Kapil Mirchandani

+1(343)558-8263

Canada

kapilm2602@gmail.com | linkedin.com/in/kapilm26 | github.com/KapilM26 | kapilm26.github.io

SUMMARY

- Machine Learning Engineer with 2+ years of industry experience, specializing in Natural Language Processing, MLOps, and machine learning model development.
- Expertise in building end-to-end machine learning pipelines, from data collection and preprocessing to model research, development, deployment, and monitoring.
- Strong experience in developing and optimizing machine learning systems, with notable contributions to multiple applied ML projects and two research publications in international conferences.
- Skilled in building scalable, robust backend services and RESTful APIs (monoliths and microservices), as well as data analysis and automation.

EDUCATION

**Master of Engineering, Electrical and Computer Engineering
(Concentration in Applied Artificial Intelligence)**   September 2023 - April 2025 (Expected)
University of Ottawa, ON

Bachelor of Engineering, Electronics and Telecommunication Engineering   August 2017 - May 2021
Pune Institute of Computer Technology, Pune, India

RELEVANT WORK EXPERIENCE

Data Scientist Intern – Quantum AI Research   January 2025 - Present
Natural Resources Canada, Ottawa, Ontario, Canada

- Building classical neural networks with parameter constraints to classify image data for comparative analysis with quantum neural networks.
- Conducting detailed comparisons to explore the interplay and advantages of classical and quantum neural networks.

Software Developer – SYNC HMI Intern   May 2024 – December 2024
Ford Motor Company, Ottawa, Ontario, Canada

- Implemented an automated bug ticket triaging system using word embeddings and a machine learning classifier, achieving 80% accuracy in routing tickets based on titles and descriptions, saving 15 hours of total developer time per week.
- Analyzed over 2 billion log lines, extracting insights such as high-memory processes and low-impact crashes, leading to fixes that reduced crashes by 20% and memory usage by 10%.
- Automated configuration validations and detection of common issues, integrating these solutions with Jira, which reduced the number of tickets requiring manual triaging by 50%.
- Parsed over 1 billion log lines to build a classification dataset of over 2000 data points for machine learning, extracting key features based on developers' manual analysis processes.
- Applied machine learning classifiers to the dataset, effectively handling imbalanced data using ADASYN, and achieved an F1 score of 82%.

Machine Learning Engineer   March 2022 - July 2023
Avoma, Palo Alto, CA, USA (Remote)

- Fine-tuned transformer-based embeddings, using bi-encoder and cross-encoder architectures to build a semantic similarity model which achieved an accuracy of 88% on meeting transcript data.
- Leveraged LLMs, along with RAG, particularly GPT-3.5 and GPT-4, for various summarization, rephrasing, question-answering, and sentence classification tasks, to deploy 6 new features, including transcript

summarization and automated notetaking.

- Implemented image processing and OCR techniques to identify the current speaker from virtual meeting interfaces (Zoom, Google Meet, MS Teams), improving the accuracy of the current speaker recognition system from 67% to 88%.
- Developed a feature using clustering on over 100,000 meetings to rank top FAQs and reveal key themes; its deployment increased average CSAT score from 7 to 8.
- Involved in system and data platform design for deployment of these models, using Django, PostgreSQL and AWS EC2 and ECS to build end-to-end Machine Learning inference and training pipelines for real-time sentence classification and transcription.
- Optimized inference speed of Tensorflow models, using the gRPC protocol and AVX acceleration achieving a 30% decrease in latency.
- Set-up and maintained automated CI/CD pipelines, using Tensorflow serving, Docker and Github actions, automating and reducing deployment times from approximately 10 minutes to 2 minutes.

Software Engineer I (Machine Learning)   July 2021 - February 2022
Helpshift, Pune, Maharashtra, India

- Developed advanced machine learning solutions to automate end-user issue resolution in customer service.
- Conducted research, scoping, benchmarking, and deployment of cutting-edge NLP algorithms, including user message classification and language detection.
- Engineered scalable microservice inference pipelines for machine learning models using Django, Flask, and FastAPI, which supported over 1 million requests per day.
- Diagnosed and resolved critical bugs, optimizing underperforming models for clients, by using oversampling or adjusting class weights on imbalanced datasets, boosting classification accuracy from 70% to 90%.

TECHNICAL SKILLS

- Languages: Python, C/C++, MATLAB, Java, Clojure, SQL, Javascript
- Frameworks: Tensorflow, Pytorch, Scikit-Learn, Django, Flask, FastAPI, Selenium
- Libraries: Pandas, NumPy, Matplotlib, Seaborn, OpenCV
- Developer Tools: Git, Docker
- Databases: MySQL, PostgreSQL, MongoDB
- Cloud technologies: AWS (S3, EC2, Lambda, ECS/ECR), Tensorflow Serving
- Operating Systems: Linux, Windows, MacOS

APPLIED PROJECTS

- Worked as part of a five-member team, using Deep Learning to detect the Cell-Free Layer in blood vessels, under the guidance of Prof. Ali Abbas and Prof. Marianne Fenech at the University of Ottawa. Tech stack – MATLAB, Tensorflow
- As a part of uOttaHack 6, developed an application for the Ford Motor Company to optimize routes for Electric Vehicles, providing required charging stations along the route and minimizing the time and distance required for the journey. My team was awarded first place at the hackathon. Tech stack – React.js, Django

PUBLICATIONS

- **DPSRGAN: Dilation Patch Super-Resolution Generative Adversarial Networks**, 6th International Conference for Convergence in Technology (I2CT), 2021
- **Big Data Analytics for Sustainable Cities: Pune Tree Census Data Exploratory Analysis**, 11th International Conference for Computing, Communication and Networking Technologies (ICCCNT), 2020