речи, отражающих психоэмоциональное состояние человека; проектированию систем автоматическому распознаванию эмоций человека по речи. Но, несмотря на большое количество проведенных исследований в данной области, ряд проблем все еще остается нерешенным и требуют дальнейшей проработки:

- отсутствие речевой базы данных, зарегистрированной в реальных жизненных ситуациях естественного переживания эмоций человеком (различных по выраженности, длительности, осознанности, рациональности и уровню проявления);
- отсутствие оптимальных наборов паттернов речевых сигналов, максимально «родственных» человеческому восприятию и релевантных нарушениям работы органов речевого аппарата вследствие естественного переживания эмоций человеком в реальных отрицательных, нейтральных и положительных психоэмоциональных состояниях;
- отсутствие высокоэффективных алгоритмов, способов и средств выделения релевантных паттернов речевых сигналов и классификации психоэмоциональных состояний.

Данная НИР направлена на поиск решений упомянутых выше научных проблем, достижение которых позволит значительно повысить эффективность обнаружения и классификацию психоэмоциональных состояний человека.

Основные технические параметры, определяющие количественные, качественные и стоимостные характеристики продукции (в сопоставлении с существующими аналогами, в т.ч. мировыми): В соответствии с заявленной научной новизной предлагаемых в проекте решений предполагается, что разрабатываемая технология обнаружения и классификации психоэмоционального состояния человека по речевым сигналам будет иметь следующие технические параметры:

1. Технология будет обеспечивать обнаружение и классификацию психоэмоциональных состояний человека по речевым сигналам в режиме реального времени.
2. Технология будет обеспечивать классификацию психоэмоционального состояния человека на отрицательное, нейтральное и положительное.
3. Технология будет обеспечивать визуализацию результатов классификации психоэмоционального состояния человека в режиме реального времени с помощью специализированного программного обеспечения.
4. Технология будет обеспечивать точность классификации для:
 - отрицательного психоэмоционального состояния человека с ошибкой первого рода не более 5 %;
 - нейтрального психоэмоционального состояния человека с ошибкой первого рода не более 7 %;
 - положительного психоэмоционального состояния человека с ошибкой первого рода не более 10 %;
5. Технология будет функционировать с операционной системой Windows 7, 8, 10.

6. Технология будет функционировать со следующим аппаратным обеспечением:

- процессор с тактовой частотой - не менее 1 ГГц;
- объем свободной оперативной памяти - не менее 512 Мб;
- объем свободной памяти на жестком диске - не менее 2 Гб.

7. Технология будет функционировать в режиме доступа к сети «Интернет». Конструктивные требования (включая технологические требования, требования по надежности, эксплуатации, техническому обслуживанию, ремонту, хранению, упаковке, маркировке и транспортировке): Технология будет функционировать как программное обеспечение в двух вариантах: с открытым и закрытым исходными кодами.

В первом случае функционал будет адаптирован под нужды разработчиков, доступ к технологии будет бесплатным, частично ограниченным по функционалу (представлены основные инструменты для построения функционала и архитектуры собственной системы).

Во втором случае функционал будет полноценным, адаптированным под нужды бизнеса и различных сфер жизни, без возможности изменения исходного кода программы.

Программное обеспечение будет состоять из базы данных эмоциональной речи; инструментов анализа информативных паттернов речевых сигналов, релевантных естественно выраженным эмоциям; нейросетевого программного модуля, обученного на информативных паттернах, соответствующих отрицательным, положительным и нейтральным психоэмоциональным состояниям.

Требования по патентной защите (наличие патентов), существенные отличительные признаки создаваемого продукта (технологии) от имеющихся, обеспечивающие ожидаемый эффект: Все достигнутые результаты в рамках выполнения НИР будут зарегистрированы в качестве программного обеспечения в Федеральном институте промышленной собственности в виде свидетельства о государственной регистрации программы для ЭВМ.

В рамках выполнения НИР ни чьи права на интеллектуальную собственность в России и в других странах нарушаться не будут.

