

# Липатов Данила Вячеславович

Data Analyst / Data Scientist (Middle)

 danchik-lip@mail.ru |  [Telegram](#) |  [LinkedIn](#) |  [GitHub](#)

## Обо мне

Аналитик данных с 3-летним опытом в области анализа данных, машинного обучения, архитектуры баз данных и разработки веб-сервисов. Ответственный и инициативный специалист, любитель математики и олимпиадных движений. Ищу позицию Middle в Data Analytics / Data Science.

## Навыки

Стэк: Python, TensorFlow, SQL, Docker, Streamlit, Selenium, Git, scikit-learn, VBA, Metabase, Django, Jinja2, FastAPI, Dask, QA (unit-tests)

ML: Computer Vision, классический ML (catboost / xgboost / randomforest / extratrees / CNN / VotingClassifier / stacking)

Аналитика: A/B тестирование, тестирование гипотез, аналитика в Metabase

## Опыт Аналитик — Отдел валидации, НКР / РБК Июль 2022 — по настоящее время

- **Разработка инфраструктуры валидации и аналитики кредитных рейтингов:**
- Участвовал в разработке основного сервиса для валидации рейтингов, внедрил расчёт метрик PSI, VIF, PD. Провёл A/B-тесты и проверку гипотез новых методов, что повысило точность валидации моделей на ~8 %; более 5 улучшений были приняты в методологию агентства.
- **ML / Применение и реализация:**
- Создал модуль обхода CAPTCHA на базе CNN (Keras/Tensorflow), метрика качества классификации достигла ~91%, полная автоматизация ручного труда при парсинге необходимых данных.
- Реализовал с нуля скоринговую модель классификации кредитных рейтингов (voting classifier/ catboost/ xgboost/extratrees(rf) с дополнительным анализом фичей). **F1 — 87%**.
- Реализовал сервис для анализа кредитных рейтингов горизонте 1-3 года. Относительная стабильность улучшилась в среднем до 75%.
- **Backend и внутренние инструменты:**
- Разработал микросервис генерации отчётов (FastAPI + Jinja2 + Streamlit), сокративший время подготовки документов ~70%.
- Создал защищённый интерфейс редактирования таблиц (Streamlit + SQLAlchemy), позволивший упростить работу с базой для тех, кто не знает SQL. Безопасность и контроль изменений на уровне базы.
- Реализовал 3 крупных парсера на базе Selenium + BS, которые напрямую коррелируют с моделями анализа кредитных рейтингов. В среднем по всем парсерам ускорение получения данных достигло ~65%
- **DevOps / MLops:**
- Все сервисы были упакованы в Docker, развёрнуты на stage-сервере, с настройкой nginx и подключением доменов. Настроил Metabase, включая подключение к аналитическим хранилищам и настройку ролей.
- Для скоринговых моделей собран единый ML пайплайн на базе mlflow, обеспечивающий непрерывный мониторинг обучения моделей, отбора качественных и успешных запусков. Данный инструмент снизил нагрузку на разработчиков в постоянном изменении моделей и гиперпараметров + позволил манипулировать моделями тем, кто не разбирается в ЯП.

## Куратор проекта – НИУ МГСУ

Октябрь 2022 — Май 2023

- **Разработка методических курсов:**
- Разрабатывал методические курсы для иностранных студентов.
- Написание прикладных программ по реализации методов из курса на Python.
- Руководство группой студентов в составе 4 человек.
- Разработка сервиса для размещения и хранения курсов.

## Образование и курсы НИУ ВШЭ — магистратура, Прикладная математика и информатика (2024–2026) НИУ МГСУ — бакалавриат, Прикладная математика (2020–2024)

GPA: 8.8. Победы в олимпиадах по математике и моделированию, статьи ВАК

GPA: 4.97 / 5. Победы в олимпиадах (математика, физика, ICPC), публикации РИНЦ-ВАК, руководство разработкой учебных курсов для иностранных студентов.

- Я-профессионал (Математическое моделирование) – призер 2025
- ICPC (1/4) – призер 2023
- ICEF HSE Summer School 2024: Econometrics / Mathematics