

УДК 004.85

СИСТЕМА ОПТИМИЗАЦИИ РАБОТЫ КЛИЕНТСКИХ СЛУЖБ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

Алисултанова Э.Д., Моисеенко Н.А., Тасуев У.Р., Юсупова Р.В.

*ФГБОУ ВО «Грозненский государственный нефтяной технический университет
имени академика М.Д. Миллионщикова», Грозный, e-mail: esmira59@mail.ru,
namoi@mail.ru, tasuevusman@gmail.com, yu.raikhann@mail.ru*

В данной статье рассматриваются инструменты для создания системы распознавания речи малораспространённых языков, в частности чеченского языка. Технологии распознавания речи – одна из актуальных проблем современного цифрового общества. Данные технологии в программной реализации дают возможность вычислительным системам воспринимать устную речь, интерпретировать ее и генерировать в текстовом формате. Инструментарий распознавания речи уже стала частью нашей повседневной жизни, но пока все еще ограничивается относительно простыми отработанными командами. Учитывая современные тенденции, технология распознавания речи будет быстрорастущим (и меняющим мир) подмножеством обработки сигналов на долгие годы. Программная разработка предназначена для автоматизации сбора входных данных и их качественной обработки для реализации системы эффективной обратной связи с потребителем товаров и услуг в call-центре. Поставленная задача разработки программного приложения для записи и проверки аудиозаписей на чеченском языке решена технологией машинного обучения в анализе больших аудиоданных. Описаны разработка интерфейса для анализа и мониторинга работы Call-центра, проектирование приложения сбора данных для записи и проверки аудиозаписей на чеченском языке.

Ключевые слова: машинное обучение, call-центр, сервер, чеченский язык, автоматизированные системы, распознавание речи, дуальная модель распознавания, большие данные, стенографирование, фреймворк, Python, обучение модели

NEURAL NETWORK-BASED CUSTOMER SERVICE OPTIMIZATION SYSTEM

Alisultanova E.D., Moiseenko N.A., Tasuev U.R., Yusupova R.V.

*Grozny State Oil Technical University named after Academician M.D. Millionshchikov, Grozny,
e-mail: e-mail: esmira59@mail.ru, namoi@mail.ru, tasuevusman@gmail.com, yu.raikhann@mail.ru*

This article discusses tools for creating a speech recognition system for minority languages, specifically the Chechen language. Speech recognition technologies are one of the pressing problems of the modern world of digital society. Software implementation of such technology enable computing systems to perceive spoken speech, interpret it and generate it in text format. Speech recognition tools have already become a part of our daily life, but are still limited to relatively simple, generic commands. Given current trends, speech recognition technology will be a fast-growing (and world-changing) subset of signal processing for years to come. The software under development is designed to automate the collection and processing of input data for the implementation of an effective feedback system with the consumer of goods and services in the Call Center. The set task of developing a software application for recording and checking audio recordings in the Chechen language is solved by machine learning technology in the analysis of large audio data. The development of an interface for analyzing and monitoring the operation of the Call-center, designing a data collection application for recording and checking audio recordings in the Chechen language are described.

Keywords: Machine learning, call center, server, Chechen language, automated systems, speech recognition, dual recognition model, big data, verbatim, framework, Python, model training

В последнее десятилетие наблюдается бурное развитие сферы машинного обучения в анализе больших данных. Различные организации внедряют технологии, основанные на машинном обучении. Они решают задачи автоматизации, ведения статистики и анализа деятельности. Крупные IT-компании предоставляют доступ к проприетарному программному обеспечению, функционирующему на их собственных серверах. Это является проблемой для организаций, в которых действуют особые требования к защите конфиденциальной информации. Кроме того, стоимость пользования такими услугами может быть

достаточно высокой, что может служить препятствием для использования этих технологий для малых организаций.

Другая проблема – это отсутствие системы распознавания малораспространённых языков, в частности чеченского языка. Для носителей таких языков нет возможности перевести записанную речь в текст. Пример ситуаций, при которых это критично – общение с туристами, коммуникация с людьми с ограниченными возможностями, работа региональных клиентских сервисов и т.д.

В данной статье будут рассмотрены инструменты для создания системы рас-

познавания речи на русском и чеченском языках, а также описана разработка интерфейса для анализа и мониторинга работы колл-центра.

