



Verbetsky Edward

Мужчина, 29 лет, родился 29 ноября 1994

+7 (930) 9472725 — предпочитаемый способ связи • Пишите сразу в Telegram:

https://t.me/Verbetsky_E □

verbasik2018@gmail.com

Другой сайт: <https://github.com/Verbasik>

Проживает: Москва

Гражданство: Россия, есть разрешение на работу: Россия

Готов к переезду, готов к командировкам

Желаемая должность и зарплата

Data science

Специализации:

- Программист, разработчик
- Научный специалист, исследователь

Занятость: полная занятость, частичная занятость, проектная работа

График работы: полный день, сменный график, гибкий график, удаленная работа

Желательное время в пути до работы: не более часа

Опыт работы — 3 года 4 месяца

Июль 2023 —
настоящее время
8 месяцев

Департамент информационных технологий Москвы

www.mos.ru/dit/

Team Lead Data Science

Ведущий Data Scientist / Разработчик Big Data

- Ответственный за end-to-end процессы создания data-продуктов в контуре BigData, от исследования релевантных данных, построения и проверки гипотез, создания признаков до тестирования и подготовки готового продукта к промышленной эксплуатации.

- Разрабатывал и внедрял сервис обработки запросов на основе больших языковых моделей (LLM), охватывающий весь цикл от исследования данных до интеграции в продукт.

- Разработка и внедрение моделей машинного обучения для решения бизнес-проблем.

- Управление командой и менторство младших сотрудников.

Стек знаний и технологий, которые использую:

Языки программирования:

PythonI, SQL, Bash

Библиотеки и инструменты:

Data Analysis & Visualization: Pandas, NumPy, SciPy, Matplotlib, Seaborn, statsmodels, Plotly.

Machine Learning & Deep Learning: Keras, TensorFlow, Scikit-learn, PyTorch, MLLib.

Big Data Technologies: Hadoop, PySpark, Apache, GreenPlum.

NLP & LLM: BERT, GPT, Transformers, NLTK, NLU, NER, RAG, NLP4Code, LLM4Code.

Инструменты совместной работы и управления проектами: Jira, Confluence, Git, GitHub, GitLab, Docker, Kubernetes.

Инструменты для работы с данными и анализа: JupyterLab, Vertica, Hive.

Инструменты для автоматизации и развертывания: Airflow, Jenkins, CI/CD pipelines.

Дополнительный опыт и проекты:

Участвовал в создании сервиса, использующего большие языковые модели (LLM). В его состав входили такие компоненты, как фильтр запросов, классификатор NLU, модель NER, генератор вопросов, система проверки синонимов, RAG хранилище, извлекатель контекста и генератор подсказок. Каждый микросервис был разработан как независимый модуль для масштабируемости и высоко производительности.

Ноябрь 2022 —
Ноябрь 2023
1 год 1 месяц

**Национальный научно-исследовательский институт
общественного здоровья имени Н.А. Семашко**

nriph.ru/

Data analytics

Работал в НИИ общественного здоровья имени Н.А. Семашко как эксперт в области анализа данных. Моя роль заключалась в оказании поддержки научным исследованиям, используя методы Data Science и Data Analytics. Работал сдельно.

Ноябрь 2021 —
Ноябрь 2022
1 год 1 месяц

BIOCAD

biocad.ru/

Data science/analytics

Опыт работы биоинформатиком - биостатистиком в BIOCAD:

Работал в роли аналитика данных с акцентом на анализ данных, статистическое моделирование и интерпретацию результатов. Специализировался на применении методов машинного обучения, включая разработку ML-моделей для задач прогнозирования, классификации и кластеризации для научных исследований. Транслировал аналитические данные и результаты обученных ML моделей для исследовательской команды и управляющего состава, облегчая понимание и применение сложных статистических концепций в рамках проектов.

Стек знаний и технологий, которые использовал:

Языки программирования: Python, SQL
Библиотеки и инструменты: Biopython, Pandas, NumPy, SciPy, Matplotlib, Seaborn, statsmodels, Plotly, Keras, TensorFlow.

