Ivan Pogrebnyak

Data Scientist

PhD in High Energy Physics

ivanp@msu.edu
(919) 215-5546
ivankp.github.io
Garner, NC

Experience

Over 10 years of experience with statistical analysis, data science, and programming obtained from working in experimental particle physics.

Higgs boson data analysis at CERN Large Hadron Collider Jun 2013 – Present

- Analyzed terabytes of experimental and simulated data.
- Developed interactive web-based analysis tools eliminating the need to manually perform recurring data categorization tasks.
- Implemented and applied (non-)linear statistical regression models.
- Optimization and significance testing; LLR, Wald and F tests.
- Uncertainty analysis using complex constrained likelihood models.
- Monte Carlo techniques for physics simulations and statistical tests.

DOE Science Graduate Student Research (SCGSR) fellowship Jan 2015 – Dec 2015

- Delivered requirements for improvement of an electronics monitoring system crusial for continued operation of the ATLAS experiment.

Statistical analysis of nuclear data at UNC May 2012 – Jun 2013

- Applied non-linear regression modeling.
- Optimized statistical models using likelihood and χ^2 methods.

Technology skills

- `C++`, `C`, `CERN ROOT`
  Application and implementation of data analysis tools, system programming, expert knowledge of modern `C++`, including `C++20`
- `Python`: `scikit-learn`, `numpy`, `matplotlib`, `requests`
  Data mining, data analysis, data visualization, and general scripting
- `Javascript`, `HTML`, `SVG`, `CSS`, `PHP`, `http`, `websockets`
  JS tools: `jQuery`, `d3.js`, `Highcharts`, `Plotly`
  Web development, frontend and backend
- `SQL`, `sqlite`: Database management, data aggregation
- `Git`, `GitHub`, `GitLab`, `SVN`: Software version control and collaboration
- `HTCondor`, `DAGMan`: Distributed computing
- `GNU+Linux` OS and utilities
- Other: `Java`, `Mathematica`, `VHDL`, `Xilinx ISE`, `Geant4`.

Education

2013 – 2021
Michigan State University
PhD, High Energy Physics

2009 – 2013
University of North Carolina at Chapel Hill
BS with Honors, Physics
BA, Mathematics

Affiliation

2013 – 2021
CERN
Large Hadron Collider
ATLAS Experiment

Data science skills

- Data mining
- Regression modeling
- Hypothesis testing
- Machine learning
- Multivariate analysis
- Uncertainty analysis
- Optimization algorithms
- Big data
- Distributed computing
- Monte Carlo simulations
- Gaussian processes
- Data visualization
- Interpretation and presentation

Language fluency

English, Russian, Ukrainian

General skills

- Collaboration: worked in small (1–5 people) and large (10–50 people) analysis groups.
- Written communication: directly contributed to 17 published articles and reports.
- Verbal communication: conducted over 50 presentations at meetings at different levels of collaboration.
- Initiative and creativity: learned web development, server programming, and database operation on my own to simplify, automate, visualize, and share data analysis algorithms and results.
- Fast learner: demonstrated history of learning technology for new data analysis and programming projects.
- Supervision and mentoring: trained and mentored undergraduate and graduate students in programming, physics, and research techniques, such as using `C++` and `python` for data analysis.
- Project and time management: always met deadlines on concurrent and collaborative time-critical projects.