

Скоробогат Кирилл

Москва, м. Лефортово • skorobogat2017@yandex.ru • +7 (951) 712-64-64

О себе

Студент 4-го курса, мотивирован развиваться в области Data Engineering, особенно в построении ETL-процессов и работе с распределёнными системами. Имею практический опыт с Python, SQL, HDFS, Hive, Spark. Быстро осваиваю новые технологии и умею работать в команде.

Образование

Национальный исследовательский университет «МЭИ»

Высшее образование, Прикладная математика и информатика, GPA: 4.57

Москва, Россия

2022 – 2026

Школа 21 (Образовательный проект Сбера)

Peer-to-peer обучение, программирование и разработка ПО

Москва, Россия

2024 – 2025

Опыт

Techno Diasoft

Москва, Россия

Участник корпоративного обучения на позицию Data Engineer

Май 2025 – Август 2025

- Реализовал скрипт для SCD2-инкрементальной загрузки, обеспечив эффективное обновление данных в хранилище.
- Оптимизировал структуру БД до третьей нормальной формы, что позволило снизить избыточность и уменьшить объём хранения.
- Построил витрину данных для анализа эффективности промоакций, обеспечив удобный доступ к метрикам бизнеса.
- Реализовал парсинг логов веб-сервера в Hive для анализа активности пользователей.
- Получил практический опыт работы с HDFS, Hive, YARN, Spark.

Школа 21 (Образовательный проект Сбера)

Москва, Россия

Участник обучения

Сентябрь 2024 – Июль 2025

- Организовывал командные встречи и дополнительные лекции для участников обучения.
- Руководил групповыми проектами в рамках программы: распределял задачи, контролировал сроки, помогал решать технические вопросы.
- Развивал навыки командной работы, коммуникации и управления проектами в среде peer-to-peer обучения.

Ренессанс Банк

Москва, Россия

Участник хакатона «РентонХакатон 2.0»

- Работал в команде над решением задачи кредитного скоринга в ограниченные сроки.
- Участвовал в распределении задач и координации работы в команде.
- Получил опыт командной работы в условиях высокой динамики и неопределённости.

Учебные проекты

ETL-пайплайн для обработки погодных данных

- Реализовал асинхронный сбор погодных данных по городам России из API Open-Meteo (Python, asyncio) и выгрузил полученные данные в HDFS.
- С помощью PySpark подготовил датасет для обучения ML-моделей и агрегированный датасет для аналитики.
- Настроил автоматизированный пайплайн в Apache Airflow: загрузка данных, очистка, обработка и агрегация.

Навыки

Программирование и анализ данных:

- Python (Advanced), SQL (Advanced), Pandas (Intermediate)

Big Data и распределённые системы:

- HDFS (Intermediate), Hive (Intermediate), Spark (Intermediate), YARN (Basic)

Системные инструменты:

- Linux (Intermediate), Git (Intermediate), Docker (Basic)

Иностранные языки:

- Английский — уровень C1 (CAE — Certificate in Advanced English)