Для сбора данных, необходимых при реализации услуг в call-центре, который является системой эффективной обратной связи с потребителем товаров и услуг (заказчиком), ставится задача разработки программного приложения для записи и проверки аудиозаписей на чеченском языке.

Технология распознавания речи позволяет компьютерам принимать устную речь, интерпретировать ее и генерировать из нее текст. Она уже стала частью нашей повседневной жизни, но пока все еще ограничивается относительно простыми командами. Учитывая современные тенденции, технология распознавания речи будет быстро развиваться (и меняющим мир) подмножеством обработки сигналов на долгие годы.

Для организации работы контакт-центров необходимо обеспечить отдел определенным оборудованием, в список которого входит: сервер, VoIP-шлюзы, GSM-шлюзы, IP-телефоны, гарнитура для операторов. Ключевая задача сервера в данном случае – это функция управления телефонными вызовами и распределения звонков между операторами центра. Все входящие и исходящие звонки предполагается осуществлять на основе внедрения в телекоммуникации серверного оборудования.

В случае если организация хочет проводить анализ и мониторинг содержания записей, она должна выделить персонал, который будет монотонно прослушивать записи и заполнять анкеты для каждого из них. Для этого необходимо использовать информационную систему для сбора этих анкет и анализа их результатов.

Другая значительная проблема – региональный язык. В данном случае чеченский язык. В настоящее время отсутствуют автоматизированные системы распознавания чеченской речи.

Существует потребность в разработке такой системы, которая должна соответствовать следующим требованиям:

- должна функционировать оффлайн;
- иметь дуальную модель распознавания;
- должна распознавать чеченский язык.

Система, описываемая в данной статье, решает одну из проблем – стенографирование аудиозаписей. Для анализа диалога требуется разработать алгоритмы для поиска и фильтрации диалогов, анализа корпоративной этики, появления конфликтной ситуации и др.

Для решения поставленной задачи нужно решить следующие задачи:

– определить систему для разработки системы распознавания и перевода речи в текст (проблема «Speech2Text») дуальной языковой модели – русской и чеченской;

– разработать решения для сбора аудиоданных чеченского языка;

– разработать приложение для мониторинга работы call-центра, а именно – стенографирование разговоров для дальнейшего анализа аудио.

Материалы и методы исследования

Среди систем, предназначенных для разработки системы перевода речи в текст, наиболее распространены практики применения следующих инструментов:

- CMU Sphinx;
- Kaldi;
- Deepspeech.

CMU Sphinx сейчас является крупнейшим проектом по распознаванию человеческой речи, в инструментарий которого входят следующие программы и библиотеки [1]:

– Kaldi – это набор инструментов для распознавания речи на основе математических функций, предназначена для исследований в области распознавания речи;

– Deepspeech – система, предпочтительная в применении для дуальной модели, так как предлагает «end to end» – модель для распознавания двух языков.

В связи с возникшей ситуацией необходимости формирования огромного набора данных Deepspeech была выбрана в качестве подсистемы для распознавания чеченской речи в случае, если русский язык распознан с низким значением confidence – то есть уверенности в адекватном отображении [2].

Для обучения систем распознавания речи используются наборы данных со следующей структурой: аудиофайл с частотой 16 кГц, моноканальный, в формате wav; документ в формате csv/tsv со следующими данными – id, название аудиофайла, id пользователя, дата, количество подтверждений.

В открытом доступе отсутствует набор данных на чеченском языке, в связи с чем возникла необходимость разработки программного приложения для сбора данных. Размещение данного приложения рекомендовано на собственном сервере в связи с высокой себестоимостью использования способа аренды больших объемов памяти.

Для мониторинга деятельности call-центра были определены ключевые показатели эффективности, к которым относятся:

– индекс потребительской лояльности (NPS) – отражает вероятность того, что по-

ребитель порекомендует товар или услугу другому человеку;

- среднее время обработки контакта (АНТ) показывает, насколько быстро операторы обрабатывают контакты с абонентами.

Скорость обработки данных, в свою очередь, тесно связана с экономическими показателями деятельности call-центра. В общем случае, чем быстрее операторы работают, тем меньшее количество таких сотрудников необходимо, и, следовательно, тем меньший фонд оплаты труда потребует для содержания такого офиса.

