

Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина» (УрФУ)

Институт радиоэлектроники и информационных технологий - РТФ

ОТЧЕТ

о проектной работе

по теме: Распознавание объектов

по дисциплине: Проектный практикум 1A

Команда: MASTERMINDS

Екатеринбург

2022

СОДЕРЖАНИЕ

[Введение 3](#_Toc106146277)

[Команда 5](#_Toc106146278)

[Целевая аудитория 6](#_Toc106146279)

[Календарный план проекта 9](#_Toc106146280)

[Определение проблемы 11](#_Toc106146281)

[Подходы к решению проблемы 12](#_Toc106146282)

[Анализ аналогов 13](#_Toc106146283)

[Требования к продукту и MVP 14](#_Toc106146284)

[Стек для разработки 15](#_Toc106146285)

[Прототипирование 16](#_Toc106146286)

[Разработка системы 17](#_Toc106146287)

[Заключение 18](#_Toc106146288)

[Список литературы 22](#_Toc106146289)

# ВВЕДЕНИЕ

         Каждый из нас получал хоть раз хоть что-то на почте - письма, посылки, открытки. С каждым годом этих почтовых отправлений все больше и больше. Как показывает статистика опросов, интернет-магазины в России хотя бы раз в месяц посещают 70% населения страны старше 12 лет, а ежедневно это делают 33% россиян. Соответственно, возникает множество новых почтовых отправлений, из-за сделок, совершаемых в Интернете. А если возрастает поток посылок, то возрастает и затраченный труд на обработку данных с них, их сортировку и пересылку до получателей. На наш взгляд, самое сложное в этом процессе как раз-таки и есть обработка данных с конвертов и посылок, ведь вручную обработать огромное количество данных очень сложно и времязатратно.

Работники почты при распознавании данных с писем проделывают огромный, и самое главное, рутинный труд. Ведь перебрать десятки, а иногда и сотни писем – работа не из легких. Люди устают, начинают ошибаться, снижается эффективность.

      Но мы живем в 21 веке – в эпоху цифровизации и автоматизации. Сегодня люди стараются сложную и рутинную работу переложить на роботов, искусственный интеллект, компьютер, чтоб облегчить себе жизнь.  Проанализировав ситуацию, мы с командой пришли к выводу, что если применить искусственный интеллект для распознавания данных с конвертов, то это намного упростит рабочий процесс работникам логистических организаций. Не нужно будет вручную проделывать по многу раз одну и ту же работу.

Искусственный интеллект экономит время и трудозатраты, а также он работает быстрее и точнее людей, а соответственно обрабатывает во много раз больше писем, чем человек за это же время. Это позволит повысить прибыль компании, снизит затраты на обработку логистических данных, и помимо всего этого, снизится стоимость выполняемой работы за счет автоматизации.

      К тому же, как мы и сказали выше, искусственный интеллект работает точнее людей – ведь при его использовании снижается количество ошибок, связанных с человеческим фактором.

 Актуальность этой работы заключается в том, что сейчас, как было сказано выше, век цифровизации и автоматизации. Каждые большие организации, особенно основанные на логистических операциях, будут рассматривать не только количество сил, стоимость выполнения, но и время затраченное на реализацию, в нашем случае, на распознавание и ввод информации с почтового отправления в базу. А чем меньше эти показатели, тем более прибыльным будет предприятие. С помощью нашего продукта работникам больше не придется тратить время на рутинную работу, ведь за них будет делать её наш искусственный интеллект.

Целью работы является разработка телеграм-бота, позволяющего распознавать данные с почтовых отправлений, такие как: ФИО отправителя и/или получателя, адрес, куда адресована посылка. Телеграм-бот получает на вход картинку, а на выходе получаем данные в текстовом формате, для дальнейших операций с ними.

Исходя из поставленной цели, в работе определены следующие задачи:

* Изучить целевую аудиторию.
* Выявить проблемы целевой аудитории и как их решает наш продукт.
* Проанализировать и определить плюсы и минусы решений конкурентов.
* Найти и переобучить нейросеть, которая будет способна распознать рукописный и печатный русский текст.
* Разработать телеграм-бота позволяющего распознавать данные с почтовых отправлений.

# КОМАНДА

Тимлид - Керимов Камран Имранович РИ-110912

Программист - Романов Вадим Юрьевич РИ-110950

Дизайнер - Числов Степан Игоревич РИ-110915

Аналитик - Илькаева Карина Булатовна РИ-110912

Тестировщик Юдицкий Семен Денисович РИ-110912

# ЦЕЛЕВАЯ АУДИТОРИЯ

Для определения целевой аудитории используется методика 5W Марка Шеррингтона. Это наиболее распространенный способ определения целевой аудитории и психологических характеристик, которыми обладают потенциальные потребители.

