Harsh Upadhyay

upadhh1@mcmaster.ca LinkedIn GitHub Portfolio

Education

McMaster University | Honours Computer Science Co-op + Minor in Statistics — Sep 2022 – Apr 2027

- **Courses:** Data Struct & Algorithms, Operating Systems, Databases, Compilers, Machine Learning, NLP, Cryptography

Research

OCAD University | Research Assistant — April 2026 – December 2026

- Evaluated **subword tokenization** strategies (BPE, unigram LM) for polysynthetic Indigenous languages, benchmarking against **morphologically-informed, phoneme-aware** segmentation to preserve semantic integrity in agglutinative structures.
- Engineered a real-time **vocal collaboration pipeline** using **YIN/RMVPE pitch tracking**, rhythm analysis, and adaptive **harmony generation**, with DSP synthesis and an isolated **RVC neural voice-conversion** service for low-latency performance.

Experience

Government of Ontario - MPBSDP | IT Systems Assistant — May 2026 – August 2026

- Supported **OAC DV** rollout for **200+ users** through technical workshops, workbook development, and **AI feature evaluation**.
- Optimized **OAC subject areas** and **workbooks** to improve performance and streamline ministry-wide self-service analytics.
- Performed **SQL analysis** and developed **Python validation scripts** for **Detailed Schedules of Payments (DSP)** reports, automating checks across multi-ministry extracts and reducing manual validation for **Public Accounts**.
- Developed an interactive analytics dashboard for the ministry's **APEX-based Procurement and Contract Tracking (PCT)** platform, transforming data into **actionable insights** for streamlined monitoring, analysis, and reporting.

Government of Ontario - MPBSDP | Business Analyst Co-op — May 2025 – December 2025

- Enhanced Match and Merge workflows, increasing automated throughput from **30% to 70%** by leading **Agile grooming & refinement** sessions, translating requirements into technical tasks, and delivering implementation-ready **Jira** user stories.
- Performed end-to-end **system and integration validation**, exercising **REST APIs** and third-party integrations via **Postman**, simulating workflows to detect defects, validate edge cases, and enforce business rules.
- Conducted **data validation** and **migration feasibility analysis** via **Python-SQL analytics**, comparing legacy postal records against external datasets and **GeoJSON boundaries** to identify reliable replacement opportunities.

Projects

Archer-Robot Vocal Collaboration System — *Python, FastAPI, PyTorch, RVC, RMVPE, YIN*

- Engineered a **real-time vocal AI pipeline** combining **YIN/RMVPE pitch tracking**, rhythm analysis, and adaptive harmony.
- Architected a **DSP synthesis engine** using sinusoidal, wavetable, and formant synthesis for multi-voice accompaniment.
- Integrated **PyTorch/RVC voice conversion** via an HTTP sidecar to prevent heavy inference blocking real-time audio.
- Deployed **FastAPI AI microservices** driving Whisper/Vosk transcription, RoBERTa sentiment analysis, and Cree translation.

ActiveTrack — *Python, Pandas, NumPy, Scikit-Learn, Matplotlib*

- Built a **machine learning model** with **98.9%** accuracy for identifying barbell exercise types from multi-axis **IMU sensor** data.
- Engineered **25+ features** from **accelerometer and gyroscope** time-series using **FFT**, statistical analysis, and clustering.
- Optimized model performance with exhaustive **grid search** across **Random Forest**, **Decision Tree**, and **SVM** classifiers.
- Developed automated rep-counting using **LowPassFilter** and peak detection algorithms, achieving **95%+** counting accuracy.

Arena — Next.js, TypeScript, Tailwind CSS, Go, WebSockets, OpenAI, Anthropic, Groq

- Engineered a real-time **Go game server** hosting multi-model **AI competitions** streamed via **WebSocket rooms**.
- Designed concurrent **multi-game engines** for turn-based and simultaneous AI matches via isolated **goroutine loops**.
- Developed an LLM parser that strips **reasoning model think blocks** via automated **3x fallback retry loops**.
- Built a **Next.js 15 dashboard** visualizing live match states alongside an automated, dynamic **ELO leaderboard**.

Skills

- **Languages:** Python, JavaScript, TypeScript, C, C++, C#, Java, Scala, SQL, NoSQL, Go, Bash / Shell, Haskell, MATLAB, R
- **Frameworks & Libraries:** Next.js, Angular, Node.js, .NET, Django, Spring Boot, FastAPI, TensorFlow, PyTorch, OpenCV, Pandas
- **Tools & Infrastructure:** Git, GitHub, Linux, Docker, Kubernetes, AWS, Cloudflare, GitHub Actions, Jenkins, CI/CD, REST APIs, WebSockets, GraphQL, Postman, Swagger