Marnix Koops

marnixkoops@gmail.com | +31622750810 | Amsterdam, Netherlands | marnixkoops.github.io

EXPERIENCE

QuantumBlack, AI by McKinsey

Principal Data Scientist — Sep 2020 — Present

- Leading projects and prototyping, developing, and implementing ML models across industries to solve business problems
- Building AI solutions with cross-functional teams to innovate and accelerate R&D in domains like drug- and materials discovery

Coolblue

Lead Data Scientist — ML — Aug 2019 — Sep 2020

- Played ping pong, researched and built auto-regressive neural networks for sequence-based embedding recommender systems
- Implemented recommender systems on website and in the app together with multi-disciplinary product teams

Machine Learning Engineer — Apr 2018 — Aug 2019

- Researched, built, and implemented ML models to make customers smile :)

Blue Field Agency

Research Scientist — Feb 2017 — Jan 2018

EDUCATION

Erasmus University Rotterdam

MS in Mathematical Statistics — Sep 2016 — Dec 2017

- Research paper rewarded 9/10 on Gaussian mixture modeling with likelihood penalization

BS in Mathematical Statistics, Econometrics — Sep 2015 — Jul 2016

- Courses include: multivariate stats, prob theory, statistical learning, ML theory, optimization, Bayesian stats, timeseries modeling

Delft University of Technology

BS in Engineering & Geosciences — Sep 2011 — Jul 2014

- Research paper rewarded 9/10 on oil and gas reservoir simulation modeling to translate lab experiments to field scale
- Courses include: calculus, linear algebra, numerical maths, signal processing, geophysics, thermodynamics, mineralogy, matlab

PROJECTS

Molecule discovery — Quantumblack, 2023

- Partnered with one of the biggest US mining companies in worlds first AI-driven metal leaching R&D effort to improve efficacy while reducing environmental impact on a massive scale
- Responsible for building AI system, combining GNN- and Transformer-based molecular foundational models with more traditional chemical representation algorithms to discover, rank, and evaluate molecules
- Built a closed-loop research system together with chemists and other domain experts, including a dedicated testing wet lab

Compound embedding — QB Labs, 2023

- Generalised Python asset with production-ready models to predict structure, function, or reaction properties of molecules
- Codebase has two main pipelines; embeddings generation, and fine-tuning on biochemical data for downstream modeling
- Framework is used for experimentation, and as starting point in partnerships with major biochemical companies

mRNA vaccines — QuantumBlack, 2022

- Partnered with new research centre of a leading PharmaCo to design, develop and industrialise AI solutions to research and produce the next generation of mRNA vaccines
- Part of multi-disciplinary team from engineers to computational biologists to researchers and UI designers
- Responsible for in-silico modeling of lipid nanoparticle properties, from representation learning to down-stream target prediction to identify and rank the most promising candidates to sent to the wet lab for in-vitro testing

Bioluminescence — QuantumBlack, 2021

- Implemented computer-vision algorithms such as Faster R-CNN for object detection and image segmentation applied for in-vivo bioluminescence imaging experiments to develop novel cancer treatments
- Built codebase to enabled rapid model prototyping and experimentation with new ideas while speeding up research timelines

Lightning-MF — Fun, 2021

- Implemented the classic Matrix Factorization for Recommender Systems algorithm as pytorch lightning module

Embedding vector search — Fun, 2020

- Built a simple and lightweight Python package for fast vector similarity search using Approximate Nearest Neighbors

Sequence embedding recommendation — Coolblue, 2019

- Researched, built and compared traditional matrix factorization algorithms and sequence-based embedding neural nets like LSTMs and GRUs for personalized item recommendation

SKLLS

- tensorflow, keras, pytorch, pytorch-lightning, sklearn, xgboost, lightgbm, numpy, pandas, scipy, mlflow