

Горбунов Егор Ильич

 godofflipping |  eigorbunov@edu.hse.ru |  godofflipping |  +7-965-134-13-85

SUMMARY

Data Scientist (МФТИ, выпускник 2024). Работал над проектами от ЦОСиВТ по анализу IT рынка РФ и разработке приложения для классификации клеток, где обучал различные модели CNN для классификации изображений. Также реализовывал различные ML/DL-модели, рекомендательные системы и A/B-тесты на датасетах с Kaggle. Владею следующими навыками, характерными для позиции Data Scientist: Python, SQL, Basic ML/DL, RecSys, A/B-тесты.

ОПЫТ РАБОТЫ

АО ЦОСиВТ

Июнь 2022 - Июнь 2024

1) Участвовал в проекте по анализу IT рынка РФ за 2019 - 2021 гг. На Python проводил ABC-анализ доходов компаний и их налогов, по индексу Херфиндаля — Хиршмана проверял IT отрасль на наличие монополии на рынке, а также использовал регрессионные модели для оценки влияния наличия аккредитации на экономические показатели IT рынка ([ссылка на публикацию](#)).

2) Участвовал в проекте по разработке приложения для сегментации и классификации клеток (в частности костного мозга). Занимался реализацией графического интерфейса на PyQt, а также подбором и обучением модели свёрточной нейронной сети для классификации клеток. Участвовал с этим продуктом в конкурсе УМНИК ([сертификат](#)).

ОБРАЗОВАНИЕ

Высшая Школа Экономики

Сентябрь 2024

Факультет: ФКН (Науки о Данных)

Московский Физико-Технический Институт

Сентябрь 2020 - Июнь 2024

Факультет: Физтех-школа Радиотехники и Компьютерных Технологий

Дополнительное образование ([Ссылка на профиль в Stepik](#))

- Основы статистики
- Свободное погружение в СУБД
- Интерактивный тренажер по SQL
- Deep Learning School

ДРУГИЕ ПРОЕКТЫ

Пет-проекты

[Ссылка на репозиторий в Github](#)

Анализировал датасеты с помощью алгоритмов машинного обучения таких, как Random Forest, Gradient Boosting, KNN и другие. Проводил A/B-тесты на различных датасетах с kaggle, связанных с маркетингом, e-commerce, мобильными играми и другими. Для поиска статистически значимых результатов в данных и проверки гипотез использовал метод bootstrap, z-тест, t-тест, U-критерий Манна — Уитни и другие. Также Реализовывал и сравнивал различные рекомендательные системы (Collaborative filtering на основе матричного разложения, Content-based filtering с TF-IDF преобразованием) на датасете по взаимодействию пользователей со статьями.

НАВЫКИ

Python	Pandas, NumPy, Matplotlib, SciPy, OpenCV, PIL, scikit-learn, PyTorch, PySpark
ML/DL	KNN, Linear Models, DecisionTree, RandomForest, Gradient Boosting, CNN Classification and Segmentation, RecSys
SQL	PostgreSQL