Saruultugs Batbayar

sbatbaya@ur.rochester.edu | +1(234) 296-3598 | saruulmeister.github.io/portfolio | saruulmeister | saruulmeister

Professional Summary

Data professional with 4 years' experience in banking and fintech, specializing in ML, quantitative methods, and big data analytics. Skilled in predictive modeling and real-time data solutions that drive strategic decisions. Expanding expertise in Generative AI, image processing, and network science.

Education

Master's in Data Science — University of Rochester, New York, USA *Aug 2024 – Dec 2025*

Bachelor's in Data Science — American University of Mongolia, Ulaanbaatar, Mongolia *Aug 2018 – May 2021*

Honors and Awards

- Fulbright Scholarship: Fully funded by the U.S and Mongolian governments
- Magna cum laude: American University of Mongolia (GPA: 3.83/4.00)

Technical Skills & Certifications Certifications Link

Languages: Python, R, SQL, Matlab, JavaScript **NLP & GenAI:** BERT, GPT-4, PEFT/LoRA **Libraries:** Scikit-learn, TensorFlow, Pytorch, pandas, numpy, networkX, nltk **Algorithms:** Linear Regression, Decision Tree, Clustering, Random Forest, XgBoost, Gradient Descent, GNN, CNN, GAN, DDPM, RePaint, LDM **Cloud:** Azure, IBM Cloud, Docker **Big Data & Distributed Computing:** Spark, Databricks, Delta Lake, cuda **Certifications:** The Science of Uncertainty and Data (MIT), SQL for Data Science (IBM), Visualizing Data with Python (IMB)

Work Experience

Information Technology Auditor — Khan Bank, Ulaanbaatar, Mongolia *Jun 2021 – Jul 2024*

- Developed a fraudster detection model using employee transactions, behavior, and demographics; flagged **80 cases**, with **13 confirmed policy violations** after audit review.
- Worked as a Full Stack Developer to build the Internal Audit website, enabling auditors to securely file critical documents and track audit operations in real time.
- Conducted **1 departmental audit** (Data Analytics) and **2 operational audits** (Server Operations, ML model reliability), providing recommendations to improve model performance, regulatory compliance, and operational efficiency
- Replaced 50+ Excel-based audit rules with a Python/Oracle system, reducing data manipulation time by **83 %**.

Software Architect Intern — Tus Solution, Ulaanbaatar, Mongolia *Jun 2020 – Sep 2020*

- Assisted the development team by designing ERD diagrams and C4 models for accounting and operations systems, following business requirements; delivered a complete accounting system based on my proposed architecture.

Data Science Intern — Khaan Insurance, Ulaanbaatar, Mongolia *Jun 2020 – Sep 2020*

- Built a Django-based system with a REST API to serve real-time Power BI dashboards, replacing manual PowerPoint reports and saving an estimated **450+ person-hours annually** by eliminating twice-weekly status meetings.

Project and Research

Volumetric Inpainting for LSM Images *May 2025 – Present*

- Built custom masked dataset pipeline; evaluated models with SSIM, PSNR, and LPIPS metrics.
- **Developing a 3D inpainting method** using diffusion models (PatchDDM + RePaint) on Light Sheet Microscopy data.

Apriori Algorithm and Markov Chain based Sequential and Dynamic Association Rules *Sep 2024 – Dec 2024*
GitHub Link

- Developed a hybrid association rule mining algorithm combining Apriori pruning and Markov Chains to capture sequential customer behavior; reduced dataset scans by **99% (from 8,489 to 43)** while preserving **66.7%** of top 3-item rules on a 38,765-row grocery dataset.

Network Analysis on Neurons of Drosophila Melanogaster in Feeding Motor Sequences *Jan 2025 – May 2025*
GitHub Link

- Analyzed neural circuitry of Drosophila melanogaster feeding initiation using the FlyWire connectome; applied network science techniques including **motif analysis**, **spectral clustering**, and **edge percolation** to uncover non-random recurrent patterns and neurotransmitter-specific structural modules in synaptic graphs.

A Comparative Study of Cellular Automata and Alternative Algorithms in Medical Imaging GitHub Link *Oct 2024 – Dec 2024*

- Developed a mammogram tumor detection pipeline improving sensitivity by **30%** over a baseline cellular automata method; integrated CLAHE, median filtering, and Canny edge detection to enhance robustness under 3% simulated noise.

Elite Scraper – Market Intelligence Data Extraction Tool GitHub Link *Jun 2021 – Aug 2021*

- Developed a web scraping solution to collect and structure market data on real estate, automobiles, and electronics for the non-banking sector.
- Automated competitor analysis and demand forecasting workflows using large-scale, real-time datasets.
- Reduced data collection time from **2–3 days to just 7 minutes (a 99% decrease)** for a target market segment, drastically increasing research efficiency.