

ПРИЛОЖЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Адилбеков Еркебулан

магистрант, НАО «Восточно-Казахстанский университет имени Сарсена Аманжолова»,
Казахстан, г. Усть-Каменогорск

Забиров Алмас Олжасович

магистрант, НАО «Восточно-Казахстанский университет имени Сарсена Аманжолова»,
Казахстан, г. Усть-Каменогорск

Жолболдин Кайрат Бериккалиевич

магистрант, НАО «Восточно-Казахстанский университет имени Сарсена Аманжолова»,
Казахстан, г. Усть-Каменогорск

Тлебалдинова А.С.

научный руководитель, PhD доктор, НАО «Восточно-Казахстанский университет имени
Сарсена Аманжолова», Казахстан, г. Усть-Каменогорск

Сыздыкпаева А.А.

научный руководитель, канд. техн. наук, ТОО «Rating», Казахстан, г. Усть-Каменогорск

ARTIFICIAL INTELLIGENCE APPLICATION

Yerkebulan Adilbekov

Master's student of non-profit joint stock company "S.Amanzholov University", Kazakhstan, Ust-Kamenogorsk

Almas Zabirow

Master's student of non-profit joint stock company "S.Amanzholov University", Kazakhstan, Ust-Kamenogorsk

Kairat Zholboldin

Master's student of non-profit joint stock company "S.Amanzholov University", Kazakhstan, Ust-Kamenogorsk

A. Tlebaldinova

PhD doctor, non-profit joint stock company "S.Amanzholov University", Kazakhstan, Ust-Kamenogorsk

A. Syzdykpaeva

Аннотация. В статье кратко изложены материалы о разработке голосового ассистента, информационной системы учета поставок, система документооборота медицинского учреждения, которые собраны авторами в процессе разработки программных приложений на основе методов интеллектуального анализа. Авторы определили недостатки и последовательность разработки таких приложений.

Abstract. The article summarizes the materials on the development of a voice assistant, an information system for accounting supplies, a document management system of a medical institution, which are collected by the authors in the process of developing software applications based on intellectual analysis methods. The authors identified the shortcomings and the sequence of development of such applications

Ключевые слова: искусственный интеллект, bNLU, TTS, NER, ASR, Google Assistant, Яндекс, голосовой помощник, машинное обучение, управление данными, SOJA.

Keywords: artificial intelligence, blu, TTS, NET, ACER, Google Assistant, Yandex, voice assistant, machine learning, data management, SOJA.

Достижения информационных технологий отражаются в улучшении приборов при использовании их в производстве. Раздел искусственного интеллекта помогает нам общаться в режиме реального времени вне зависимости от нашего местонахождения, учиться онлайн, послушать информацию в любое время, упорядочить график поставок, оптимизировать документооборот. Актуальность нашего исследования обусловлена тем, что информационные технологии являются частью в лингвистике, логистике.

С разработками Интернета вещей [1-5] и машинного обучения компьютер самостоятельно изучает привычки пользователя, его шаблоны поведения, действия и становится ассистентом. IT-корпорации создают голосовые ассистенты, логистические информационные системы. Например, Google Assistant от Google, Siri от Apple, Amazon Echo (Alexa) от Amazon, Cortana от Microsoft, Алиса от Яндекса. Из-за модернизации интеллектуальные помощники улучшают алгоритмы распознавания и обработки речи, позволяют спрогнозировать спрос, упорядочить информацию о клиентах. Понятие цифровой ассистент означает систему автоматизации взаимодействия с пользователем (клиентом или сотрудником компании), реализованная на основе искусственного интеллекта (AI) в диалоговом формате (чат-бот, управление данными, самообучающиеся алгоритмы). Сервис осуществляет текстовые и голосовые консультации. Пользователь отправляет запрос в чате в свободной текстовой или голосовой форме, а система предоставляет быстрый подходящий ответ.

Перечислим технологии, используемые в системах интеллектуальных голосовых ассистентов: Voice technology (Voice Activation, ASR Automatic Speech Recognition, Text-To-Speech (TTS), Brain technology (Voice Biometrics, Dialog Management, Natural Language Understanding (NLU), Named Entity Recognition (NER). Эти голосовые ассистенты отличаются, так как разработчики используют специфичные подходы и разные алгоритмы: отличие по качеству распознавания для разных языков, выполнение некоторых задач без дополнительных объяснений. Универсального решения для выполнения любой задачи распознавания голосовым ассистентом нет [6-8, 11]. Общий принцип построения голосовых ассистентов остается одинаковым. Технологии активации по голосу (Voice Activation), автоматическое распознавание речи (Automatic Speech Recognition), синтез речи (Text-To-Speech), голосовая биометрия (Voice Biometrics, распознавание пола или возраста говорящего), диалоговый менеджер (Dialog Manager), понимание естественного языка (Natural Language Understanding), распознавание именованных сущностей (Named Entity Recognition) [8-10, 12].

