

ЧАТ-БОТ ДЛЯ ЛЮДЕЙ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

А. С. Круглик

kruglik.a.s@mail.ru

ФГБОУ ВО «Уфимский государственный авиационный технический университет» (УГАТУ)

Аннотация. Развитие информационных технологий должно приносить пользу людям. Особенно людям с ограниченными возможностями здоровья. Одна из самых популярных форм взаимодействия человека и цифрового устройства – диалоговая. Но информационные системы практически всегда разрабатываются без учета ограниченных возможностей здоровья пользователей с особенностями – инвалидов. Поэтому важной задачей является разработка приложения, удовлетворяющее подавляющую часть требований комфортного взаимодействия и нацеленное на помощь в решении некоторых возникающих задач у людей с ограниченными возможностями здоровья.

Ключевые слова: чат-бот; люди с ограниченными возможностями здоровья; искусственный интеллект; роботизированная автоматизация процессов.

ВВЕДЕНИЕ

Согласно статистике, в России каждый 12-й человек имеет инвалидность, а в мире – каждый 7-й [1, 2]. По внешнему виду человека не всегда можно понять является ли он инвалидом или нет. Кажется, что в окружении мало людей с ограниченными возможностями здоровья, или вовсе их нет. И многие считают, что утверждение о наличии сегрегации общества (разделение людей в повседневной жизни) является преувеличенным [3]. И их мнение не изменится, пока они не проведут некоторое время с инвалидами.

Сегрегация существует, хоть она и не всегда видна. Развитие новых технологий должно учитывать требования особой группы общества – людей с ограниченными возможностями здоровья. Но во многих случаях адаптация под эти требования осуществляется в качестве дополнения к существующим технологиям, а не чем-то интегрированным с самого начала [4]. Программное обеспечение, направленное на удовлетворение хотя бы части потребностей людей с ограниченными возможностями здоровья, может избавить последних от многих трудностей при взаимодействии с другими людьми, поисковыми системами, сайтами и цифровыми устройствами. Одной из такой программ является чат-бот, являющийся технологией виртуального общения.

Технологии используются не в полной мере, когда дело доходит до инвалидности. Многие компании в действительности не учитывают мнение людей с ограниченными возможностями здоровья. Многие средства реабилитации, такие как инвалидная коляска, которая может помочь людям стоять, по словам исследователя в области изучения инвалидности в Университете Висконсин-Мэдисон Эми Гейты, на самом деле не столько вспомогательные технологии, сколько корректирующие технологии [4]. Таким образом, такие корректирующие технологии направлены на то, чтобы изменить людей с ограниченными возможностями здоровья так, чтобы они соответствовали окружающему миру, а не наоборот.

ПОПУЛЯРНОСТЬ МЕССЕНДЖЕРОВ И ЧАТ-БОТОВ

Мессенджер (программа для мгновенного обмена сообщениями) является одним из самых популярных приложений [5]. Также в исследовании «Медиапотребление в России – 2020» от аудиторской компании «Deloitte» было заключено, что 42% респондентов стали

пользоваться мессенджерами чаще, чем в прошлом году. Благодаря популярности мгновенного обмена сообщениями стали популярны и чат-боты, в которых форма общения тоже представляет собой чат.

В 2019 году «Deloitte» провела исследование, посвященное ожиданиям работников от чат-ботов. В нем было спрогнозировано, что к 2022 году ежедневно с чат-ботами будут взаимодействовать 70% работников, занимающиеся умственным трудом. А в 2017 году консалтинговая и исследовательская компания «Gartner» в своем исследовании «Лучшие стратегические прогнозы Gartner на 2018 год и далее» сделала предположение, что к 2021 году более 50% предприятий будут ежегодно тратить на создание ботов и чат-ботов больше, чем на разработку традиционных мобильных приложений.

ЦЕЛИ И ЗАДАЧИ

Целью исследования – достичь в чат-бот приемлемой точности распознавания в сообщении пользователя тематики, тональности и контекста, прогнозирования поведения и последующих сообщений собеседника, а также реализовать в чат-боте сбор информации о пользователе, образование полезных намерений и адаптацию к особенностям лиц с ограниченными возможностями здоровья.