Сегментация рынка проводится по 5 вопросам:

Что?

Это сегментация по типу товара: что вы предлагаете потребительской группе? Какие товары/услуги?

Мы предлагаем телеграм-бота, позволяющего распознавать данные с конвертов с помощью искусственного интеллекта.

Кто?

Это сегментация по типу потребителя: кто приобретает товар/услугу? Какой пол/возраст?

Нашими основными клиентами будут работники логистических организаций, в частности почтовых отделений.Пол нашей целевой аудитории, очевидно, любой, потому что в логистических организациях работают как женщины, так и мужчины.

Возраст нашей аудитории – возраст официального трудоустройства в России. Проведя глубокое исследование, мы поняли, что в большинстве логистических организациях, а именно почтовых отделениях, возраст работников варьируется с 30 до 65 лет.



Почему?

Это сегментация по типу мотивации к совершению покупки и потребления: какова потребность или мотивация клиента? Какую проблему решает товар/услуга?

Наш продукт ускоряет процесс обработки данных с конвертов, также уменьшается вероятность ошибки, в том числе связанной с человеческим фактором.

Также снижаются затраты на обработку данных и снижается стоимость выполняемой работы за счет автоматизации

Когда?

Это сегментация по ситуации в которой приобретается продукт: когда потребители хотят приобрести товар/услугу?

Нашим продуктом потребитель будет пользоваться непосредственно во время работы. Наш продукт позволит упросить и ускорить работу потребителю и тем самым сэкономит время и сохранит ментальное здоровье.

Где?

Это сегментация по месту покупок: в каком месте происходит принятие решения о покупке и сама покупка? — имеются в виду точки контакта с клиентом, где можно повлиять на решение.

Так как софт разрабатывается для логистических организаций, то взаимодействие с клиентом происходит за его рабочим местом. Наш продукт будет рассчитан для использования за компьютером. Но следует обратить внимание, что для MVP мы реализуем продукт в виде телеграм-бота, чтоб можно было продемонстрировать работу нашего продукта.

# КАЛЕНДАРНЫЙ ПЛАН ПРОЕКТА

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | Название | Ответственный | Длительность | Дата начала | Временные рамки проекта |
| 1 нед | 2 нед | 3 нед | 4 нед | 5 нед | 6 нед | 7 нед | 8 нед |
| Анализ |
| 1.1 | Определение проблемы и способы ее решения | Числов Степан | 1 неделя | 09.03,2022 |  |  |  |  |  |  |  |  |
| 1.2 | Выявление целевой аудитории | Романов Вадим | 2 недели | 09.03,2022 |  |  |  |  |  |  |  |  |
| 1.3 | Анализ аналогов | Керимов Камран | 1 неделя | 09.03,2022 |  |  |  |  |  |  |  |  |
| 1.4 | Определение платформы и стека для продукта | Юдицкий Семен | 1 неделя | 16.03,2022 |  |  |  |  |  |  |  |  |
| 1.5 | Формулирование требований к MVP продукта | Числов Степан | 1 неделя | 16.03,2022 |  |  |  |  |  |  |  |  |
| 1.6 | Определение платформы и стека для MVP | Романов Вадим | 1 неделя | 16.03,2022 |  |  |  |  |  |  |  |  |
| 1.7 | Создание календарного плана | Илькаева Карина | 1 неделя | 16.03,2022 |  |  |  |  |  |  |  |  |
| Проектирование |
| 2.1 | Поиск готовой модели нейросети | Керимов Камран | 2 недели | 23.03.2022 |  |  |  |  |  |  |  |  |
| 2.2 | Проектирование оптимального алгоритма | Романов Вадим | 1 неделя | 23.03.2022 |  |  |  |  |  |  |  |  |
| Разработка |
| 3.1 | Написание кода | Романов Вадим | 3 недели | 30.03.2022 |  |  |  |  |  |  |  |  |
|  | Разработка Телеграм Бота | Юдицкий Семен | 1 неделя | 13.04.2022 |  |  |  |  |  |  |  |  |
| 3.2 | Тестирование приложения | Илькаева Карина | 1 неделя | 27.04.2022 |  |  |  |  |  |  |  |  |
| Внедрение |
| 4.1 | Оформление MVP | Юдицкий Семен | 1 неделя | 27.04.2022 |  |  |  |  |  |  |  |  |
| 4.2 | Внедрение MVP | Романов Вадим | 2 недели | 11.05.2022 |  |  |  |  |  |  |  |  |
| 4.3 | Написание отчета | Числов Степан | 1 неделя | 25.05.2022 |  |  |  |  |  |  |  |  |
| 4.4 | Оформление презентации | Числов Степан | 1 неделя | 25.05.2022 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Защита проекта |  |  |  |  |  |  |  |  |  |  |  |