Например, структура голосового ассистента «Алиса» от компании Яндекс, который позволяет решать общие задачи пользователей (поиск информации в интернете, поиск мест на карте, прокладывание маршрутов, сообщение прогноза погоды, может развлекать пользователя и т.д.). использует облачные средства компании «Яндекс» через API посредства сети Интернет.

На первом этапе активации произнесенной пользователем ключевой фразы ассистент анализирует наличие ключевых слов, распознает, переходит в активный режим. Пользователь произносит текст, чтобы объяснить помощнику что хочет сделать. Система распознавания (Automatic Speech Recognition) превращает текст в N-лучших гипотез того, что сказал пользователь. Система распознавания естественного языка (Natural Language Understanding) превращает текст в N-лучших вариантов понимания фразы пользователя, диалоговый менеджер интерпретирует и классифицирует эти фразы и определяет, что сделать на основе полученной информации.

После получения необходимых данных система производит процесс возвращения информации пользователю. Система генерации естественного языка (Natural Language Generation) генерирует текст ответа пользователю, далее система генерации голоса (Text-To-Speech) на основе обученных моделей генерирует звуковую информацию, которая объявляет пользователю в качестве ответа. Может вместо ответа происходить действие на мобильном телефоне или компьютере, например запуск приложения или поиска информации в поисковой системе. Диалоговый менеджер - это важный компонент голосового ассистента. Существуют простые сценарии, на которые можно сразу извлечь из NLU модели и воспроизвести через NLG. Сложные сценарии основываются на форме. Форма повторяет форму обычного пользовательского интерфейса (UI), где существуют обязательные и необязательные поля для заполнения. В таких сценариях используется подход Form Filling, в процессе диалога с пользователем форма заполняется ответами, причем эти ответы могут заполняться как пользователем, так и самой системой на основе информации, которую она сможет получить у пользователя. Процесс заполнения интеллектуальный и сама система часть полей может заполнить самостоятельно. После заполнения форма отправляется на обработку, в котором принимается решение о точном ответе на запрос пользователя или переключение на нейронный диалог.

Ассистенты не получили широкого распространения из-за существования ряда ограничений. Практика разработки ассистента показала нам, что можно осуществить алгоритм решения общих задач, приложение зависит от Интернета, облачных сервисов, сложно объединить функции со сторонними сервисами, незащищены персональные данные пользователя. Для систем умного дома ассистент справляется с многими функциями. Значит в медицине или в сфере безопасности ассистент не применяется. Но исследования по разработке ассистента как выражения искусственного интеллекта приводит к пониманию того, что универсальные системы не смогут реализовать запросы потребителя. Создание специализированных персональных помощников с узкими задачами, с минимальными требованиями к инфраструктуре является актуальной задачей.

Метод создания интеллектуального голосового ассистента для специфических задач взаимодействия основан на использовании существующих систем. Например, проект PocketSphinx как основное средство распознавания голоса. PocketSphinx - это инструментальный для автоматического распознавания голоса [14], который неплохо работает на различных маломощных встраиваемых системах (Raspberry Pi), а также является кроссплатформенным [13]. Для генерации голоса был выбран наиболее используемый движок Festival, который работает на операционных системах Linux и имеет достаточно неплохие характеристики генерации голоса. Постановка задачи машинного обучения для интеллектуального анализа намерений поставлена так: выходными характеристиками может быть набор различных классов (т.е. N-лучших гипотез намерений пользователя), задание представляет собой задачу мультиклассовой классификации, где в одном ответе могут содержаться метки из разных классов. Например, в качестве ответа может быть «on; light; bathroom», где каждая часть ответа принадлежит одному из нескольких классов.

Сбор необходимых данных - самый важный процесс, который позволит наиболее точно

прогнозировать результаты. Для сбора данных необходимо составить таблицу синонимов и различных вариантов произношения ключевых слов, по которым система сможет строить прогнозы. Слова и синонимы указываются в формате, в котором они обычно употребляются в разговорной речи. Программа определяет список ответов, которые необходимо прогнозировать на основе входных данных. Следующий этап – построение словаря, из которого создается обучающая выборка. Так как большая часть алгоритмов машинного обучения оперирует числовыми данными, мы должны сопоставить в соответствие каждому слову в словаре определенное уникальное в пределах словаря число. Целесообразно также будет уменьшить размер словаря с помощью стемминга [13].

