

В. А. Шкарубо
(ГГУ имени Ф. Скорины, Гомель)
Науч. рук. **Е. И. Сукач**, канд. техн. наук, доцент

РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ЗАПИСИ И ОБРАБОТКИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ РАСПОЗНАВАНИЯ РЕЧИ

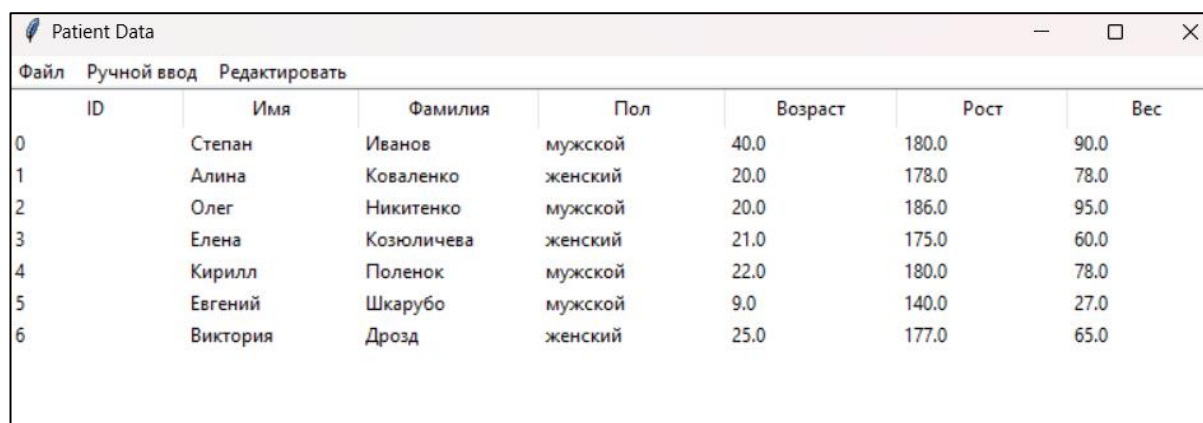
Введение. Автоматизация обработки данных играет важную роль в современных информационных системах, повышая эффективность работы и снижая вероятность ошибок. Одним из перспективных направлений является использование технологий распознавания речи, которые позволяют упростить взаимодействие с программами за счёт голосового управления [1].

В данной работе рассматривается разработка приложения для операционной системы Windows, которое обеспечивает запись и обработку данных с применением технологии распознавания речи. В основе системы лежит язык программирования Python и библиотеки Vosk, PyAudio, Pandas и Tkinter. Применение этих инструментов позволяет пользователям вводить данные голосом, минимизируя потребность в ручном вводе и повышая удобство работы с приложением.

Основные компоненты системы. Разрабатываемое приложение включает в себя следующие ключевые компоненты.

Модуль распознавания речи. Отвечает за преобразование голосовых команд в текстовый формат с использованием библиотеки Vosk. Это позволяет системе интерпретировать голосовой ввод и преобразовывать его в управляемые команды для взаимодействия с пользователем.

Графический интерфейс пользователя (GUI). Реализован с использованием библиотеки Tkinter и предоставляет удобные инструменты для ввода, редактирования и просмотра данных. Пользователь может взаимодействовать с системой как голосом, так и через стандартные элементы интерфейса (кнопки, текстовые поля и выпадающие списки) (рисунок 1).



Patient Data						
Файл Ручной ввод Редактировать						
ID	Имя	Фамилия	Пол	Возраст	Рост	Вес
0	Степан	Иванов	мужской	40.0	180.0	90.0
1	Алина	Коваленко	женский	20.0	178.0	78.0
2	Олег	Никитенко	мужской	20.0	186.0	95.0
3	Елена	Козюличева	женский	21.0	175.0	60.0
4	Кирилл	Поленок	мужской	22.0	180.0	78.0
5	Евгений	Шкарубо	мужской	9.0	140.0	27.0
6	Виктория	Дрозд	женский	25.0	177.0	65.0

Рисунок 1 – Окно приложения

Модуль обработки данных. Основан на библиотеке Pandas, что обеспечивает структурированное хранение информации и её анализ. Данный модуль позволяет агрегировать и фильтровать данные, а также строить отчёты по введённым записям.

Функционал сохранения и загрузки данных. Реализован с возможностью экспорта и импорта данных в формате CSV. Это даёт возможность пользователям работать с сохранёнными записями и анализировать информацию в сторонних системах, таких как Excel или базы данных.

Алгоритм работы приложения. Программа функционирует по следующему алгоритму.

Инициализация модуля распознавания речи и графического интерфейса. На данном этапе загружается библиотека Vosk, инициализируется модель распознавания речи, а также создаётся графический интерфейс, который отображает основные элементы управления и список записанных данных.

Ожидание голосового ввода или ручного ввода данных через GUI. Программа находится в режиме ожидания команд от пользователя. Пользователь может либо ввести данные вручную через интерфейс, либо использовать голосовые команды.

Обработка голосовых команд, идентификация ключевых слов и передача их в систему управления данными. Распознанный текст анализируется, выделяются ключевые команды (например, «записать данные», «обновить запись», «сохранить файл») и передаются соответствующим модулям системы.