# ОПРЕДЕЛЕНИЕ ПРОБЛЕМЫ

Для определения проблем и болей потребителя, мы провели небольшое исследование. Мы пользовались несколькими способами выявления основных проблем клиента:

1. Мы проанализировали работу работников логистических организаций исходя из открытых источников (сайтов, форумов, соцсетей). Выявили конкретные боли потребителя, способы решения их.
2. Также мы проанализировали конкурентов. Выявили плюсы и минусы их решений.

Исходя из всего этого мы выявили главную проблему - выполнение рутинной работы работниками почтовых организаций при распознавании данных вручную с почтовых отправлений.

# ПОДХОДЫ К РЕШЕНИЮ ПРОБЛЕМЫ

Так как тема нашего проекта - распознавание объектов с помощью искусственного интеллекта, то мы, определив проблему, примерно наметили способ ее решения. На первый взгляд казалось, что никакой невероятно сложной работы не предстоит, но приступив к реализации задуманного убедились в обратном. Решением нашей проблемы станет автоматизация процесса распознавания данных, а именно нейросеть, которая будет распознавать данные с конвертов, и к тому же ручная проверка и согласование после этого не потребуются.

Также, мы нашли похожий подход у наших конкурентов, но об этом подробнее мы напишем в блоке с анализом аналогов.

# АНАЛИЗ АНАЛОГОВ

Проанализировав рынок, мы обнаружили несколько прямых и косвенных конкурентов, у которых аналогичное или похожее программное обеспечение на аналогичном рынке и работающее с такой же целевой аудиторией, что и наш разрабатываемый продукт. Но мы не совсем считаем их за конкурентов, ведь их решения используются достаточно долгое время, причем успешно, а наш продукт находится на стадии разработки. Другими словами, наше решение, по сравнению с их – всего лишь капля в море.

К аналогам мы отнесли решения таких компаний, как “Почта России” и “DHL”. На зарубежном рынке логистических компаний в разы больше, присутствуют даже полностью автоматические. Но поскольку мы делаем аспект на российском рынке, мы и будем рассматривать решения, представленные на нём. Проанализировав способ реализации решений аналогов глубже, мы выявили у них небольшие недостатк. Например, на “Почте России” данные с конверта вносит работник почты вручную, и позже клеит штрих-код, чтоб машина хоть как-то его определила. Наше же решение позволяет пропустить шаг с вводом данных вручную, ведь это будет делать наш продукт автоматически. Это экономит время и силы, которые могут быть затрачены на понимание написанного отправителем и ввод этих данных в базу.

Также у нашего продукта, в отличие от конкурентов, открытый исходный код. Любой желающий может скачать решение с репозитория на “GitHub”, и изменить программный код под свои нужды.

# ТРЕБОВАНИЯ К ПРОДУКТУ И MVP

1. Продукт и MVP должны уметь принимать фотографии на вход.
2. Продукт и MVP должны уметь обрабатывать данные с фотографии
3. Продукт и MVP должны определять как и печатный так и рукописный текст на фотографии.
4. Продукт и MVP должны выводить на выход данные с фотографии конверта в текстовом формате.
5. Продукт и MVP должны иметь простой и понятный интерфейс.

# СТЕК ДЛЯ РАЗРАБОТКИ

Для разработки продукта был использован следующий стек:

* Язык программирования – Python.
* Для распознавания машинописного текста – библиотеки PyTesseract, OpenCV.
* Для распознавания рукописного текста – измененный вариант модели Simple HTR, с использованными в нем библиотеками; YandexSpeller.
* Для создания телеграм-бота - библиотека aiogram

# ПРОТОТИПИРОВАНИЕ

1. Интерфейс телеграм бота



1. Список команд.
	* /start – запускает бота и выводит приветственное сообщение с возможностями бота
	* /help – список возможностей бота
	* /demo – демонстрация работы нейросети (сделано для предзащиты и защиты проекта)
	* /info – о проекте

# РАЗРАБОТКА СИСТЕМЫ

Алгоритм работы продукта

1. Пользователь запускает бота, ему приходит приветственное сообщение со списком возможностей бота.
2. Пользователь отправляет боту на вход фотографию с конвертом
3. Телеграм бот, обрабатывает полученную картинку и отправляет текстовые данные с нее в ответном сообщении.

