








Владислав Суматохин

● Инженер / Аналитик данных

Аналитик и инженер данных с опытом работы с классическим машинным обучением, разработки ETL процессов и управления базами данных.

КОНТАКТЫ

-  +7 993 890 91 02
-  @dotsum
-  me@dotsum.tech
-  @dotsum
-  <https://dotsum.github.io/>

ОБРАЗОВАНИЕ

Москва, Россия 2020 - 2022

Магистратура (закончил с отличием)
МГТУ им. Н.Э. Баумана

Направление обучения:
Организация и управление
научноёмким производством

Москва, Россия 2020 - 2022

Магистратура (закончил с отличием)
Московский политех

Направление обучения:
Машиностроение

ЗНАНИЕ ЯЗЫКОВ

- Русский
родной
- Украинский
родной
- Английский
C1
- Немецкий
A1

ОПЫТ РАБОТЫ

ML / Data Engineer

Москва, Россия
06.2022 - 05.2023



РТК-Элемент
Предыдущая компания
была поглощена
Ростелекомом

- Разрабатывал **ETL-процессы** для обучения и использования ML-моделей. Разработал **20+ процессов с нуля**.
- Разрабатывал и **оптимизировал загрузки** исходных данных из XML-файлов. Увеличил скорость загрузки **более чем в 2 раза**.
- Создал предиктивные модели для оценки риска смерти и прогрессирования болезни с использованием данных EHR.
- Улучшил точность моделей**, сократив ошибку второго рода до 20%, **обеспечив интерпретируемость** для медицинских специалистов.
- Принимал участие в **миграции с PostgreSQL на GreenPlum**, оптимизировал процедуры и структуры данных для GreenPlum.
- Администрировал **инфраструктуру ML**: JupyterHub, MLFlow, S3, GreenPlum для команды из четырех человек.

Аналитик данных

Москва, Россия
05.2021 - 05.2022



Лаборатория элемент
Государственный подрядчик
по цифровизации
здравоохранения

- Разрабатывал **ETL-процессы** для аналитического хранилища данных с использованием PostgreSQL и Apache Airflow. Разработал **20+ процессов с нуля**.
- Администрировал сервер **Tableau для 15+ пользователей**.
- Разработал и поддерживал **10+ управленческих дашбордов** в Tableau. Общался с конечными пользователями, **улучшая пользовательский опыт**, в результате чего большинство дашбордов использовались **ежедневно**.
- Подготавливал скрипты для регулярной и adhoc отчетности. Улучшил частоту обновления регулярной отчетности с месячной на ежедневную.
- Принимал участие в **миграции с Oracle DB на PostgreSQL**, рефакторил PL/SQL процедуры и ETL-процессы.

ТЕХНИЧЕСКИЕ НАВЫКИ

ЯП	Python, JavaScript/TypeScript
Инфраструктура	Linux, Docker, k8s, JupyterHub
Базы данных	PostgreSQL, GreenPlum, MongoDB, S3, Redis
Обработка данных	Pandas/Polars, Dask, Spark, Airflow
Машинное обучение	NumPy, SciPy, Sk-learn, CatBoost, PyTorch, TensorFlow, MLFlow, DVC
Визуализация	Tableau, matplotlib, plotly, bokeh, D3.js
Другие навыки	HTML, CSS, React, Redux, GraphQL, FastAPI, git, Jira, prompt engineering

