Arjan Waraich

647-473-1175 | waraicharjan97@gmail.com | https://www.linkedin.com/in/arjwaraich/ | https://github.com/RJN25

Education

University of Toronto Schools — Toronto, Canada
*Cumulative GPA: **4.0/4.0** (Math, Computer Science, Physics, Chemistry, English, History)* — *Sept. 2024 – June 2026*

The Knowledge Society — Toronto, Canada
Alum — *Sept. 2023 – June 2024*

Technical Skills

Languages: Python, C++, HTML/CSS, JavaScript
Frameworks: Flask, CrewAI, EdgeImpulse, Tkinter
Developer Tools: Google Colab, VS Code, Git, Figma, Linux, LaTeX
Libraries: pandas, Scikit-learn, NumPy, Matplotlib, Seaborn, Tensorflow, Pytorch

Certifications

Machine Learning by Andrew Ng - Stanford University (Online)
Deep Learning Using Tensorflow - IBM
Deep Learning Essential - IBM
Data Science Foundations - IBM

Conference & Research Papers

(CUCAI) ProphetJet — Predictive Maintenance for NASA Turbofan Engine RUL-Metrics March 2025

- Developed and feature-engineered a hybrid **LSTM-XGBoost-Random Forest** pipeline to forecast Remaining Useful Life **(RUL)** of turbofan engines using NASA's C-MAPSS dataset
- Integrated time-series deep learning with tree-based regression to model engine degradation, achieving **87.4%** RUL prediction accuracy (±2% std)
- Optimized feature selection via correlation matrix heatmaps and prioritized high-impact sensors to refine input vectors and boost model reliability
- Authored and presented a research paper on the model's development and results, successfully **accepted** into the 2025 **CUCAI** (Canadian Undergraduate Conference on AI) proceedings - as a **high-schooler**

Projects

Predictra — Industrial Manufacturing Intelligence AI Platform June 2025

- Developed a modular ML pipeline using **LSTM Autoencoders**, **Bayesian bootstrapping**, and **XGBoost** to estimate machine-specific Remaining Useful Life (RUL) without labeled failure data
- Built an anomaly detection engine that parses raw logs into structured sequences, detecting deviations using **reconstruction error thresholds**, and assigning per-timestep **risk scores [0–1]**
- Engineered a **generalizable framework** capable of adapting to **any industrial machine type** by learning patterns from **healthy data alone**, requiring no machine-specific tuning or failure labels
- Integrated industry-grade datasets (SCADA logs, IEEE, UCI, NASA CMAPSS), standardizing **multi-sensor time series** from **15+ machine types** with minimal pre-processing
- Translated anomaly curves into **RUL bands** (<5 days = urgent) for real-time health visualization via a secure, full-stack dashboard
- Back-end built with **Python, TensorFlow, Pandas, Joblib**; deployed models with JWT-authenticated endpoints and React-based UI for protected machine views
- **$5.5K Top AI-Startup Prize @ SpurHacks (2000+ hackers)** — Canada's largest business-focused hackathon, as high schoolers

ByteBite — ML-Triage Inventory Management System for Foodbanks February-March 2025

- Built a machine learning **triage**-like model from scratch using **Random Forest** regression with Pandas and Scikit-learn to prioritize food distribution and reduce spoilage, trained on **real-world** food bank inventory data
- Designed a scoring system to rank food items by urgency and nutritional value, improving distribution efficiency for **AVENUE ROAD Foodbank Toronto** and others in the GTA
- Engineered backend data pipelines using Pandas and Scikit-learn for standardization across **9** nutritional and logistic features (e.g., expiry, quantity, weekly demand)
- Built a responsive, full-stack web interface with user authentication (MongoDB-JSON UX Architecture), dynamic inventory tracking, and Nutritionix API integration for real-time nutrient data

AI Marketing & Datacenter Optimization Consultation (GOOG + MSFT) Dec. 2023 & Apr. 2024

- Contributed to **two** independent AI-focused innovation initiatives with **Microsoft** and **Google** through **TKS-led collaborations**
- With **senior Microsoft engineers**, researched solutions and created conceptual framework to increase efficiency within AIOps, and datacenter energy/cooling costs optimization
- Formulated an extensive solution regarding integration of **underwater datacenters** leveraging seawater cooling properties, heat dissipation & geothermic energy, and focused on the AI budget for growth & environmental sustainability
- With **Google product managers**, consulted and collaborated to increase efficiency within marketing & product management ecosystem with AI-led solutions
- Created **3** primary recommendations to ensure communication between GTM (go-to-market) teams; genAI product portfolio updates, 3rd party service AdScraping with **NLP** filtration, automated universal product-update terminal connected to engineers' deployments for clarity & communication
- Compiled findings into detailed slide-decks pitched to engineers, product managers, and stakeholders for review and feedback.