Календарный план проекта в рамках договора по программе «УМНИК»

Первый этап: 1) Поиск, анализ и систематизация аудио и видео материала на информационных ресурсах с открытым доступом (аудиохостингах и видеохостингах), содержащего речь на русском языке, зарегистрированную в условиях естественного переживания эмоций человеком в реальных

жизненных ситуациях (отрицательных, нейтральных и положительных психоэмоциональных состояниях, различных по выраженности, длительности, осознанности, рациональности, уровню проявления и др.).

2) Обработка звуковых записей, извлеченных из аудио и видео материала, включающая фильтрацию, сглаживание, временную и частотную коррекцию, удаление эхо и ревербераций, преобразование стерео в моно, нормализацию, монтаж и компоновку, оценку разборчивости речи, преобразование в единый формат WAVE PCM для хранения.

3) Формирование речевой базы данных, включающей не менее 10000 зарегистрированных звуковых записей (слов, фраз и предложений), релевантных нарушениям работы органов речевого аппарата вследствие естественного переживания эмоций человеком в реальных отрицательных, нейтральных и положительных психоэмоциональных состояниях.

4) Исследование паттернов речевых сигналов: поиск скрытых особенностей речи и формирование оптимального набора параметров, релевантных естественно выраженным эмоциям посредством применения новых адаптивных методов частотно-временного анализа информативно-значимых амплитудных, временных, частотных и энергетических характеристик речи.

Второй этап: 1) Разработка оригинальных алгоритмов и способов обработки речевых сигналов, предназначенных для применения в реальных условиях эксплуатации, направленных на фильтрацию и подавление шума, сегментацию на информативные участки и определение паттернов, релевантных естественно выраженным эмоциям посредством применения новых адаптивных методов анализа и комбинирования известных способов обработки речи.

2) Разработка оригинальных способов классификации отрицательных, нейтральных и положительных психоэмоциональных состояний человека, посредством применения известных алгоритмов классификации и объединения алгоритмов статической и динамической классификации в комбинированные способы.

3) Экспериментальные исследования, отражающие объективную оценку эффективности разработанных алгоритмов и способов обработки речевых сигналов для определения релевантного набора паттернов естественно выраженных эмоций; системы обнаружения и классификации отрицательных, нейтральных и положительных психоэмоциональных состояний человека.

КОММЕРЦИАЛИЗУЕМОСТЬ НАУЧНО-ТЕХНИЧЕСКИХ РЕЗУЛЬТАТОВ

Область применения: Своевременное обнаружение и классификация психоэмоционального состояния человека особенно важно в тех отраслях деятельности, которые сопряжены с повышенным риском возникновения техногенных и биогенных аварий, террористических угроз с серьезными последствиями. В частности постоянный мониторинг психоэмоционального

состояния необходим специалистам, деятельность которых связана с повышенной ответственностью: пилоты, космонавты, военнослужащие, диспетчеры аэропортов, диспетчеры опасных производственных объектов (АЭС, ТЭС, объектов химической промышленности и т. п.).

Данная НИР является актуальной с медицинской точки зрения - в совокупности с другими методиками, разрабатываемая технология может быть использована в целях диагностики и профилактики психических расстройств населения.

Объем внебюджетных инвестиций, собственных средств и иных источников, источники средств и формы их получения, распределение по статьям: Отсутствует.

Имеющиеся аналоги: Проведен краткий обзор продуктов на рынке речевых технологий, позволяющий сформулировать некоторые обязательные требования к разработке технологии обнаружения и классификации психоэмоционального состояния человека по речи. Обзор содержит исследование шести продуктов, получивших широкую практическую популярность в области речевой аналитики. В широком понимании речевая аналитика представляет собой анализ речи и смысловых высказываний собеседников, отдельный анализ голоса и его характеристики, интонационный и просодический анализ, эмоциональных анализ, анализ окружающей среды, определения количество дикторов и ряд других.

