



Марченко Андрей
Программист
21 год

+7-967-123-25-25
marcheanin1576@gmail.com
GitHub Profile

ОБРАЗОВАНИЕ






- МГТУ им. Н. Э. Баумана** 2021-2025
ИУ9 Теоретическая информатика и компьютерные технологии (ПМИ) Москва
 - ООП (Java, c++)
 - Компьютерные сети
 - Алгоритмы и структуры данных
 - Базы данных (T-SQL)
 - Дискретная математика
 - Теория формальных языков
 - Многопоточное и распределенное программирование
 - Конструирование компиляторов
 - Мобильная разработка (Android)
 - Численные методы
- ВКР: Разработка алгоритма автоматического формирования модели кредитного скоринга**
- VK Education – Отделение при МГТУ им. Н.Э. Баумана (бывш. Технопарк)** 2023-2024
Машинное обучение, диплом №460-89704
 - Программирование на современном C++
 - Углубленный Python
 - Алгоритмы и структуры данных
 - Классическое ML
 - Нейронные сети в ML
 - Навыки деловых коммуникаций
 - DWH

ОПЫТ

- Компания ИНВЕРСИЯ** Июль 2022 – Февраль 2023
Программист, Java, PL/SQL Москва
 - Работал в связке Java – SQL, решал задачи как во фронтенд, так и в бэкенд частях
- Газпромбанк** Июнь 2024 – наст. время
Инженер разработки, Java, Москва
 - Микросервисная разработка на языке Java.
 - Разработка сложных структур баз данных PostgreSQL.

ПРОЕКТЫ

- Анализ мемов на hateful содержание** Pet-проект, Декабрь 2023 [Source Code](#)
 - * Курсовой командный проект в рамках курса Нейронный сетей в VK Education
 - * Сервис анализирующий входной мем(картинку) на ненавистническое содержание
 - * Занимался объединением свёрточной и модели анализа текста с помощью классификационной головы
 - * **Используемые технологии:** Python, Pandas, Lightning, PyTorch
- AutoML для кредитного скоринга** Дипломная работа, Весна 2025 [Source Code](#)
 - * ВКР на степень бакалавра, МГТУ им Баумана, ИУ9, ПМИ
 - * Система автоматического формирования полного пайплайна для задачи бинарной/многоклассовой классификации
 - * Основа оптимизации - генетический алгоритм, оптимизирующий 22-генную хромосому, кодирующую пайплайн