Среднее время обработки контакта АНТ складывается из четырех компонентов:

- Ring Time – время, отсчет которого начинается с момента, когда автомат распределения нагрузки извлекает вызов из очереди и направляет его на рабочее место;

- Talk Time – время разговора без учета времени постановки абонента на удержание;

- Hold Time – время постановки абонента на удержание;

- Wrap-Up Time – время поствызовной обработки. Оно начинается в момент завершения соединения и заканчивается, когда система открывает доступ к оператору для следующего контакта [3].

Кроме вышеперечисленных критериев, будет преимуществом в данной разработке определять следующие показатели, влияющие на эффективность работы системы: эмоциональная окраска записи, класс обращений, уровень корпоративной этики, определение признаков конфликтной ситуации.

Наиболее популярными решениями, с точки зрения оптимальной работы, для call-центров являются следующие:

- «Оки-Токи»;
- Mango Office;
- Oktell.

При выборе платформы для реализации программного решения были проанализированы современные решения этой проблемы и обоснованно сделан оптимальный выбор «Оки-Токи», который является профессиональным облачным сервисом для организации работы контакт-центра. Главным достоинством этого решения является то, что «Оки-Токи» работает из браузера без дополнительного программного обеспечения и оборудования.

Платформа рассчитана поддерживать до 500 операторов и 1000 линий одновременно, а также ее надёжность обеспечивается резервным копированием и хранением данных в разных странах мира. При этом интеграция осуществляется с внешними системами через технологию Webhooks и API.

Отметим также основной функционал, который будет использован в решении на-

шей проблемной задачи распознавания речи и обработки данных:

- мультисканальная обработка обращений;
- проектирование сценария для IVR и операторов, с помощью различных инструментов;

- поддержка черных, белых и VIP списков для каждого проекта;

- интеграции и сценарии работы в чатах (Facebook, WhatsApp и др.);

- обработка обращений в личном кабинете;

- автоматический callback по утраченным звонкам;

- импорт и экспорт телефонной базы в CRM;

- настройка разных воронок продаж под разные проекты, используя библиотеку шаблонов и сценариев;

- нормализация номеров;

- автоматическая проверка загруженных контактов на уникальность, а также на наличие ошибок;

- хранение истории взаимодействий с клиентом (звонки, переписки, чаты);

- управление доступами к контактам;

- кастомизация отчетов;

- экспорт отчетов.

MANGO OFFICE – функциональное и производительное решение среди российских облачных сервисов телефонии. Описание основных возможностей данного продукта для реализации описываемой системы можно сконцентрировать на следующем перечне:

- подключение номеров и других каналов связи – многоканальные номера, единые номера, FMC-связь, виджеты звонков, мультисканальный чат, рассылки;

- внутренние коммуникации и совместная работа – адресная книга сотрудников, мессенджер, конференции;

- работа с входящими обращениями – переадресация, создание групп сотрудников, распределение звонков по операторам, голосовое меню, голосовая почта;

- работа с исходящими звонками – адресная книга контрагентов, автоподстановка номера, плагины;

- стандартизация и автоматизация бизнес-процессов – готовые интеграции с популярными бизнес-приложениями, Google Apps, конструктор API, записи разговоров, облачное хранилище.

Кроме того, выбор Mango Office в данном исследовательском проекте обоснован его универсальным набором инструментов мониторинга, анализа и контроля: статистика и мониторинг, дашборд руководителя, доска мотивации, аналитика продаж, аналитика обслуживания, речевая аналитика.

Разработанная система распознавания чеченской речи в общем виде состоит из модулей, сформированных для исполнения функционала независимо друг от друга. Поэтому программные средства для каждого из них собственные, для которых можно привести краткую характеристику основных инструментов, использованных в данной разработке:

1) основной язык программирования – Python. Python с динамической типизацией и автоматическим управлением памяти [4];

2) JavaScript – это язык программирования, который в данном проекте использовался для добавления интерактивности веб-страницам;

3) Jupyter Notebook – удобный интерпретатор для использования языка Python, обычного текста, математических выражений и визуального представления данных;

4) Google Colab – Colaboratory, или сокращенно Colab – продукт компании Google Research, который особенно хорошо подходит для машинного обучения, анализа данных и образования;

5) Flask – фреймворк для создания веб-приложений на языке программирования Python, использующий набор инструментов Werkzeug, а также шаблонизатор Jinja2 [5];

6) CUDA – программно-аппаратная архитектура параллельных вычислений, которая позволяет существенно увеличить вычислительную производительность бла-

годаря использованию графических процессоров фирмы Nvidia [6].

