

# Горбунов Егор Ильич

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

## SUMMARY

---

Data Scientist (Москва). Закончил МФТИ в 2024. Работал над проектами от ЦОСиВТ по анализу IT рынка РФ и разработке приложения для классификации клеток, где обучал различные модели CV. 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, ML/DL, RecSys, A/B-тесты.

## ОПЫТ РАБОТЫ

---

### Data Scientist - АО ЦОСиВТ

Май 2022 - Июнь 2024

#### 1) Проект по анализу IT рынка РФ:

- Работал с СУБД PostgreSQL, с помощью библиотек Python проводил анализ данных и строил регрессионные и классификационные ML модели для определения влияния аккредитации на IT компании. В результате получил модели по определению дохода от наличия аккредитации у компании.
- С помощью pandas исследовал IT рынок на наличие монополии. Результатом стала статья в Computer Research and Modeling ([ссылка на публикацию](#)).

#### 2) Проект по разработке приложения для анализа медицинских изображений костного мозга:

- Подбирал и сравнивал различные CNN для классификации изображений клеток костного мозга. Также сравнивал и применял разные алгоритмы сегментации для медицинских изображений. Встраивал лучшие модели в MVP.
- Занимался реализацией GUI на PyQt для приложения. Результатом стал MVP, который обеспечивает возможность сегментировать и классифицировать изображения клеток костного мозга, выдавать диагноз пациента. Участвовал с этим проектом 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

## ДРУГИЕ ПРОЕКТЫ

---

### Пет-проекты

[Ссылка на репозиторий в Github](#)

- Анализировал датасеты на Kaggle с помощью моделей классического ML таких, как Random Forest, Gradient Boosting, KNN и других.
- Проводил A/B-тесты на различных датасетах с kaggle, связанных с маркетингом, e-commerce, мобильными играми и другими.
- Реализовывал и сравнивал различные рекомендательные системы (Коллаборативные на основе матричного разложения, контентные с TF-IDF преобразованием) на датасете по взаимодействию пользователей со статьями.

## НАВЫКИ

---

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