# ILIA ZAITSEV

Data Scientist / Machine Learning Engineer

20257, Hamburg, Hamburg Area, Germany
iliazaitsev.me
linkedin.com/in/ilia-zaitsev
medium.com/@iliazaitsev

## PROFESSIONAL PROFILE

Data Scientist and Machine Learning Engineer with 8 years of applied data science, machine learning, and software development experience. Advanced knowledge of software design patterns, fluent in Python, learning Rust, strong in object-oriented and functional programming paradigms and in technical writing. Currently working at Blue Yonder GmbH as a Staff Data Scientist. Excels at applying AI/ML theory to commercial development in high volume data processing, visualisation, and model development.

## KEY SKILLS

**Machine Learning**: PyTorch, Scikit-Learn, TensorFlow, Deep Learning, Computer Vision, Boosted Trees, Statistical Analysis, Model Complexity & Tuning, Regression vs Classification Problems, Reinforcement Learning
**Data Processing & Visualisation**: numpy, pandas, Dask, Jupyter, matplotlib, seaborn, opencv
**Programming**: Object-Oriented & Functional Programming, Python, Rust
**Cloud Services:** AWS SageMaker, EMR, S3, Databricks, Azure, Google Cloud Compute
**Databases, Backend & Environments**: PostgreSQL, SQLAlchemy, Flask, FastAPI, macOS, Linux
**DevOps**: Docker, Jenkins
**SW Design & IDE Tools**: PyCharm, VS Code, pdb, vim
**Test-driven development**: xUnit, pytest

## CAREER SUMMARY

**Staff Data Scientist, Blue Yonder Inc., June 2020 – Present**

- Worked on distributed data processing systems training regression models on large volume of transactional data
- Developed and maintained a set of internally used data management and data processing libraries
- Participated in building of a new Machine Learning product oriented on retail business clients selling short-lifetime goods
- Analized and built regression models for predicting utilization of goods for various retail businesses
- Participated in product's design sessions, codebase maintenance and migration to a new set of model's training tools

**Machine Learning Engineer, Smaato Inc., June 2019 – May 2020**

- Worked on scalable Machine Learning pipeline capable to train Deep Learning models on huge data volumes to improve currently established algorithms with data-driven solutions
- Participated in building a Python package unifying data management and model training processes company-wide
- Worked on building CI/CD pipeline releasing machine learning package and data pipelines
- Analyzed ads providers bidding patterns and developed Machine Learning models to reduce the outbound traffic costs without revenue impact

**Kaggle: Generative Dog Images (Competition), June 2019 – August 2019**

- Participated in data competition implementing various Generative Adversarial Models architectures and losses in Keras and PyTorch
- Won a bronze medal (67th place of 927 teams)

**Python Developer (Contractor), August 2014 – March 2019**

- Worked with programmers, data scientists, technical teams and product owners on diverse projects, authored technical articles on machine learning algorithms
- In-commodities. Developing a set of trading tools for data scientists and commodities traders
- Windsor.ai. Improved codebase tests coverage and eliminated bugs for toolset that analyses customer journeys and enables multi-touch attribution modelling to optimise client marketing budgets
- Viziphi. Developed a visualization tool for financial data plotting, leveraged knowledge of seaborn/matplotlib Python libraries to generate 2D image plots with critical financial information (loss/return heatmaps, confusion matrixes, etc.)
- ZenVPN. Ported and improved OpenVPN-based Windows VPN client to macOS, made adjustments to account for differences between platforms
- Klue. Developed a set of scripts, backend endpoints and ORM models for predictive food consumption system
- Private Client/Entrepreneur. Implemented a decision-tree based classifier for a tabular dataset classification for Proof-of-Concept
- Whippletree Research. Analyzed medical treatment information dataset using various classification algorithms (Random Forests, SVM, Gradients Boosting)

**Python Developer & Machine Learning Engineer (Contract), VectorScient, December 2015 - May 2018**

- Developed a configurable machine learning pipeline with visualizing and reporting system to preprocess and classify database-stored e-commerce information for B2B e-commerce clients of Predictally, wrote package to predict future behavior from millions of transactions per client

**Face Emotions Recognition System Engineer (Contract), December 2016 - March 2017**

- Developed a model to classify human face emotions in Computer Vision during Udacity MOOC
- Authored a technical paper on software development and results verification

**Python Developer, BASE Capital (Contract), May 2016 - March 2017**

- Fixed scripts and increased speed of script execution by factor of 5, as verified by the product owner
- Improved and refactored legacy data processing codebase
- Developed a web UI to access CLI-based tools running loan volume prediction models.

## EDUCATION & PROFESSIONAL DEVELOPMENT

- Software Engineering, 5-years undergraduate degree, Surgut State University, Russia, September 2009 - July 2014
- Artificial Intelligence Nanodegree, Udacity, September 2017 - December 2017
- Machine Learning Engineer Nanodegree, Udacity, September 2016 - April 2017
- CS188.1x Artificial Intelligence / edX, BerkeleyX, February 2015 - May 2016.