

Цепков Ярослав Алексеевич

Data Scientist

Email: yaroslav.tsepkov@gmail.com | Phone: +79608243141

Github: [yaroslavtsepkov](#) | Telegram: [zireaelv](#)

Образование

Самарский университет

Магистр

Август 2020 - Август 2022

Факультет информатики. Прикладная математика и информатика.

Разработка интеллектуальных систем.

Самарский университет

Бакалавриат

Август 2016 - Август 2020

Факультет математики. Прикладная математика и информатика.

Системный анализ.

Опыт работы

ООО Ксимад (Xim inc.)

Data Engineer

Январь 2022 - по настоящее время

- Поддержка текущей ETL системы и помощь в проектировании новой
- Разработка модулей по сбору данных различных нетворков используя API сервисов. Поддержка текущих модулей.
- Формирование витрин данных аналитикам для построения отчетов

СБЕР

Старший аудитор (Data Scientist)

Март 2021 - Август 2021

- Выгрузки данных из различных источников с использованием различных диалектов SQL
- Очистка и маскирование данных с использованием regex и расстояния Ливенштейна
- Проведение анализа с использованием графовой теории. Построение графов связей и процессов с использованием Gephi, networkx, SberPM
- Исследовательские и продуктовые проекты в области машинного обучения: задачи классификации, кластеризации и поиска аномалий в данных, распознавание документов, детекция объектов
- Написание внутренних библиотек для будущих разработок

Интегра-С

Программист

Январь 2020 - Август 2020

- Выявление бизнес требований к приложениям.
- Помощь в разработке дизайна. Преимущественно работа с векторной графикой.
- Написание тонкого клиента для мобильных приложений iOS/Android.
- Работа с REST API, написание документации к ПО.
- Работа с потоковым видеосигналом RTSP.
- Поиск оптимальной модели для детекции объектов на видео для серверной части.

Персональные проекты

Умный дом

Участие в проекте умный дом в качестве мобильного разработчика. Разработка мобильного клиента на iOS, разработка архитектуры базы данных.

Высокопроизводительные вычисления

Решение задач линейной алгебры, обработки сигналов с использованием технологии CUDA. Написание статьи на newtechaudit: [Использование технологии CUDA в решении прикладных задач](#).

Обработка изображений

Изучение библиотек для работы с изображениями и собственная реализация алгоритмов обработки изображений. Написание двух статей по этой теме на newtechaudit: [Анализ и обработка изображений с использованием операций математической морфологии](#) и [Разработка аналитических веб-приложений с использованием библиотеки Streamlit](#)

Навыки

- Языки программирования **Python, Java (EE 8 stack), C (only high performance), Scala(only spark framework)**.
- Работа с данными **Pandas, Dask, OpenCV, Spark, SQL, Big Query**.
- Машинное обучение и глубокое обучение **SciKit-learn, PyTorch, XGboost, Keras**.
- Высокопроизводительные вычисления **CUDA, Numba, Multiprocessing, Threads**.
- Визуализация **Matplotlib, Seaborn, Plotly, Dash, Altair, NetworkX, Gephi**.
- Сетевой стек **aiohttp, socket, ssl, request, wireshark, selenium, oauth 2**.
- Остальное **Linux, Bash, Git, Jira, Markdown, Latex, WolframMatematica, GNU Octave, HTML5, CSS3**.