Curriculum Vitae

Tahir Muhammad

Tahir.muhammad@mail.utoronto.ca
Homepage.com/Tahir
Linkedin.com/Tahir

Research Interests

My research interests are broadly focused on questions related to artificial intelligence, reinforcement learning, deep learning and statistical learning theory. I want to answer questions that enable us to understand deep learning on a new paradigm – allowing us to focus on AI alignment. In terms of application of AI, I want to focus on building tools which allows teams to build un-biased and equitable models, detect fraud entities and illicit behavior, and applications aimed at uplifting societies.

Education

University of Toronto Sep. 2015 – June 2021

Honours Bachelors of Science – Mathematical and Computational Sciences

- GPA: 3.7/4.0
- Specialization: Major in Applied Statistics and Computer Science, Minor in Economics
- Research Assistant: AI Researcher for Anti-Money Laundering, Entropy Estimation for Censored Data
- Graduate Courses: Machine Learning, Deep Learning, Advanced Statistical Learning, Multivariate Regression

Work Experience

Software Engineer – Big Data Platform Feb. 2023 – Present

BlueCat Networks

- Develop scale-able data pipelines using modern tech stack to ingest 10 million rows of DNS data / day
- Optimize SQL queries to reduce report generation latency time by 27% for all company wide reports
- Replace dependency of large reports from Hadoop to Apache Iceberg, reducing AWS glue-crawlers usage by 87%, leading to a reduction of our datalake costs by 31%

Data Scientist – Data Strategy & Engineering Dec. 2021 – Feb. 2023

Toronto Dominion Bank

- Strategize with VPs, Business Stakeholders, and engineers to optimize different departments using data science
- Develop data visualization dashboards & automation tools, utilizing Python, Databricks, and Tableau
- Finance and Risk Data Lead, helping transfer terabytes of TD's on-perm data onto the MS Azure Cloud

Junior Machine Learning Engineer Sep. 2021 – Dec. 2021

Toronto Dominion Bank

- Development of NLP and Optical Character Recognition (OCR) models PyTorch & TensorFlow
- Developed data pipelines, created ML model web apps (Streamlit), and deployed to departmental platforms
- Utilize advanced models such as BERT, ELMo to reach text classification accuracy of 91%

Innovation Data Science Intern May 2020 – May 2021

Royal Bank of Canada

- Perform data mining, analysis, visualization, and model development to create innovative solutions for GRM
- Develop recommendation engine for data access management with a NLP chatbot for user-interaction
- Collaborate with Borealis AI team to utilize deep learning to conduct entity recognition on entire GRM datalake
- Utilize tools such as Python, PowerShell, R, Spark, Elastic Search, Hadoop & Hive, Dremio, SQL and Tableau

Co-Founder & Software Engineer May 2019 – Jan. 2020

Nox

- Developed a full-stack web app to allow professors to receive feedback on audience's understanding in real-time
- Utilized JavaScript, React, MongoDB, APIs and Express server to build the site with a graphical dashboard
- Scaled to 20,000 students across the 3 campuses, and verified by 5 professors at UofT

Research & Publications

FruadIT | *AI Research Engineer* May 2020 – Apr. 2022

- Led a team of 4 ML engineers to investigate whether a company is requesting a fraudulent loan, to detect money laundering transactions & human trafficking entities given terabytes of transaction data
- Implement End-to-End Data Science workflow with final model reaching a false positive rate $< 2\%$

Monte Carlo Methods | *Researcher & Engineer* Sep. 2020 – Sep. 2021

- Develop & implement algorithms for estimation of entropy and extropy for type II censored data (cancer patients)
- Co-Author and published a paper in *Journal of Monte Carlo Methods and Applications*

Relevant Community Involvement

MedicAI | *Machine Learning Engineer* Oct. 2023 - Present

- Develop and deploy a deep learning model (CNNS, VGGNet, ResNets) to detect Pneumonia in images
- Leveraged TensorFlow, React, Express, MongoDB, NextJS, NodeJS and AWS – Sagemaker, Lambda, & EKS

QuranJourney | *Full Stack Software Engineer* May 2021 – Jan. 2022

- Build, design app architecture, implement and deploy advanced features for the biggest mosque in Canada (ISNA)
- Developed sophisticated tools for analysis and studying the Quran and Hadiths, utilized by over 1000 users/month

Glucose Vision | *Consultant - ML Engineer* Sep. 2021 – Dec. 2021

- Guide development team through computer vision tasks such as object detection and image segmentation
- Eliminated team's blocker of past 4 months within 4 weeks of consultation

RBC Technical Curriculum | *President, Software Engineering Instructor* Sep. 2020 – May 2021

- Instructed multiple python boot camps to help the business teams become more technically advanced
- Founded the Enterprise Risk chapter where I organized, lead, taught and empowered others to build with code

Teaching Experience (University of Toronto)

- CSC311: *Machine Learning* — Teaching Assistant (Enrollment: 36 students) 2021
- STA238: *Data Analysis II* — Teaching Assistant (56 students) 2021
- STA218: *Statistics for Management* — Head Teaching Assistant (70 students) 2021
- STA130: *Introduction to Data Science* — Teaching Assistant (37 students) 2020
- STA107: *Probability and Reasoning*— Teaching Assistant (46 students) 2020
- MAT223: *Linear Algebra I* — FSG Leader (12 students) 2020
- CSC108: *Introduction to Programming* — FSG Leader (15 students) 2020

Certifications & Achievements

- Generative AI Course | Grade: 100% 2023
- Supervised Learning Course | Grade: 100% 2022
- Advanced Learning Algorithms Course | Grade: 100% 2022
- Unsupervised Learning, Recommenders, Reinforcement Learning Course | Grade: 99.18% 2022
- Neural Networks and Deep Learning Course | Grade: 97.3% 2021
- $4^{th}/300$ students for Big Data and AI Case Competition 2021
- Best Research Implementation Project for Nox — (Implemented in UofT CS classrooms after) 2019

Technical Skills

Languages: Python, Java, C/C++, SQL (Postgres), JavaScript, Linux (bash), Spark & Scala
Packages: pandas, NumPy, Matplotlib, PyTorch, scikit-learn, Statsmodel, Tensorflow, MLFlow
Software: AWS (S3, Lambda, EKS), Apache Flink, Docker, Airflow, Kubernetes, Databricks & Gitlab CI/CD
Statistics & ML: Supervised & Unsupervised Learning, Time-series, Neural Networks, Bayesian Stats, A/B Testing

References

Prof. Luai Al Labadi
UofT | Statistical Sciences
luai.allabadi@utoronto.ca

Dr. Carlo Lisi
TD Bank | Sr. Manager
carlo.lisi@gmail.com

Roxana Zamfir
RBC | Director
zamfir.roxana@gmail.com