Задачи исследования:

1. Разработать алгоритмы, сконструировать и модели искусственного интеллекта, а также создать методы из объединения их в информационную систему.
2. Адаптировать алгоритмы к большинству видов заболеваний людей с ограниченными возможностями здоровья и обучить и/или дообучить модели на данных, соответствующих формам и видам общения инвалидов (речь с отклонениями, язык жестов).
3. Сформировать архитектуру программного обеспечения и реализовать ее. В состав программного обеспечения должны входить следующие компоненты: система для взаимодействия с пользователем, система обработки сообщений и поддержки принятия решений, а также, при необходимости, система для автоматизации обмена данными с внешними информационными системами.
4. Удостовериться в полезности разработанной информационной системы для людей с ограниченными возможностями здоровья и осуществить исправление и доработку алгоритмов и моделей.

ОПИСАНИЕ ЧАТ-БОТА

Чат-бот позиционируется как интеллектуальный онлайн-помощник для людей с ограниченными возможностями здоровья. Он будет давать полезные советы пользователям – формировать намеренья. Настроение собеседника и паттерны его поведения будут определяться чат-ботом во время общения. Для облегчения ввода текстового сообщения будет доступна функция прогнозирования оставшейся части сообщения, чтобы человеку с нарушениями моторики не пришлось набирать текст полностью. На основе сохраненных данных будет формироваться профиль человека с общей и временной информацией.

Чат-бот будет сохранять информацию о пользователе, получаемую из его сообщений. Формат сообщений может быть любой: текст, звуковое сообщение, демонстрация языковых жестов. Для отправки текстового сообщения пользователю необходима клавиатура, для отправки голосового сообщения – микрофон, а для взаимодействия с чат-ботом с помощью языковых жестов – веб-камера. Эффективность системы зависит от качества микрофона и веб-камеры при их необходимости.

Вне зависимости от формата сообщения оно будет преобразовано в текст. Преобразование нетекстового формата сообщения в текст осуществляется с помощью алгоритмов по распознаванию голоса людей, в том числе тех, у которых имеется нарушение речи, либо с помощью алгоритмов распознавания языка жестов. Далее текст анализируется с помощью методов по обработке естественного языка (Natural Language Processing, NLP).

Для функционирования чат-бота будут применены модели искусственного интеллекта, основанные на машинном обучении. Цель применяемых моделей – распознавание изображений языковых жестов и аудиопотока голоса с нарушениями речи. В случае с распознаванием речи планируется в качестве варианта реализации взять модель, распознающую голос без нарушения речи. А в дальнейшем дообучать эту модель на персональном голосе каждого пользователя при такой необходимости с помощью подготовленных и внедренных инструментальных средств. Таким образом, система будет адаптирована к как можно большему числу пользователей.

Чат-бот будет иметь подсистему роботизированной автоматизации процессов, позволяющую искать и извлекать информацию с различных источников, например, сайтов, а также отправлять данные туда. Кроме того, это позволит автоматически заполнять поля на форме сайта на основе данных, предоставленных чат-боту от пользователя.

ОБЗОР РАБОТ

Авторы в своей публикации [6], посвященной аудиовизуальному распознаванию речи людей с нарушением речи, описали применение мультимодального распознавания речи. У предлагаемой системы, как и у типичной аудиовизуальной системы распознавания речи, в качестве входных данных принимает видео. На видеозаписи изображен говорящий, который произносит слова. Видео делится два рабочих потока: аудиозапись и соответствующий упорядоченный массив изображений. Затем выполняются процессы линейной интерполяции и интеграции аудиовизуальных признаков. На финальном этапе применяется классификатор на базе скрытой марковской модели для классификации слов к их соответствующим классам с использованием как звуковых, так и визуальных признаков. В качестве звуковых признаков использовались Мел-частотные кепстральные коэффициенты (Mel-frequency cepstral coefficients, MFCC). Предложенная система тестировалась на людях с нарушениями речи. Перед извлечением визуальных признаков происходила предобработка:

1. поиск фрагмента изображения, на котором изображено лицо;
2. поиск фрагмента изображения лица, на котором изображен рот;
3. масштабирование изображения рта до размера 128×128 пикселей;
4. преобразование цветного изображения рта в оттенки серого.

