



Владислав Вододохов

Мужчина, 28 лет, родился 21 февраля 1995

+7 (915) 6533061 — Telegram: @vododokhov21
vododoh21@gmail.com — предпочитаемый способ связи

Проживает: Венгрия

Гражданство: Беларусь, есть разрешение на работу: Россия, Венгрия, Беларусь

Не готов к переезду, не готов к командировкам

Желаемая должность и зарплата

Data Scientist

Специализации:

- Дата-сайентист
- Аналитик

Занятость: полная занятость, стажировка

График работы: полный день, гибкий график, удаленная работа

Желательное время в пути до работы: не имеет значения

Опыт работы — 3 года 2 месяца

Декабрь 2021 —
Октябрь 2022
11 месяцев

Яндекс.Практикум

Информационные технологии, системная интеграция, интернет

- Интернет-компания (поисковики, платежные системы, соц.сети, информационно-познавательные и развлекательные ресурсы, продвижение сайтов и прочее)

Специалист по Data Science (учебный опыт)

- Овладел практическими навыками предобработки, исследовательского анализа, и визуализации данных с помощью библиотек языка Python Pandas и Seaborn.
- Применил библиотеку Scikit-learn для построения моделей машинного обучения для решения задач регрессии и классификации.
- Научился оценивать качество моделей машинного обучения с помощью кросс-валидации и различных метрик, а также улучшать его.
- Изучил получение данных из баз данных с помощью SQL.
- Использовал нейронные сети для построения модели классификации токсичных комментариев (BERT, nltk), а также определения возраста человека по фотографии (Keras).

Выполненные проекты:

- Выбор локации для скважины

(https://github.com/uladzislau21/YandexPracticum-projects/tree/main/oil_rig). Провел анализ данных нефтедобывающей компании и построил модель машинного обучения для выбора наиболее прибыльного региона для добычи нефти. С помощью техники Bootstrap оценил риски. Использованные технологии: Pandas, Seaborn, Scikit-learn.

- Предсказание стоимости автомобилей

(https://github.com/uladzislau21/YandexPracticum-projects/tree/main/car_price). На основе различных характеристик автомобилей построил модель, предсказывающую их стоимость. Применил градиентный бустинг. Сравнил несколько моделей между собой по скорости обучения и предсказания. Использованные технологии: Pandas, Seaborn, Scikit-learn, Lightgbm, Hyperopt.

- Определение токсичных комментариев

(https://github.com/uladzislau21/YandexPracticum-projects/tree/main/cmmnts_classif). Провел подготовку и анализ текстовых данных. Обучил логистическую регрессию на TF-IDF и эмбедингах из BERT, сравнил обе модели. Используемые технологии: Nltk, Scikit-learn, Numpy, Pandas, Torch, BERT.

Сентябрь 2019 —
Август 2022
3 года

Biocad

Москва, biocad.ru

Медицина, фармацевтика, аптеки

- Фармацевтическая продукция (производство)

Научный сотрудник

- Принимал активное участие в поддержке доклинических исследований генно-терапевтических препаратов, что позволило вывести один из продуктов на уровень клинических исследований.
- Проводил выделение ДНК/РНК из тканей и экскрементов животных.
- Проводил qPCR и qPCR-RT для количественного определения вирусных частиц в образцах.
- Внедрил роботизированную систему Тесла для автоматизации исследований, что позволило повысить эффективность работы до 400 образцов в день.
- Провел валидацию компьютеризованной системы.
- Написал 2 СОПа.

Получили звание лучшего сотрудника 2020 года за создание приложения для анализа данных с использованием языка программирования R (ссылка на репозиторий приложения: <https://github.com/uladzislau21/qPCR-Data-Wizard>).

Образование

Магистр

2026

University of Debrecen

Faculty of Medicine, Molecular Cellular and Immune Biology

2019

University of Debrecen

Faculty of Medicine, Molecular biology

Повышение квалификации, курсы

2022

Яндекс.Практикум

Яндекс, Специалист по Data Science

Электронные сертификаты

2021

Математический анализ
Теория вероятностей

2020

Введение в Linux
Линейная алгебра

Ключевые навыки

Знание языков

Белорусский — Родной
Английский — B2 — Средне-продвинутый
Немецкий — A1 — Начальный
Русский — C2 — В совершенстве

R Python SQL Jupyter Notebook Git Кластерный анализ
Статистика Теория вероятностей Математический анализ sklearn
seaborn PyTorch pandas Linux Visual Studio Code Bash (basic)
GitHub PyCharm RStudio shiny (R) tensorflow keras numpy
Визуализация данных Предобработка данных
Построение моделей машинного обучения CatBoost XGBoost
Computer vision (basic) NLP (basic)

Дополнительная информация

Обо мне

Моя область интересов – применение Data Science к анализу биологических данных, в частности, применение машинного обучения для лучшего понимания масс-спектрометрических данных. Работа в Data Science позволит мне развить новые навыки, которые позволят не только лучше понимать биологию, стоящую за данными, но и создавать новые, более эффективные методы обработки биологических данных (применение deep learning). Однако, в независимости от типа данных, я с большим энтузиазмом отношусь к их анализу.

Свои знания в данный момент расширяю чтением книги "Python для сложных задач: наука о данных и машинное обучение" и участием в соревновании <https://www.kaggle.com/competitions/amp-parkinsons-disease-progression-prediction>. Любимый канал на YouTube – Data Professor. В планах - поучаствовать в хакатоне или соревновании на Kaggle.

Люблю получать новые знания и опыт в короткие сроки – за полгода изучил язык программирования R и с использованием пакета shiny разработал приложение, сократившее анализ данных от нескольких дней до нескольких минут.