Mihael Tunik

✉ *mihael.8112@yahoo.com*
https://github.com/mihael-tunik/

About me

Programmer with versatile experience in IT and computer science. I do believe that modern scientific research process requires significant programming skills (and ready to provide them). Seeking a position as a developer-researcher to enhance my career growth.

Education

2013 — 2017 **Bachelor degree**, *Saint-Petersburg*, Peter the Great St.Petersburg Polytechnic University, departament of applied mathematics and mechanics.

2017 — 2019 **Master degree**, *Saint-Petersburg*, Peter the Great St.Petersburg Polytechnic University, departament of applied mathematics and mechanics.

Master thesis

2019 **Special kernel density estimator for finite sample size conditions**.

Work was dedicated to research of theoretical accuracy of statistical kernel density estimator of special type for finite sample size conditions.

Experience: 4 years and 10 months

august 2019 — now **Saint-Petersburg State University, Chebyshev Laboratory**, *engineer-researcher*.

- Here, I started as an intern in the small team, where we're developing statistical instruments for geo-data analysis and seismic inversion. There we extensively used various Gaussian process based regression models and various techniques for data-processing.

  Typical tasks:
  - research for relevant scientific articles;
  - automate research pipeline;
  - integrate and test new submodule in codebase;

- Then, I continued to work as engineer-researcher on development the tool for fine-tuning advanced hydrodynamic simulations in Dumux with Bayesian optimization techniques. Among other things as a researcher I took part in implementing experimental software for solving Riemann problems.

  Typical tasks:
  - build custom desktop UI for our application (5K codebase from scratch);
  - reorganize project codebase, fix architecture issues;
  - rewrite algorithmic core for optimization;

- Latest project, where I work mostly with ML-pipelines for classification/recognition timeseries data from sensors of gas-analyzer. Developed baseline window-based approach for classification in combination with classic ML-tools like LDA/QDA or logistic regression. Took part in designing and implementation of multistaged classifier based on classic and gradient boosting models.

  Typical tasks:
  - Explore the data and develop strategies for handling it;
  - Develop project research pipeline completely from scratch;
  - Propose and develop different models for solving stated ML-problems;

- Actually during my work I've created even more things: like microservices for convenient remote access to advanced simulator software or gradio dashboards to make research process even more convenient.

Technical skills

Basically, I know three programming languages: **Python** and **C/C++**. During my work experience, which is long enough, I encountered many more: SQL, Javascript, R and Matlab.
I'm competent enough in fields of statistics and probability theory, linear algebra and calculus. Studied in university and still didn't forget algorithms and numerical methods.
In the latest projects I had a lot of practice with **statistical data analisys** and **ML** (hypothesis testing, binary/multiclass table data classification, linear models).

- General purpose skills:

  - Many years of experience with different **Linux** distributions, system configuration, terminal (bash, Unix commands);

  - Proficient with **Git**, managing repositories: pull-requests, **Github Actions** CI, reviews; **Notion** for task-tracking;

  - Remote access via **ssh**, familiar with **Docker** and **docker-compose**;

  - Extensive experience with **Python** toolchain and ecosystem: building up Python packages from scratch with **setuptools**, managing things with **venv** or **Anaconda**;

  - Testing with **pytest**, profiling with **cProfile** and **Sphinx** for automated documentation;

- Experience as engineer-researcher:

  - Work on project sketches in **Jupyter Notebooks** and **Google Colab**;

  - Proficient with **numpy, scipy, sklearn**; very familiar with **Pandas and Polars** dataframe libs; work with GP models via GPflow, botorch, bayes_opt, gradient boosting with **CatBoost/LightGBM/XGBoost**; certain experience with Tensorflow, Keras and Torch;

  - Work with model ensembles, familiar with **model stacking** techniques;

  - Advanced **LaTeX** for scientific texts and presentations;

- Some experience from desktop-dev:
  - UI development with **PyQt5**, Qt Creator IDE, PyInstaller for bulding binaries;
  - Experience in writing detailed documentation for code and UI;
- Some experience from web-dev:
  - Some experience from backend: HTTP protocol, Nginx, **Flask**, Django, testing APIs with **Postman**;
  - Basic experience with databases (PostgreSQL, ClickHouse, SQLite) and key-value stores;
  - Some experience from frontend: HTML, CSS/SCSS, static site generators like Hugo and Javascript basics;
- Some experience with C/C++ (parallel computations with OpenMP, make-files and CMake, building small .so libs), linking Python and C/C++ with **Python C API** or **Ctypes**;

Languages

Russian	C2, Native speaker
English	B2, Upper-Intermediate
German	A2, Beginner

Online courses

Stepik October 2022 ***Hadoop. System for big data processing.***

Learned basic things about Hadoop ecosystem, including HDFS and MapReduce.

Result: certificate with distinction, >90% score.

Stepik January 2023 ***Apache Airflow for analysts.***

Learned basic things about Apache Airflow, DAGs and ETL in general.

Result: certificate.