1. Система речевой аналитики «SVQA»

Производитель: частотное предприятие «Речевые системы»

Страна производитель: Беларусь, г. Минск

Назначение: безопасность, бизнес (банковская сфера, контакт-центры)

Доступ: платный

Стоимость: по запросу

<https://svqa.ru>

2. Подсистема автоматической оценки качества «Smart Logger» многоканальной системы регистрации телефонных вызовов и речевых сообщений «НЕЗАБУДКА II»

Производитель: группа компаний «Центр речевых технологий» (ЦРТ) (компании ЦРТ, ЦРТ-инновации и SpeechPro)

Страна производитель: Россия, г. Санкт-Петербург

Назначение: безопасность, бизнес (банковская сфера, контакт-центры)

Доступ: платный

Стоимость: по запросу

<https://www.speechpro.ru>

3. Профессиональная система диагностики голоса и речи «lingWAVES»

Производитель: компания WEVOSYS medical technology GmbH

Страна производитель: Германия, г. Баунах (Baunach, Germany)

Назначение: медицина

Доступ: платный

Стоимость: по запросу

<https://www.wevosys.com>

4. Технология анализа голоса «Sense»

Производитель: компания «NemeSYSCO»

Страна производитель: Израиль

Назначение: безопасность, бизнес (банковская сфера, контакт-центры)

Доступ: платный

Стоимость: по запросу

<https://www.nemesysco.com>

5. Технологии для анализа эмоций по голосу «Empath» / «Web Empath API» / «SDK Empath API» / «Beluga Box» / «Emo Value Generator»

Производитель: компания «Empath Inc.»

Страна производитель: Япония, г. Токио (Japan, Tokyo)

Назначение: бизнес (банковская сфера, контакт-центры)

Доступ: платный

Стоимость: по запросу

<https://www.webempath.com>

6. Технология аналитики речи «SmartSpeech»

Производитель: компания ПАО «Сбербанк»

Страна производитель: Россия, г. Москва

Назначение: бизнес (банковская сфера, контакт-центры)

Доступ: платный

Стоимость: по запросу

<https://sberdevices.ru/smartspeech/>

В рамках представленного обзора наибольший интерес представляют алгоритмы и способы обработки речевых сигналов и классификации отрицательных, нейтральных и положительных психоэмоциональных состояний человека. Однако в силу коммерческой тайны производители не предоставляют данную информацию. Подробное описание перечисленных выше продуктов можно найти на сайтах производителей и в приложении 1, 2 к данной заявке.

План коммерциализации проекта: Проект реализуется по модели «Software as a Service» (SaaS) - программное обеспечение как услуга - одна из форм облачных вычислений, модель обслуживания, при которой подписчикам предоставляется готовое прикладное программное обеспечение, полностью обслуживаемое провайдером. Поставщик в этой модели самостоятельно управляет приложением, предоставляя заказчикам доступ к функциям с клиентских устройств, как правило, через мобильное приложение или веб-браузер. Главное преимущество данной модели - это отсутствие затрат, связанных с установкой, обновлением и поддержкой оборудования и программного обеспечения.

Типы клиентов:

- компании (частные и государственные), для которых предоставляется программное обеспечение;
- разработчики, которым предоставляются все инструменты для

реализации аналогичных проектов по обнаружению и классификации психоэмоциональных состояний по речевым сигналам.

Планируется реализации двух бизнес-моделей:

- b2b, b2g (создание готовых решений для бизнеса и государства, прибыль с интеграции и последующей абонентской платы; предоставление инструментов для создания аналогичных проектов прибыль с последующей абонентской платы);
- b2c (создание готовых решений с ограниченным функционалом для конечных пользователей);

Рынок систем обнаружения и классификации эмоций активно развивается. По оценкам ряда экспертных и маркетинговых компаний, мировой рынок продемонстрирует среднегодовой рост в 27,4 % и достигнет планки 29,1 млрд. долларов к 2022 году.

По опыту множества конкурентов эффективна будет система ежегодного лицензирования, стоимость лицензии будет составлять 9690 Р/год. В настоящий момент стоимость разрабатываемой технологии обнаружения и классификации психоэмоционального состояния человека по речевым сигналам определена не в полной мере. Корректно оценить стоимость разрабатываемой технологии, а также объем рынка, планируется по завершению НИР.