

# КОПТЕВ ДМИТРИЙ

АНАЛИТИК

## НАВЫКИ

- PYTHON, MATPLOTLIB, PANDAS, PLOTLY, NUMPY, SCIPY, KERAS, OPENCV, KORNIA
- SQL, POSTGRESQL, SQLITE
- MS EXCEL, МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, ТЕОРИЯ ВЕРОЯТНОСТЕЙ
- LINUX UBUNTU, КОМАНДНАЯ СТРОКА
- АНГЛИЙСКИЙ (C1)

Обучение нейронных сетей различной сложности, предобработка данных, статистический анализ выборки, визуализация данных средствами Python, построение баз данных.

## ОБРАЗОВАНИЕ

СПбГУАП бакалавриат (3 курс)

Сентябрь 2021 - Июнь 2025

Факультет информационных технологий и программирования

## ПРОЕКТЫ

Github URL: <https://github.com/very-cool-dude>

Прогнозирование временного ряда (time\_series\_forecasting)

Свёрточная нейронная сеть распознавания сгенерированных изображений (ai\_generated\_image\_classifier)

Свёрточная нейронная сеть (convolutional\_neural\_network)

Распознавание образов (image\_recognition)

Полносвязная нейронная сеть (dense\_neural\_network)

## ДОПКУРСЫ

Базовый курс по анализу данных. Тинькофф образование

Интерактивный тренажер по SQL. Stepik

Intro to Statistics. Udacity

## ДОСТИЖЕНИЯ

- Лидер и участник стартап проектов программы "Акселератор Техно Питер"
- Участник научных конференций ГУАП, МАИ, ФАОУ
- Участник хакатонов

## КРАТКАЯ ИНФОРМАЦИЯ

Здравствуйте! Занимаюсь анализом данных и машинным обучением

## КОНТАКТНАЯ ИНФОРМАЦИЯ

Телефон:

+7 (937) 980 26 87

e-mail:

dmitrii.koptev76@gmail.com

Телеграм:

@mitrey

# АКСЕЛЕРАЦИОННЫЕ ПРОЕКТЫ

## • CozyNet

Проект, направленный на автоматизацию сбора показаний счётчиков в многоквартирных домах для ведения отчётности, статистики и оптимизации трат.

Предполагалось устанавливать наши счётчики в квартиру, а также один общий счётчик на дом. Прототип умного счётчика должен был быть изготовлен на основе Arduino. Благодаря данному цифровому решению показания с уже имеющегося “простого” счётчика жкх в квартире попадали бы на наш, после чего происходила бы запись расходов и их учёт. В дальнейшем из всей информации формировались бы метрики и графики расходов, которые отображались бы в удобном онлайн дашборде. В данном проекте мы провели каст-дев нескольких отделений жкх по Санкт-Петербургу. Мы закрыли все необходимые компетенции в команде и начали разработку прототипа. В проекте я выполнял менеджерские и организаторские обязанности. Был проведён анализ рынка, формирование бизнес модели, проверка гипотез и проведение каст-девов.

## • Автоматизированная система учёта товаров в упаковке ‘Каунтер’

Система, которая позволит быстро настроить учет для конкретной упаковки и будет автоматически подсчитывать количество мелких предметов в крупном контейнере.

Реализация предполагала использование нейронных сетей “YOLO”, так как они отлично подходят для задачи детекции. Снизить процент ошибки предполагалось за счёт грамотной установки света и большого количества снимков.

В данном проекте занимал лидерскую роль, занимался маркетингом и занимался организацией взаимодействий внутри команды и вне её.

Нашим достижением является то, что мы смогли договориться с действующим производством (фармацевтическое производство “Эллара”) для предоставления материала, чтобы мы могли обучить нашу модель.

Проект был отменён, в связи с наличием большого количества подобных решений на рынке и его небольшой размер.

