# Gleb Lukicov

Data Scientist | PhD in Physics


## Portfolio

https://glukicov.github.io

## Contact

London, UK


## Programming

`Python`, `C++`, `Cython`, `SQL`, `Bash`, `LaTeX`

## Data analysis

**Techniques**

Fourier transform
Monte Carlo methods
Iterative optimisation

**Tools**

`NumPy`, `pandas`, `SciPy`, `Matplotlib`, `seaborn`, `Dash`, `JupyterLab`

## Machine learning

**Techniques**

Regression
Classification
Neural networks
GPU utilisation

**Tools**

`sklearn`, `MLflow`, `CatBoost`, `SHAP`, `TensorFlow`, `Keras`

## Software

`Linux`, `PostgreSQL`, `Docker`, `VSCode`

## Languages

English (native)
Russian (native)
Latvian (intermediate)

## Interests

**Professional**

Education outreach
Technology blogging

**Personal**

Observational astronomy
Thai kickboxing


GitHub: glukicov
LinkedIn: glukicov
Medium: @lukicov
Twitter: @Gleb_Lukicov


## Profile

- Industry and research experience as a **full-stack data scientist**.
- 6 years of experience in applying advanced statistical methods to large datasets.
- Proficient in using `Python`, `C++`, `SQL`, as well as `sklearn` and `TensorFlow` pipelines.
- Comfortable with Unix development environment and modern DevOps tools.
- Practical experience with big-data collection, storage, processing, and analysis.
- Applied expertise with distributed computing systems: cloud, servers, grid, IoT.
- Strong communicator with 9 years of public speaking experience.

## Experience

**Just Eat Takeaway** | *Data Scientist*
October 2020 – November 2020 | London, UK
Six-week project, as part of the Faculty Fellowship, with the aim of increasing revenue by improving restaurants' menu-conversion.

- Analysed datasets of 45K restaurants and 9M products extracted from `BigQuery`.
- Developed `CatBoost` prediction model, based on the restaurant's segmentation level.
- Utilised the Shapley model-explainability method to rank feature importances.
- Deployed the model as a web-based API with robust visualisations
- Delivered an analysis report on data-driven insights into menu optimisation.
- Presented the project results in front of 200 companies.

**Fermilab** | *Researcher*
March 2017 – November 2019 | Chicago, USA
Collaborated on a physics experiment with 200 scientists and engineers.

- Led the development of *data optimisation software* (`Python`, `C++`) in an 18 person team.
- Derived calibration constants into the production `PostgreSQL` database, enabling the processing of 2 PB of data.
- Managed the effort to add extra distributed computing resources to the "common pool", and designed tools for data quality monitoring.
- Involved in the design, testing, and operation of detectors and data acquisition system, that recorded over 6 PB of raw data.

**Paul Scherrer Institute** | *Research Trainee*
June 2015 – September 2015 | Villigen, Switzerland
Worked as part of a team of hardware and software experts to ensure optimal data collection.

**University College London** | *Research Intern*
July 2014 – September 2014 | London, UK
Developed a QR-coded online database for over 300 research devices.

## Qualifications

2020 **Data Science Fellowship** **Faculty**
Highly competitive funded fellowship, providing project experience and intense eight-week training in machine learning and commercial skills:

- Unsupervised learning and dimensionality reduction
- Bayesian statistics and causality methods
- NLP, reinforcement learning and deep learning
- Visualisation tools, API and model deployment
- Project management, negotiation, data security

## Education

2016–2020 **PhD** in Experimental Particle Physics **University College London**
Thesis work focused on data collection, processing and analysis.
Graduate-level computer science courses in machine learning:

- Logistic regression, SVM, random forests, neural networks, CNN

2012–2016 **MSci** in Physics (1st class) **University College London**