Для двух первых перечисленных шагов применялся алгоритм Виолы — Джонса. На выборке в 1 900 случайных изображений спикера точность поиска фрагментов лица и рта составляла 100% и 96.5% соответственно. После предобработки визуальные признаки были получены с помощью дискретного косинусного преобразования (Discrete Cosine Transform, DCT). Авторы провели эксперименты, показавшие увеличение точности после добавления визуальных признаков в 3-7% в зависимости от качества видеозаписывающего устройства. В эксперименте приняли участие 5 человека с нарушениями речи.

Публикация [7] также посвящена мультимодальному распознаванию речи с нарушениями артикуляции. Авторы исследовали Активные модели внешнего вида (Active Appearance Models, ААМ). Их разработанная система основана на акустических и визуальных признаках. Аудиозапись конвертируется в множественные акустические кадры (МАФ) в качестве акустической динамической характеристики. Вектор акустических признаков строится путем объединения МАФ, а для уменьшения размерности используется PCA. Кепстр (вид гомоморфной обработки сигналов) и полученная комбинация признаков используется в качестве акустического признака. Для получения визуальных признаков сначала область изображения с лицом извлекается из входной последовательности изображений с помощью алгоритма AdaBoost на основе Нагг-подобных характеристик, и происходит установка на начальные позиции поиска. Затем используется ААМ для обнаружения детальных точек и направления лица автоматически без ручного указания. Далее, векторы визуальных признаков извлекаются из мозаичного изображения области губ с помощью DCT, а для уменьшения размерности используется метод главных компонент (Principal Component Analysis, PCA). На последнем

этапе обучаются скрытые марковские модели отдельно на акустических и визуальных признаках. Аудиовизуальная интеграция позволила не только распознавать без помех, но и подавлять колебания речи. Предложенный метод позволил повысить уровень распознавания на 7,3% (с 51,6% до 58,9%) по сравнению с методом, использующим только распознавание аудиозаписи. В тестировании приняла участие только одна персона.

В статье [8] проводится обзор распознавания языка жестов с помощью машинного обучения. Кроме того, был представлен метод распознавания американского языка жестов с помощью машинного обучения. Это подход к решению проблем, с которыми сталкиваются люди с нарушениями слуха и речи. Алгоритм состоял из двух основных компонентов: анализа жестов с изображений и классификации изображений. В данном исследовании набор данных содержит изображения американского языка жестов и 24 метки статических жестов от букв А до У, за исключением J. На каждый класс приходится в среднем 100 изображений. Для тестирования было использовано 20% изображений из каждого класса. Классификация выполнялась с помощью нескольких моделей, однако наилучший результат показала модель сверточной нейронной искусственной сети Inception v3 – 90% точности.

Работа [9] описывает примененные алгоритмы машинного обучения для распознавания языковых жестов Индии. Методология исследования заключалась в разделении задачи на три подзадачи (этапа). Первый этап – сегментация фрагментов с кожей лица на собранном наборе изображений; оставшиеся фрагменты изображений рассматривались как шумы. На втором этапе происходило извлечение соответствующих признаков из сегментированных изображений кожи лиц. Эти признаки необходимы были для третьего этапа. На третьем этапе использовались извлеченные признаки в качестве входных данных для обучения различных моделях машинного обучения с учителем, а затем осуществлялась классификация с помощью обученных моделей. Для обучения использовалось 70% данных, а для тестирования моделей – 30%. Наилучшим оказался метод опорных векторов (Support Vector Machines, SVM), показавший точность в 100%.

ОБСУЖДЕНИЕ

Существует методы и модели, позволяющие обеспечить эффективную работу чат-бота, предназначенного для людей с ограниченными возможностями здоровья. Однако, для этого необходимо учитывать мнение пользователей. Как считает исследователь в области изучения инвалидности в Университете Висконсин-Мэдисон Эми Гейта, лучший способ оценить приверженность стартапа или технологической компании вспомогательным технологиям для людей с ограниченными возможностями здоровья – спросить, нанимают ли они дизайнеров-инвалидов. И активно ли они прислушиваются к их мнению? Насколько хорошо работают эти решения? Какие отзывы нравятся пользователям? По словам Гаэты, ключевым моментом является установление отношений с сообществом и обратная связь [4]. Технологические компании должны работать с сообществом, а не производить для них. Наиболее успешные решения, утверждает она, приходят, когда людей с ограниченными возможностями здоровья приглашают за стол с самого начала.

