

MEHARJI A

PhD, Senior Data Scientist - Predictive and Recommender Systems | Applied AI/ML Engineer

Helsinki, Finland

SUMMARY

Senior Data Scientist and Applied AI/ML Engineer (PhD) with 6+ years of experience taking applied ML problems end-to-end, from problem framing and data analysis through model development, offline evaluation, experimentation, production deployment, and monitoring, for high-stakes business decisions. Experienced in building predictive regressors, classifiers, and recommender systems that balance model quality with real-world product constraints, including data availability, business rules, and changing conditions. Hands-on experience developing AI agents using Google ADK and integrating LLMs via OpenAI APIs into production data and ML workflows, including prompt engineering, data generation, and enrichment. Strong track record of partnering with engineering, product, and business stakeholders to translate ambiguous problems into robust ML/AI solutions and measurable business impact.

CORE SKILLS

- Applied ML & Modelling:** predictive modelling (regressors, classifiers), recommender systems (retrieval, ranking, personalisation), and deep learning (NLP, CNN, ANN).
- Experimentation & Evaluation:** A/B testing & experiment design, offline evaluation metrics, statistical inference, hypothesis testing
- Programming & Large-Scale Data:** Python, R, R Shiny Server, SQL/PostgreSQL, FastAPI, TensorFlow, PyTorch, LightGBM (regressors, classifiers & LambdaMART), pandas, scikit-learn, NumPy
- MLOps & Production:** AWS, MLflow (experiment tracking, model registry, deployment), Docker, Apache Airflow, Argo Workflows, CI/CD, production monitoring
- Agents & GenAI:** AI agents (Google ADK, LangChain, CrewAI), prompt engineering, LLM integration (OpenAI)
- Collaboration:** JIRA, Confluence, Miro, Slack

PROFESSIONAL EXPERIENCE

Senior Data Scientist, SkenarioLabs, Helsinki	<i>05/2022 - Present</i>
- Drove end-to-end development of quantitative and categorical predictive models for bank real estate portfolio valuations, covering model development and offline evaluation, experimentation, and authoring white papers on modelling methodologies, results, and business applications. - Designed and built a recommender system for a real estate PropTech platform, covering retrieval, ranking, and personalization to surface relevant property listings to users. - Integrated LLMs via OpenAI APIs into production ETL pipelines, designing prompt-engineering workflows for automated data generation, enrichment, and transformation. - Own the MLOps lifecycle end-to-end with MLflow (experiment tracking, model registry) monitoring, and performance tracking, ensuring models remain reliable in production at scale. - Collaborate closely with engineering, product, and external stakeholders to turn ambiguous business problems into production-ready, measurable ML/AI solutions; mentor team members on code review, testing, and Git/GitFlow.	
Data Scientist, CSC - IT Center for Science, Helsinki	<i>01/2022 - 05/2022</i>
- Led a 2-person team building a full-stack data analytics dashboard platform, owning requirements, backend, and deployment end-to-end. - Architected production ETL pipelines (SQL, Apache Airflow) for large-scale data processing, improving the reliability and scalability of analytics workflows. - Delivered hands-on ML training workshops, providing technical guidance to analysts on Python (pandas, scikit-learn, NumPy) and production ML best practices.	
Data Scientist, University of Helsinki, Helsinki	<i>2020 - 2022</i>
- Led an EU-funded project from design through implementation with a 4-person cross-functional team. - Developed and open-sourced a full-stack interactive dashboard application for exploratory data analysis and visualization.	

Doctoral Researcher, University of Helsinki, Helsinki*2015 - 2020*

- Applied statistical modelling and hypothesis testing to large-scale observational datasets, including predictive modelling of gene-disease associations.
- Developed ETL/analysis pipelines and interactive dashboards to communicate statistical findings to research stakeholders.

Data Analyst (Part-Time), Fimlab Laboratoriot Oy, Tampere*2016 - 2023*

- Build automated data pipelines (Python/Shell/R) and produce analytical reports with actionable, statistically grounded recommendations for stakeholders.

SELECTED PROJECTS

Predictive Analytics & Scoring Platform, SkenarioLabs*2022 - Present*

- Built end-to-end predictive models with scalable, cloud-based ETL pipelines and geospatial feature engineering to rank and prioritize real estate opportunities by predicted market value; deployed FastAPI endpoints for real-time predictions

Recommender System for Real Estate Platform, SkenarioLabs*2022 - Present*

- Built an end-to-end recommender system for a real estate PropTech platform, covering candidate retrieval, ranking, and personalization to surface relevant property listings tailored to individual user preferences and behavior.

ML-Driven Classification System, SkenarioLabs*2022 - Present*

- Developed classification models to prioritize and score energy consumption/solution predictions for client decision-making; deployed production-ready ML services on AWS with monitoring and performance tracking.

Multi-Label Document Classification (Deep Learning / NLP), KPMG, Helsinki*2018*

- Built a deep learning NLP solution (TensorFlow, transformers) for automated content categorization and relevance tagging, with feature engineering pipelines extracting semantic representations from unstructured text and a production API for real-time classification.

Computer Vision Classification System (Deep Learning), University of Helsinki*2021*

- Built a CNN-based image classification system (PyTorch) for automated detection and diagnosis, using data augmentation to improve robustness on limited datasets.

Predictive Healthcare Analytics (Deep Learning), University of Helsinki*2021*

- Developed ANN models for multi-class prediction in medical diagnostics, with a full training/evaluation/optimization pipeline and monitoring for prediction reliability.

EDUCATION

PhD, Bioinformatics, University of Helsinki, 2020

MSc (Tech), Bioinformatics, Aalto University, 2013

MSc (Tech), Biotechnology, Andhra University, 2009

CERTIFICATIONS

AWS Certified AI Practitioner (AIF-C01), In Progress

PUBLICATIONS

30 peer-reviewed journal publications.