

Haozhen Shen

[in LinkedIn](#) | [647-505-0656](tel:647-505-0656) | haozhenshen.com | haozhen.shen@outlook.com | [GitHub](#)

Skills

- **Development:** Python | C | C++ | Java | JavaScript | React | TypeScript | Node | Bootstrap | Tailwindcss | Express | Next | MongoDB | PostgreSQL | Kafka | Redis | Docker | CI/CD | Jenkins | Git | Bitbucket | Jira | Unit Testing | OOP | Distributed Systems
- **Machine Learning:** R | NumPy | Pandas | scikit-learn | Prophet | XGBoost | CatBoost | LightGBM | Pytorch | Pytorch Lightning | Tensorflow | Keras | Hugging Face | Transformers | VAE | EBM | W&B | Neptune | MLFlow | W&B | ONNX | Kedro | Airflow | Spark | Ray
- **Platform:** AWS | Databricks | Vercel

Experience

Machine Learning Scientist

BluWave-ai

Ottawa, ON, Canada

05/2022 - Current

- Lead Scientist for energy forecasting products.
- Built an end-to-end machine learning pipeline for time series forecasting of electricity load data using **ECS, Elasticache, Lambda, Ray, Kedro, Airflow, MLFlow, Tensorflow, Pytorch**, and **LightGBM** resulted in a 22% improvement in performance (MAE) and an 80% reduction in inference time compared to the previous method.
- Led the development of the ML system for load prediction products. From identifying system requirements and partner dependencies, and the entire life cycle of data engineering, model development, testing, and configuring metrics, alarms, and dashboards (**Grafana**).
- Took the initiative to refactor existing machine learning pipelines into modularized components leveraging Kedro accelerating model development and lowering the cost of maintenance.

Research Assistant

University of Toronto

Toronto, ON, Canada

02/2022 - 06/2022

- Applied deep learning methods to principal-agent mean-field games. Modeling the Renewable Energy Credit market using McKean-Vlasov Forward-Backward Stochastic Differential Equations. Solving the stochastic control problem using deep learning.

Data Mining Engineer, Intern

CRRC Academy

Beijing, China

04/2019 - 08/2019

- Joined the algorithm team dedicated to analyzing the vehicle's condition for the world's largest rolling stock manufacturing company
- Accelerated data preprocessing pipelines using MATLAB which increased preprocessing speed by 33%.
- Designed an LSTM-CNN classification algorithm to validate and identify potential vehicle failures.
- Developed a threshold analyzing algorithm, which helps distinguish valid data from noise caused by a sensor failure.

Projects

MarketSentinel (Full Stack Machine Learning, NLP)

- A stock market sentiment visualization website. Web component implemented using **Three.js, React, Typescript**, and **Next.js** and Deployed on **Vercel**. Backend logic is written in Vercel serverless functions. DistilRoberta from **Huggingface** and fine-tuned on financial news is used to perform text sentiment classification for news data. Inference is done through **Huggingface** Inference API. [GitHub Link](#)
- **Mini Redis:** Build a simplified version of Redis in C++ that handles multiple concurrent clients with Echo, Set, and Get commands.
- **Feedback Prize - NLP:** Fine-tuning Deberta models to assess the language proficiency of 8th-12th grade English Language Learners.
- **Adaptive Noise Score Network:** Designed a Score Based Generative Model, inspired by Adaptive MCMC techniques.
- **Energy-Based VAE:** Image generation by jointly training VAEs and Energy Based models (EBMs) through Contrastive divergence.

Education

Master of Science

University of Toronto (St. George)

Toronto, ON, Canada

09/2021 - 04/2022

- Statistics, Focus on Generative Modeling, Probabilistic Models, and Statistical Learning Theory.

Bachelor of Science

University of Toronto (St. George)

Toronto, ON, Canada

09/2016 - 04/2021

- Specialist, Computer Science, Machine Learning path.
- Specialist, Applied Mathematics, Probability, and Statistical Modeling path.

Mentorship

- UTFUN Tutor: Tutored CSC108 to undergraduate students responsible for teaching basic algorithms and data structures in Python.

Others

- **National second-level athlete** in the game of GO.
- **First Place:** Ranked first place in the three-dan promotion competition of the game of GO in, Shenzhen, China.
- **Professional fingerstyle guitarist.**