

NGOC MINH TRAN

Dublin, Ireland | +353892493332 | minhitbk@gmail.com | Irish and Vietnamese Citizen

PhD. in computer science in the field of applied statistics and machine learning, 15+ years experience with data science and engineering, pursuing a career path in this area. Open for opportunities in Dublin (Ireland) or Ho Chi Minh City (Vietnam) or remote (US, EU with possible traveling). **Links within the CV are provided as track records.**

EDUCATION

- | | |
|--|-------------|
| • Postdoc, Institute for Research in Computer Science and Automation (INRIA), France | 2012 – 2013 |
| • PhD. in Computer Science, Leiden University, The Netherlands | 2008 – 2012 |

ACHIEVEMENTS

Awards

- [IBM Research Division Award in Security AI 2019.](#)
- [IBM Research Division Award in Auto AI 2019.](#)
- [IBM Outstanding Technical Achievement Award in Security AI 2019.](#)
- [IBM Outstanding Technical Achievement Award in Federated AI for Hybrid Cloud and Edge 2021.](#)
- [IBM First Plateau Invention Achievement Award 2019.](#)
- [IBM Second Plateau Invention Achievement Award 2019.](#)
- [IBM Third Plateau Invention Achievement Award 2020.](#)

Open Source Recognitions

- [Open Source Strategic Committer in 2019.](#)
- [Pull-Shark achievements 2 times in Github.](#)
- [Arctic Code Vault contributor in Github.](#)

Publications and Patents

- First author (10) and co-author (6) of 16 scientific papers in artificial intelligence (AI) and clouds.
- Lead inventor (12) and co-inventor (9) of 21 US and EU patents in AI domains like security AI, federated AI, auto AI, generative AI and Fintech AI.
- See [Google Scholar](#) for the full list of publications and patents.

Professional Certificates

- [Developer Certification for Apache Spark](#), certified by Databricks and O'Reilly Media, Jan 2016.
- [Oracle Certified Professional, Java SE 6 Programmer](#), Oct 2015.
- [Revolution R Enterprise Certified Specialist](#), Aug 2015.
- [Cloudera Certified Specialist in Apache HBase](#), Dec 2014.
- [Cloudera Certified Developer for Apache Hadoop](#), Oct 2014.
- [Cloudera Certified Administrator for Apache Hadoop](#), Oct 2014.

TECHNICAL AND PROJECT MANAGEMENT SKILLS

• AI, machine learning, deep learning, genAI, LLM	Very good
• Python	Very good
• Hadoop, PySpark	Good
• Tensorflow, Pytorch, Scikit Learn, CICD, MLOps, Docker container	Good
• Hugging Face, Mlflow, Databricks, SQL	Experienced
• Agile Scrum, Jira, Git	Experienced
• Kubernetes, AWS	Knowledgeable

PROFESSIONAL EXPERIENCE

Data Science Tech Lead at Mastercard in Ireland (2 years and 8 months)	Oct 2021 – Current
<ul style="list-style-type: none"> Tech lead of several generative AI / LLM / data science projects such as: <ul style="list-style-type: none"> Build a RAG system on Databricks, fine-tune foundational LLM models on internal data to support cross-functional departments to onboard new projects. Predict credit risk score for SMEs, estimate merchant utilization, forecast liquidity, predict churn customers. Lead a globally distributed team of data scientists across the US and EU. Collaborate with multiple stakeholders including product managers across the projects. 	
Staff Research Engineer at IBM Research Lab in Ireland (4 years and 8 months)	Feb 2017 – Oct 2021
<ul style="list-style-type: none"> Lead research projects for papers and patents in AI such as Automation of Image Data Preprocessing (AutoDP). Be a key member of the open source AI project Adversarial Robustness Toolbox (ART), play the role of architect, maintainer, reviewer and contributor with 1355 commits, 73423++ and 45949-- of lines of codes. ART is supported by the US Defense Advanced Research Projects Agency (DARPA), contract number HR001120C0013. ART has received 4500+ stars, 1100+ forks on Github and has been cited by 570+ research papers. Contribute to the key AI project of IBM, the Federated Learning (IBMFL). IBMFL is a Python framework for federated learning in an enterprise environment. It has 2 versions: an enterprise version and an open source version, which has received 480+ stars, 130+ forks on Github and has been cited by 130+ research papers. Contribute to the key AI project of IBM, the One Button Machine, which belongs to the AutoAI domain and is implemented as a Watson Machine Learning service of IBM Watson Studio. 	
Senior Data Scientist at ARIMO (Panasonic, Remote)	2016 – 2017
ARIMO is California-based leader in deep learning and AI, ranked among the ten most innovative startup companies in data science in 2016 . My work includes developing IoT-centric AI models for commercial and manufacturing applications.	
Solution Architect (Data Science) at DirecTV (AT&T, Remote)	2014 – 2016
<ul style="list-style-type: none"> Manage and coach an R&D team of junior data scientists. Data science tech lead of big data project, which aims at building a cloud-based recommendation engine to recommend TV/movies/sport programs to ~40 millions end users of DirecTV. 	
Postdoc at INRIA, France	2012 – 2013
Involve in an EU-FP7 project , which aims at incorporating innovative hardware and network technologies seamlessly into data centers that provide PaaS cloud infrastructures.	
PhD candidate at Leiden University, The Netherlands	2008 – 2012
My PhD mainly focuses on statistical data analysis, modeling and machine learning, especially with stochastic random processes and long-term time series in the domain of clusters and clouds.	