Результаты исследования и их обсуждение

Сервер на данный момент выполняет 2 функции:

– обработка API запросов для распознавания речи;

– сохранение аудиозаписей для сбора набора данных.

При завершении обучения модели распознавания речи данную функцию следует отключить, поэтому доступ в Интернет при интеграции сервиса не требуется.

Для обучения моделей использовались 3 рабочие станции и облачные вычисления.

Обучение модели

На рабочую станцию для вычислений должна быть установлена операционная система Ubuntu версии 18.04 и библиотека cuDNN (рис. 1).

Для тестирования системы обучения нейронной сети следует использовать небольшой набор данных из 10 минут записей (рис. 2).

Созданы обработчики по следующим маршрутам: «/» – главная страница, отображает приветствие и счётчик записей (рис. 3). На этой странице размещается общая информация о проекте, для каких целей создано, а также текущий прогресс.

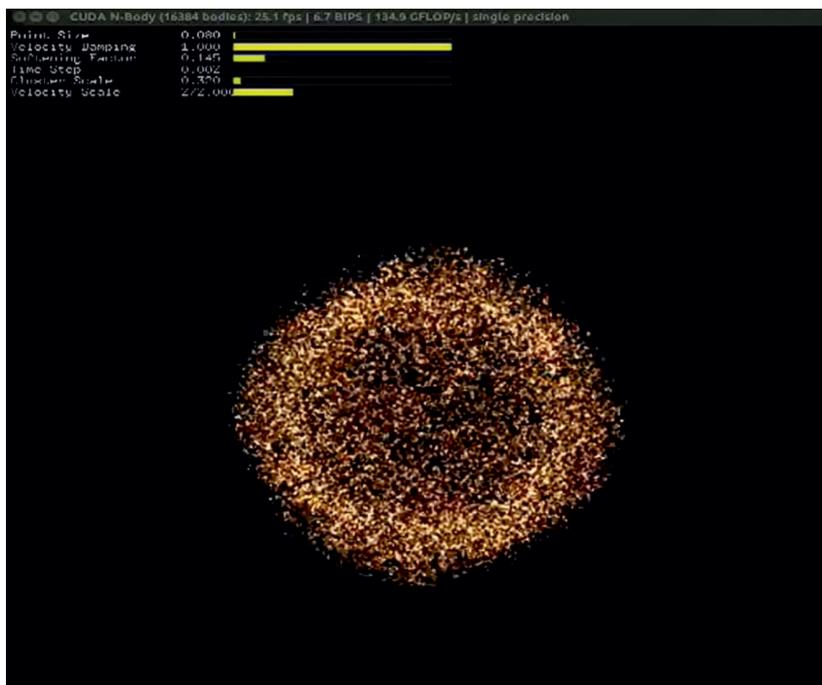


Рис. 1. Тест работы видеокарты в симуляции nbody

```
Test epoch | Steps: 349 | Elapsed Time: 0:05:46
Test on /content/drive/MyDrive/Colab_Data/clips/train.csv - WER
-----
Best WER:
-----
WER: 0.200000, CER: 0.095238, loss: 19.835438
- wav: file:///content/drive/MyDrive/Colab_Data/clips/58_2021_
- src: "нохчи мот ма хаь хьун"
- res: "нохчи мот ма х хьун"
-----
WER: 0.333333, CER: 0.083333, loss: 14.185292
- wav: file:///content/drive/MyDrive/Colab_Data/clips/36_2021_
- src: "ледер ма хил"
- res: "едер ма хил"
```