ЗАКЛЮЧЕНИЕ

Рассмотренные методы и модели могут помочь сделать жизнь инвалидов немного легче. Чат-бот, безусловно, не сможет заменить реального ассистента, помощника, друга. Но благодаря методам искусственного интеллекта интеллектуальный онлайн помощник поможет людям с ограниченными возможностями здоровья решить их некоторые задачи и проблемы.

СПИСОК ЛИТЕРАТУРЫ

1. Тема дня 2020 года: на пути к инклюзивному и устойчивому миру после пандемии с учетом потребностей инвалидов. Основные сведения [Электронный ресурс]. – Организация объединенных наций. – Режим доступа: <https://www.un.org/ru/observances/day-of-persons-with-disabilities/background>. – Дата обращения: 07.02.2022.

2. Численность инвалидов [Электронный ресурс]. – Федеральный реестр инвалидов. – Режим доступа: <https://sfri.ru/analitika/chislennost/chislennost?territory=00000000>. – Дата обращения: 07.02.2022.
3. Could AI help make the law more accessible for disabled people? [Электронный ресурс]. – ABC Radio National. 2021. – Режим доступа: <https://www.abc.net.au/radionational/programs/lawreport/disability-law-technology/13655382>. – Дата обращения: 07.02.2022.
4. Lockwood D. Tech for disabled people is booming around the world. So where's the funding? [Электронный ресурс]. – Rest of World. 2021. – Режим доступа: <https://restofworld.org/2021/innovating-for-the-worlds-largest-minority/>. – Дата обращения: 07.02.2022.
5. Медиапотребление в России – 2020 [Электронный ресурс]. 2020. Режим доступа: <https://www2.deloitte.com/content/dam/Deloitte/ru/Documents/technology-media-telecommunications/russian/mediapotrebienie-v-Rossii-2020.pdf>. Дата обращения: 07.02.2022.
6. Salama E. S., El-Khoribi R., Shoman M. Audio-Visual Speech Recognition for People with Speech Disorders // International Journal of Computer Applications. 2014. Vol. 94. № 2. P. 51-56.
7. Miyamoto C., Komai Y., Takiguchi T. Aiki Y., Li I. Multimodal speech recognition of a person with articulation disorders using AAM and MAF // 2010 IEEE International Workshop on Multimedia Signal Processing. 2010. P. 517-520.
8. Srivastava A., Malik V. Review on Sign Language Detection Using Machine Learning // Journal of critical reviews. 2020. Vol. 7. № 10.
9. Shirbhate R. S., Shinde V. D., Metkari S. A., Borkar P. U., Khandge M. A. Sign language Recognition Using Machine Learning Algorithm // International Research Journal of Engineering and Technology. 2020. Vol. 7. № 3. P. 2122-2125.

ОБ АВТОРЕ

КРУГЛИК Андрей Сергеевич, аспирант 1-го курса АВИАТ.

METADATA

Title: Chatbot for disabled people.

Author: A. S. Kruglik

Affiliation: Ufa State Aviation Technical University (UGATU), Russia.

Email: kruglik.a.s@mail.ru

Language: Russian.

Source: Molodezhnyj Vestnik UGATU (scientific journal of Ufa State Aviation Technical University), no. 1 (26), pp. 42-46, 2022. ISSN 2225-9309 (Print).

Abstract. The development of information technologies should benefit people. Especially for disabled people. One of the most popular forms of interaction between a person and a digital device is dialog. But information systems are almost always developed without taking into account the limited health capabilities of disabled people. Therefore, an important task is to develop an application that satisfies the vast majority of the requirements for comfortable interaction and is aimed at helping disabled people to solve some of the problems that arise.

Key words: chatbot; people with disabilities; artificial intelligence; robotic automation of processes.

About author:

KRUGLIK, Andrew Sergeevich, postgraduate student 1 year, Ufa State Aviation Technical University.