JOHN PATON

Machine Learning Engineer

Haarlem, NL
+31 6 10372314
john@johnpaton.net
johnpaton.net
github.com/johnpaton
john-paton

WHO AM I?

I'm an enthusiastic machine learning engineer with a background in theoretical physics. I like solving problems, automation, clean code, and corny jokes. During my time at Catawiki I grew into more of an MLE position when we migrated our data science platform from a PaaS to Kubernetes on GCP. I was responsible for architecting and implementing the new platform as well as ensuring maximal ease of use for all data users. This is how I discovered my passion for software that lets other data scientists do their work with minimal cognitive overhead, and for ensuring that services I am responsible for are set up for current and future success.

EXPERIENCE

2020 – present **Lead Data Scientist - Machine Learning** **Catawiki**
- Ownership of Catawiki's data science & machine learning platform (Google Kubernetes Engine, various GCP managed services, FastAPI+nginx model serving framework).
- Developed a "one-stop-shop" ML service providing anti-fraud intelligence asynchronously to various services (classical ML, Kafka, MySQL).
- Development and maintenance of several existing ML models & applications in the domains of Engagement and Fraud Prevention (classical ML, LightFM recommender).

2019 – 2020 **Senior Data Scientist - Machine Learning** **Catawiki**
- Lead migration of all Data Science infrastructure from a PaaS to an in-house platform on Kubernetes (Google Kubernetes Engine).
- Revamped recommendations model, improving coverage, training times, and predictive power, using a (customized variant of) LightFM.
- Created user trustworthiness model, allowing trusted users to have a less restrictive experience with minimized risk to the company.

2016 – 2019 **Data Scientist & Consultant** **KPMG**
Maximizing the value of data for clients through exploratory data analyses, data science proofs of concept, and production machine learning projects as part of the Advanced Analytics & Big Data team. Advisory services for organizations seeking to become more data-driven.

EDUCATION

2014 – 2016 **M.Sc. Theoretical Physics** **Uppsala University**
Followed the "Strings and Quantum Fields" track, including quantum field theory, string theory, general relativity, and various other topics from physics and mathematics. Graduated in June 2016 after the completion of a one semester (30ECT) master thesis on string theory.

2011 – 2014 **B.Sc. Physics** **Amsterdam University College**
Joint international honors academy of the University of Amsterdam and the VU University Amsterdam. Graduated *summa cum laude* with a thesis on computational plasma physics for fusion energy.

LANGUAGES

English - native
Dutch - fluent

Python - advanced
Javascript - basic

TOOLS & TECHNOLOGIES

Git, Kubernetes, GCP (Kubernetes Engine, BigQuery, Cloud Build), Docker, SciPy stack, scikit-learn, (Py)Spark, FastAPI, nginx, Kafka

TECHNIQUES & METHODOLOGIES

Mathematical background, classical ML, recommendations, microservices, CI/CD, MLOps, cloud native machine learning, open source software