Алгоритмы искусственного интеллекта снижает риски участников поставок, например в создании контрактов автоматизированная система аналитики отмечает идентификацию поставщиков, их ответственность, требования к виду товаров, защита от фрода. Алгоритмы машинного обучения зависят от задачи распределения данных в полученной обучающей выборке. Например, для обучения ассистента используют библиотеку машинного обучения Scikit-Learn для языка программирования Python. Гибкость программного обеспечения заложена в базовой архитектуре, например, аспектно-ориентированная (сервисно-ориентированная) архитектура (SOA). Она может применяться в сложных процессах, программное обеспечение расширяет спектр своих функциональных возможностей. Компания PSI Logistics усовершенствовала этот современный подход на основе SOA, создав сервис-ориентированные архитектуры на базе Java (SOJA). Благодаря SOJA разнородные ИТ-инфраструктуры приобретают способность к интеграции и динамику. Архитектуры SOJA обеспечивают согласованную работу разнородных ИТ-систем и использование единых решений в области мониторинга в виде различных, специфических для отрасли и/или функций программных продуктов, связанных друг с другом аналогично сети с шинной архитектурой. Это позволяет различным системам работать рядом и совместно друг с другом и упрощает решение проблемы интерфейсов.

Список литературы:

1. Поляков Е.В. Исследование методов машинного обучения для анализа и принятия решений на основе данных интернета вещей // В кн.: Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В. Арменского / Под общ. ред.: А.Н. Тихонов, С.А. Аксенов, У.В. Аристова, Л.С. Восков, А.А. Елизаров, М.В. Карасев, В.П. Кулагин, Ю.Л. Леохин, А.Б. Лось, И.С. Смирнов, Н.С. Титкова. – М.: МИЭМ НИУ ВШЭ, 2017. – С. 66-68.
2. Восков Л.С. Социальные сети WEБа вещей // В кн.: XXI Международная студенческая школа-семинар «Новые информационные технологии». Тезисы докладов / Отв. ред.: А.Н. Тихонов, В.Н. Азаров, Ю.Л. Леохин, Н.С. Титкова, С.С. Фомин. – М.: МИЭМ НИУ ВШЭ, 2013. – С. 53-58.
3. Восков Л.С. Интернет вещей // В кн.: Новые информационные технологии. Тезисы докладов XX международной студенческой конференции-школы-семинара / Науч. ред.: В.Н. Азаров, С.А. Митрофанов, Ю.Л. Леохин, Н.С. Титкова. – М.: МИЭМ, 2012. – С. 89-94.
4. Ролич А.Ю., Мартюкова Е.С., Арзамасова А.И. Интернет вещей: актуальность, решения, проблематика // В кн.: Научно-техническая конференция студентов, аспирантов и молодых специалистов НИУ ВШЭ им. Е.В. Арменского. Материалы конференции - М.: МИЭМ НИУ ВШЭ, 2015. - С. 140-142.
5. Dvornikov A. QoS Metrics Measurement in Long Range IoT Networks //Business Informatics (CBI), 2017 IEEE 19th Conference on. - IEEE, 2017. - Т. 2.- С. 15-20.
6. Dempsey P. The teardown: Google Home personal assistant //Engineering & Technology. - 2017. - Т. 12. - № 3. - С. 80-81.
7. Chung H. et al. Alexa, Can I Trust You? //Computer. - 2017. - Т. 50. - № 9. - С. 100-104.
8. López G., Quesada L., Guerrero L. A. Alexa vs. Siri vs. Cortana vs. Google Assistant: A

Comparison of Speech-Based Natural User Interfaces //International Conference on Applied Human Factors and Ergonomics. - Springer, Cham, 2017. - C. 241-250.

9. Arriany A.A., Musbah M.S. Applying voice recognition technology for smart home networks // Engineering & MIS (ICEMIS), International Conference on. - IEEE, 2016. - C.1-6.

10. Caranica A. et al. Speech recognition results for voice-controlled assistive applications //Speech Technology and Human-Computer Dialogue (SpeD), 2017 International Conference on. - IEEE, 2017. - C. 1-8.

11. Assefi M. et al. An experimental evaluation of apple siri and google speech recognition //Proccedings of the 2015 ISCA SEDE. - 2015.

12. Natural Language Understanding Lecture 10: Introduction to Unsupervised Part-of-Speech Tagging // www.inf.ed.ac.uk URL: https://www.inf.ed.ac.uk/teaching/courses/nlu/lectures/nlu_l10_unsuptag1.pdf.

13. Lovins J.B. (1968). Development of a stemming algorithm. Mech. Translat. & Comp. Linguistics, 11 - 22-31 p.