В зависимости от распознанной команды, программа выполняет одно из следующих действий. Записывает данные о пользователе (имя, возраст, пол и другие параметры). Обновляет уже имеющуюся информацию. Выводит данные на экран или формирует отчёт. Сохраняет данные в файл.

Запись данных в файл и возможность их последующего анализа. Данные сохраняются в удобном формате, позволяя использовать их в других программах. Также реализована возможность загрузки ранее сохранённых файлов.

Создание облачной базы данных. Для хранения и обработки данных в удалённом режиме была выбрана облачная база данных Firebase. Связь приложения с Firebase осуществляется через библиотеку Firebase Admin SDK. Этот процесс включает несколько этапов.

Настройка Firebase. В консоли Firebase создаётся новый проект, настраиваются параметры базы данных, такие как URL и правила безопасности. Генерируется специальный JSON-файл с ключами доступа для аутентификации приложения.

Инициализация соединения с Firebase. В коде программы загружается библиотека Firebase Admin SDK, аутентификация производится с использованием предоставленного ключа. Это позволяет приложению взаимодействовать с базой данных и отправлять запросы.

Получение данных. Запросы к Firebase выполняются с использованием API-методов, что позволяет получать данные в режиме реального времени. Приложение может отображать данные, загружая их непосредственно из облака.

Отправка данных. Программа может отправлять структурированные JSON-объекты в базу данных с помощью методов set(), push() или update(). Это позволяет добавлять новые записи, изменять существующие данные и организовывать хранение информации в удобном формате.

Обработка ошибок. Код включает механизмы обработки ошибок при потере соединения или недоступности сервера. В случае сбоя программа повторяет попытку подключения или информирует пользователя о проблеме.

Этот процесс позволяет приложению Windows интегрироваться с Firebase для хранения, обновления и извлечения данных в реальном времени, что даёт возможность удалённого доступа к информации и её безопасного хранения.

После правильной настройки базы данных, новые записи автоматически передаются в базу данных (рисунок 2).

Реализация и тестирование. В ходе разработки были протестированы основные функции системы, включая корректность распознавания голосовых команд, точность обработки данных и удобство взаимодействия с графическим интерфейсом.

Для проверки точности работы модуля распознавания речи были проведены тесты с различными голосовыми командами, учитывая различия в произношении. Были учтены возможные ошибки и разработаны механизмы исправления некорректных распознаваний (например, добавление уточняющих команд для корректировки введённой информации).

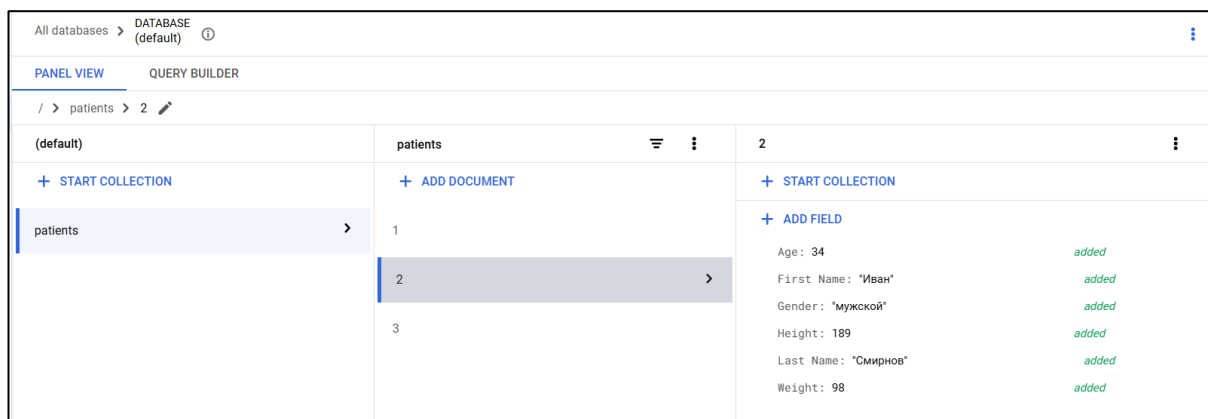


Рисунок 2 – Добавление новой записи в базу данных

Выводы. Разработанное приложение предоставляет удобный инструмент для работы с данными, снижая нагрузку на пользователей за счёт голосового управления. Это решение может найти применение в медицинской сфере, административных задачах и других областях, где требуется автоматизация ввода и обработки информации. В дальнейшем возможно расширение функционала системы, добавление поддержки дополнительных языков и улучшение алгоритмов распознавания речи.

Кроме того, система может быть интегрирована с облачными сервисами для хранения данных и работы в многопользовательском режиме. Улучшение алгоритмов машинного обучения позволит повысить точность распознавания команд, а также внедрить поддержку голосового управления на основе контекстного анализа пользовательского ввода.

Литература

1. Шкарубо, В. А. Получение, запись и обработки данных с использованием речевого интерфейса / В. А. Шкарубо // XXVII Республиканская научная конференция студентов и аспирантов «Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях», 2024. – г. Гомель : ГГУ им. Ф. Скорины. – С. 76–77.