Mihael Tunik

mihael.8112@yahoo.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).

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: >5 years

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:
  - 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 window-based method for timeseries classification 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 custom desktop UI for one of ours subprojects.

Technical skills

Started my research career in fields of statistics and probability theory. Also I'm competent enough in numerical methods and algorithms.
In the recent projects I had a lot of practice with **statistical data analisys** and **ML** (hypothesis testing, feature engineering, timeseries data classification).

- General purpose skills:

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

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

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

  - Familiar with **Docker** and **docker-compose**;

  - Familiar with testing (**pytest**), profiling and automated documentation tools;

- Experience as engineer-researcher:

  - Proficient with **numpy, scipy, sklearn**; familiar with **Pandas and Polars** dataframe engines;
  - Decision trees and gradient boosting with **CatBoost/LightGBM/XGBoost**;
  - Worked with ensembles and various **model stacking** techniques, **multi-staged classifiers**;

  - Familiar with Tensorflow and Keras;

  - 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;
- Some experience with C/C++ (OpenMP, CMake, Valgrind and building small .so libs), **Python C API** and **Ctypes**, worked with low-level C API for XGBoost and Eigen libs;

Languages

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

Articles and preprints

- Classification of Graphene-Based Electronic Nose Measurements with Gradient-Boosted Decision Trees. Available at SSRN: https://ssrn.com/abstract=5041771

Personal webpage

https://mihael-tunik.github.io/

Here I write small articles about programming and computer science and make experiments with static site generators.