Ноябрь 2020 —
Ноябрь 2021
1 год 1 месяц

BIOCAD

biocad.ru/

Системный администратор

Опыт работы системным администратором в BIOCAD:

Работал системным администратором, обеспечивая надежную работу инфраструктуры и сетевых ресурсов одного из центров BIOCAD.
Отвечал за установку, настройку и обслуживание серверного оборудования и сетевых устройств.
Осуществлял мониторинг и обеспечение безопасности сети, а также управление резервным

копированием данных.
Занимался установкой и настройкой программного обеспечения, включая операционные системы, базы данных и прикладные программы.
Оказывал техническую поддержку пользователей и решал возникающие проблемы с оборудованием и программным обеспечением.

Стек знаний и технологий, которые использовал:

Операционные системы: Windows, Linux.
Серверные платформы: Microsoft Windows Server, Linux (CentOS, Ubuntu).
Сетевые протоколы и устройства: TCP/IP, DNS, DHCP, Cisco, Juniper.
Виртуализация: VMware, Hyper-V.
Управление системами: Active Directory, LDAP.
Базы данных: MySQL, PostgreSQL.
Системы резервного копирования: Veeam, Bacula.
Установка и настройка программного обеспечения, техническая поддержка пользователей.
Мой опыт работы системным администратором и знание указанных технологий позволили мне эффективно управлять инфраструктурой и обеспечивать бесперебойную работу сети и серверов.

Образование

Бакалавр

2018 **Московский государственный университет им. М.В. Ломоносова, Москва**
Прикладная математика и информатика, Факультет вычислительной математики и кибернетики

Повышение квалификации, курсы

2024 **Профессиональная переподготовка «Data Science»**
Yandex practicum, Data Science

2024 **Математика для анализа данных и машинного обучения**
Yandex practicum , Data Science/Analysis

Ключевые навыки

Знание языков Русский — Родной
Английский — B1 — Средний

Навыки

Python

R

MATLAB

Bash

Data Analysis

Data Science

Pandas

NumPy

Matplotlib

Plotly

scikit-learn

Keras

selenium

TensorFlow

NLTK

MS Visual Studio

RStudio

Jupyter

Tableau

Git

GitHub

SQL

Jekyll

Hadoop

Jira

Airflow

PySpark

Дополнительная информация

Обо мне Магистр по наукам о данных с глубокой экспертизой в области Data Science, руковожу разработкой и внедрением data-продуктов. Моя экспертиза охватывает широкий спектр методов машинного обучения и анализа данных, применяемых для решения как научных,

так и бизнес-задач.

В сотрудничестве с "Национальным научно-исследовательским институтом общественного здоровья имени Н.А. Семашко", а так же во время работы в компании "Biocad", фокусировался на применении методов Data Science и Data Analytics для поддержки и развития научных исследований в медицине. Моя роль здесь включала в себя статистическое моделирование и интерпретацию результатов. Использовал Python и его библиотеки для анализа больших и сложных датасетов, особенно в контексте медицинских исследований. Разрабатывал ML модели биологической и медицинской направленности.

Так же имея опыт работы в роли Data Engineer, я специализируюсь на создании масштабируемых и надежных data pipeline'ов. Это обеспечивает эффективный сбор, хранение и обработку данных, их подготовку для дальнейшего анализа и моделирования. Я работаю с различными источниками данных и интегрирую их в единые системы для более эффективного и надежного анализа.

Моя особая страсть — это переплетение науки и технологий. Я активно углубляю свои знания в области биофизики и искусственного интеллекта. Моя любимая книга, "Что такое жизнь" Эрвина Шредингера, отражает мою веру в красоту и сложность научного исследования. В своем профессиональном развитии я ставлю акцент на применении математических методов для решения сложных биологических и медицинских задач. Я ценю структурированный и логичный подход к решению проблем и всегда стремлюсь быть на переднем крае современных научных и технологических разработок.