После отмены данного проекта мне и команде пришлось в кратчайшие сроки (неделя) менять проект и начинать с нуля. Мы справились с данной задачей успешно и сформулировали новую идею, бизнес модель и продуктовые гипотезы

## • ИИ-ассистент врача

ИИ-ассистент врача с распознаванием речи врачей при заполнении медицинских протоколов и других электронных медицинских документов, системой голосовой самозаписи пациентов к врачам.

Реализация находится на этапе активного обсуждения. Стоит проблема обучения модели на большом количестве медицинских терминов и достижение хорошего считывания ключевых слов с полученной аудио записи приёма

В проекте выполнял лидерские обязанности, проводил каст-дев, организовывал обсуждения, занимался анализом рынка.

Нашим достижением является то, что мы смогли пройти с нашим проектом во второй этап акселерационной программы.

Проект в данный момент находится на стадии поиска людей для каст-дева, чтобы подтвердить имеющиеся гипотезы

# НАУЧНАЯ ДЕЯТЕЛЬНОСТЬ

## • Анализ рисков контента, созданного искусственным интеллектом

Рассматриваются риски и проблемы распространения нейросетей, способных создавать комплексный контент. При анализе рисков были выделены несколько главных проблем, которые впоследствии были подробно разобраны.

## • Анализ достоверности результатов работы поисковых систем

В данной работе рассматривается создание и использование искусственного интеллекта (ИИ) для генерации контента, а также анализируется достоверность результатов работы поисковых систем в контексте распространения такого контента. Исследованы общие проблемы и риски, связанные с распространением ИИ-контента. Особое внимание уделено проблеме киберпреступности и распространению недостоверного контента. Проанализирована работа поисковых систем, обобщен процесс поиска и ранжирования информации в интернете, а также рассмотрены алгоритмы, которые определяют порядок выдачи документов. Были разобраны наиболее популярные примеры и выделены перспективы дальнейшего развития ИИ.

## • Распознавание изображений, сгенерированных нейросетями

В данной статье рассматриваются способы решения проблемы распознавания сгенерированных изображений. Выделяется наиболее благоприятная реализация и проводится работа по созданию нейронной сети, способной отличить искусственное изображение от настоящего. Рассматриваются тонкости создания обучающей выборки, настройки и оценки полученной модели.

Для решения задачи распознавания изображений был сформирован обучающий и валидационный датасет, содержащий порядка 10000 изображений, разделённых на два класса: настоящие и сгенерированные образы. При формировании набора “настоящих” изображений были использованы материалы из сети интернет раньше 2019 года, чтобы исключить возможность попадания ложного образа. Выборка искусственных изображений была отобрана среди специализированных сообществ, посвящённых нейросетям, например в социальной сети reddit. При дальнейшей предобработке были убраны все файлы, имеющие не относящиеся к изображениям расширения. Так как искомая задача заключается в классификации изображений, то была создана свёрточная нейронная сеть, состоящая из 9 слоёв. Полученная модель показала крайне высокие результаты предсказания на случайных изображениях (погрешность в пределах 5%). Для оценки модели были посчитаны следующие метрики: Precision (0.82), Recall (1), BinaryAccuracy (0.92).

Высокие значения подтверждают корректность работы алгоритма. Реализация алгоритма распознавания показала, что выбранное решение жизнеспособно и требует лишь внедрения в систему браузера.

# ХАКАТОН

Участвовал с командой из 2 человек в хакатоне, организованном Baggins Coffee. Нами был выбран кейс “Создание системы аналитики данных и прогнозирования показателей деятельности кофейни с использованием искусственного интеллекта”.

Датасет для решения задачи содержал в себе различную информацию о ежедневных продажах кофеен (выручка, количество проданных горячих/холодных напитков, температура на улице, погода, адрес).

Для решения данной задачи была выбрана модель, разработанная Facebook, под названием “Prophet”. Эта модель крайне хороша для решения задач прогнозирования временных рядов, с которой мы и имели дело. Также, благодаря выбранному решению, можно без труда определить тренд и сезонную составляющую данных, а также составить прогноз на их основе.