# ЗАКЛЮЧЕНИЕ

В ходе данной проектной работы была определена одна из проблем логистических организаций, а именно проблема ручного распознавания данных с почтовых отправлений. Из-за того, что это занятие рутинно, очень времязатратно и вызывает много ошибок, мы решили заменить это решение автоматизированным процессом, с помощью нейросети, которая будет распознавать данные с конвертов.

 Нашей целью было создание продукта, который позволит решить эту проблему с помощью нейросети.

 Проанализировав рынок, мы определили наших конкурентов и постарались учесть их недостатки. Наше решение, в отличие от других,  будет полностью независим от человеческого фактора, а именно, от человека не будет требоваться ручная перепроверка и согласование.

 Целевой аудиторией, в основном, будут являться работники логистических центров.

Чтоб достичь поставленной цели, нам пришлось найти оптимальные технические решения. Перебрав несколько вариантов, мы остановились на следующих:

* Для реализации задуманного, мы использовали язык программирования Python и его библиотеки.
* Для распознавания машинописного текста, были задействованы библиотеки, такие как  PyTesseract, OpenCV.
* В основу нашей модели распознавания рукописного текста была взята архитектура Simple HTR, разработанная австрийским учёным Гаральдом Шайдлом. В его модели использовался СТС Loss, в качестве функции потерь. Он ограничен максимальным размером входящей в него последовательности символов – 32 элемента, поэтому наши примеры содержали не более 32 символов. Если развивать проект полномасштабно, то эти ограничения пришлось бы переписать. Но поскольку мы создаем проект для учебного заведения, остановились на готовом варианте.
 Основными отличиями нашего варианта Simple HTR от оригинального стали: изменённый загрузчик данных, для более удобной работы с ними, а также отличием стало изменение в блоке вывода распознанного текста, где мы добавили вариант декодера с коррекцией правописания, основанной на работе "Спеллера" (Speller) разработанного Yandex.
 Яндекс Спеллер помогает находить и исправлять орфографические ошибки. Языковые модели Спеллера включают сотни миллионов слов и словосочетаний.

Модель обучалась на домашнем стационарном компьютере нашего программиста, c датасетом размером чуть более 38 тысяч картинок. На её обучение уходило порядка 10-12 часов.

Мы применяли алгоритмы Best Path и Beam Search. При использовании алгоритма Beam Search результаты стали существенно лучше. Ниже приведены графики изменения наших метрик (точности предсказания по символам и точности предсказания по словосочетаниям, которые мы подавали на обучение) в зависимости от эпохи обучения:





* Для реализации телеграм-бота для MVP, мы использовали библиотеку aiogram для Python.

Также, для написания презентации к защите мы использовали Figma – онлайн-сервис для разработки интерфейсов и прототипирования.

В итоге представлен эффективный продукт, позволяющий потребителю:

* С помощью фотографии почтового конверта, получить оцифрованные данные об отправителе и получателе, независимо от того была информация на конверте в рукописном или печатном виде.
* Оптимизировать свой рабочий процесс
* снизить затраты на обработку логистических данных
* снизить стоимость выполняемой работы за счет автоматизации

Целью этого проекта являлось сделать продукт, который сможет автоматизировать процесс распознавания данных с почтовых отправлений.

Исходя из поставленной цели были поставлены следующие задачи:

* изучить целевую аудиторию и выявить ее проблемы, и как их решает наш продукт
* проанализировать подходы решений этой проблемы у конкурентов
* найти и переобучить нейросеть, которая будет способна распознать рукописный и печатный русский текст.
* реализовать телеграм-бота позволяющего распознавать данные с почтовых отправлений.

Таким образом, все цели и задачи, поставленные вначале написания данной проектной работы, были полностью выполнены и освещены

# СПИСОК ЛИТЕРАТУРЫ

1. Исследование возраста работников почты:

<https://zen.yandex.ru/media/id/5ca6ead92b3ad600b22b8086/skolko-chelovek-rabotaet-na-pochte-rossii-my-poschitali-vseh-60ccdf92bb96047128fbaf51>

1. Распознавание “Почты России”

<https://www.idexpert.ru/news/pochta-rossii-ispolzuet-tekhnologiyu-opticheskogo-raspoznavaniya-simvolov/>

1. Распознавание “DHL”

<https://www.dhl.com/global-en/home/insights-and-innovation/insights/artificial-intelligence.html>

1. Исследование о росте интернет-заказов

<https://yandex.ru/company/researches/2019/market-gfk>

5. SimpleHTR

 <https://github.com/githubharald/SimpleHTR>