- * В работе использовано 5 этапов предобработки данных, каждый из которых имеет вариативность методов и гиперпараметров
 - * Использование одной из 4 моделей машинного обучения, получение полных метрик
 - * Оформление продуктового решения - сериализация пайплайна для дальнейшего использования, удобный интерфейс CLI
 - * **Используемые технологии:** Python, Pandas, Scikit-learn, Numpy, Imbalanced-learn, Gensim, Tensorflow
- **Многокритериальная оптимизация в задаче кредитного скоринга** *Курсовой проект, Декабрь 2024* **k** [Source Code](#)
- * Курсовой проект по численным методам с применением нейронных сетей
 - * Построение многокритериальной модели оптимизации разноплановых loss-функций
 - * Использование генетических алгоритмов, обработка гетерогенных данных, разработка собственного многокритериального алгоритма оптимизации, сравнение подходов
 - * **Используемые технологии:** Python, Pandas, Tensorflow, PyTorch
- **Проекты в рамках курса Теории формальных языков** *Учебные проекты, Октябрь 2023 - 2024*  [Source Code](#)
- * Разработка анализатора завершаемости произвольного количества SRS-правил переписывания
 - * Разработка преобразователя регулярных выражений с инструкциями Lookahead и Lookbehind в академические регулярные выражения при помощи операций с НКА с модулем fuzz-тестирования
 - * Приближение контекстно-свободных (и их подклассов) языков, заданных оракулами, с помощью алгоритма Angluin's NL* для построения RFSA-автоматов и построения двухместных накачек по языкам префиксов и суффиксов
 - * **Используемые технологии:** C++, CMake, Git, Algo
- **Курсовая работа по компьютерной графике** *Pet-проект, Март 2023 - Июнь 2023*  [Source Code](#)
- * Разработка приложения для Linux, способного при получении данных о мощностях сигнала в режиме реального времени строить траекторию его передвижения
 - * Спроектировано как десктопное приложение для анализа кода на уникальность сравнением с базой уже существующих исходников
 - * Используются алгоритмы расчёта отрисовки поля и системы координат для корректного и читаемого отображения нескольких объектов
 - * В основе расчета координат и погрешности лежит алгоритм трилатерации
 - * **Используемые технологии:** Python, Linux, PyOpenGL, glfw, Git, algo
- **SourcedOut** *Pet-проект, Март 2023 - Июнь 2023*  [Source Code](#)
- * Курсовой командный проект в рамках курса Программирования на современном c++ VK Education
 - * Спроектировано как десктопное приложение для анализа кода на уникальность сравнением с базой уже существующих исходников
 - * При этом может быть использовано как дополнение к сервису проверки лабораторных работ LabSpace
 - * В проекте отвечал за алгоритмы проверки на плагиат - text-based, token-based. Реализовал удобные интерфейсы и библиотеки для вызова моих метрик.
 - * Использовал библиотеку ANTLR для токенизации популярных языков, написал собственные реализации алгоритмов расчёта метрик - Расстояние Левенштейна, Метрика Жаккарда, а также W-Shingling
 - * Тесты написаны с использованием Google test
 - * **Используемые технологии:** C++, CMake, libpqxx, Boost, Google test, Git
- **LabSpace** *Pet-проект, Июнь 2022 – 2024*  [Source Code](#)
- * WEB-сервис для создания, проверки и оценки лабораторных работ по различным дисциплинам
 - * Участвовал в разработке всех частей приложения, производил деплой на сервер
 - * Участвовал проектировании базы данных, написания моделей для взаимодействия с ней.
 - * Разрабатывал почти все формы шаблонизаторов взаимодействия бекэнд с фронтэнд частью, занимался деплоем на сервер.
 - * Генерировал идеи, и воплощал из в жизнь. Продолжаю поддерживать проект для его работы на кафедре.
 - * **Используемые технологии:** Java, Maven, Spring boot, MySQL, Thymeleaf, Git, HTML, CSS
- **ClusterText** *Pet-проект, Весна 2024*  [Source Code](#)
- * Выпускной проект курса мл-разработчик в VK Education x МГТУ им. Н.Э.Баумана
 - * Web-приложение, призванное кластеризовать текст, выделять ключевые слова и расставлять тэги для лучшего понимания новости и отнесения её в соответствующий раздел.
 - * Моей частью была модель для кластеризации текстов - модель умеет относить текст к одному или нескольким заранее выделенным кластерам.
 - * Состоит из двух частей - эмбеддера на основе модели SBert и kmeans для обучения без учителя полученных эмбеддингов.
 - * Модель обучена на 60к статей, проведен анализ данных для лучшей различимости кластеров.

- * Модель различает 15 кластеров (от семьи до экономики и политики) и отлично справляется с маркировкой текстов одним или несколькими из них
- * **Используемые технологии:** Python, Pandas, Clustergram, kmeans, Transformers, pytorch

ТЕХНИЧЕСКИЕ НАВЫКИ

Языки программирования: C/C++, Python, Java, Go, Scheme, MySQL, T-SQL, PL/SQL, PostgreSQL, Dart
Среды и инструменты: Git/GitHub, VS Code, IntelliJ CLion/PyCharm/IDEA, Android Studio
Библиотеки и фреймворки: Pandas, Numpy, Matplotlib, Spring, boost, PyOpenGL, PyTorch
Языки: Русский, Английский(C1), Немецкий(A2)

Навыки: алгоритмы и структуры данных; компьютерные сети; машинное обучение; глубокое обучение; анализ естественных языков; ООП; дискретная математика; многопоточное программирование; компьютерная графика; работа с базами данных.
Дополнительно: составление документации; понимание банковского бизнеса; умение работать в команде; деловая коммуникация; автоматизация процессов

СИЛЬНЫЕ СТОРОНЫ И ИНТЕРЕСЫ

Сильные стороны: усердие, исполнительность, ответственность, трудолюбие, способность к обучению и развитию, чувство юмора
Увлечения: настольный теннис, шахматы (I взрослый разряд), футбол, робототехника.