Рис. 2. Тест модели

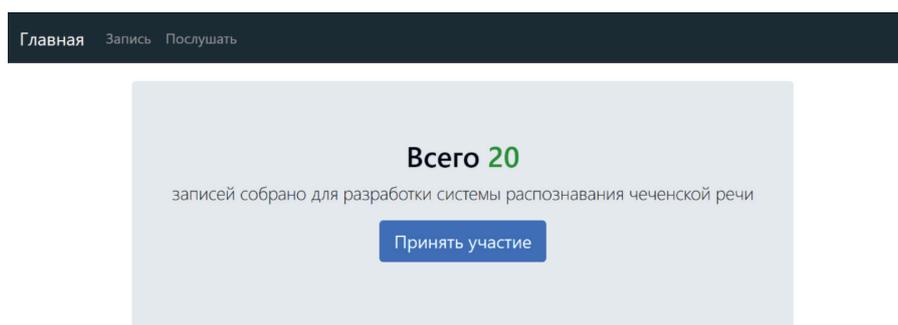


Рис. 3. Главная страница atdar

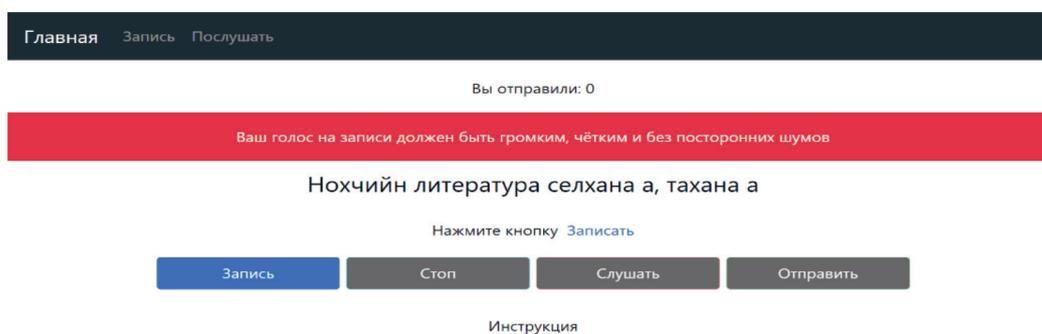


Рис. 4. Страница для записи

На рис. 4 представлен интерфейс базовой пользовательской страницы «Records», с необходимыми атрибутами: микрофоном, текстом и инструкцией.

Для работы описываемого в этой статье программного приложения рекомендовано использовать WPS-сервер и достаточно большой объём доступной памяти для создания и обработки банка данных. Для хранения файлов может декларироваться минимальный порог количественной характеристики объёма хранилища – от 1.5 Тбайта в RAID-массиве.

Заключение

Рассмотрим приложение для работы пользователя, а именно мониторинга и анализа работы call-центра, которое разработано с помощью фреймворка Django на языке Python. Программа позволяет загрузить информацию о новых записях в виде JSON-файла, данные из которых будут автоматически сохраняться в локальной базе данных SQLite.

Предусмотрена возможность поиска записей. В разделе «Общая статисти-

ка» выводится общая статистика диалогов по операторам. Показатели эффективности описывают уровень соблюдения корпоративной этики, которая описывается супервайзером в отдельном для этого разделе со списком слов и выражений. В отдельном окне прописывается показатель использования рекомендуемых слов и выражений, и все эти показатели приводятся по дням работы конкретного оператора.

Сводная статистика выводится в отдельном разделе, при которой в разделе «Мониторинг слов» формируются списки слов, а именно рекомендуемые слова и слова по предлагаемым услугам. На основе этого текста производится расчёт статистики. Отдельно можно посмотреть список активных операторов в разделе «Операторы», откуда можно совершить переходы на страницу для их редактирования.

Данное приложение уже выполняет основную функцию стенографирования и каталогизации бесед операторов. Мо-

дель, используемая для распознавания дуальной модели речи, будет улучшена при получении достаточного количества набора данных, для формирования банка которых запущено и функционирует веб-приложение.

Список литературы

1. Дронов В.А. Django: Практика создания Web-сайтов на Python. СПб.: БХВ-Петербург, 2021. 704 с.
2. Стюарт Р., Норвиг П. Искусственный интеллект: современный подход, 4-е издание. Том 1. Решение проблем: знания и рассуждения. М.: Диалектика-Вильямс, 2021. 704 с.
3. Вольский В.В. Веселова О.В. Золкина Н.К. Оператор call-центра: от найма до увольнения. Тверь: Софител, 2008. 200 с.
4. Гринберг М. Разработка веб-приложений с использованием Flask на языке Python. М.: ДМК-Пресс, 2016. 272 с.
5. Ляпунов С.М., Алымов А.С. Методология анализа текста с помощью составления частотных словарей по частям речи // Научный альманах. 2017. № 10–3 (36). С. 203–206.
6. Малявко А.А. Параллельное программирование на основе технологий openmp, cuda, opencl, mpi. М.: ЮРАЙТ-Восток, 2021. 136 с.