



SHYAM PRASATH

📍 Toronto, Ontario, Canada

🌐 www.linkedin.com/in/shyamprasath

✉ shshyam96@gmail.com

🌐 www.github.com/shyam1326

☎ (647) 673-0387

🌐 shyam1326.github.io

SUMMARY

Experienced Machine Learning professional with a **master's degree in AI&DataScience** and over 5+ years of expertise in building ML systems in production, performing model optimization, and implementing robust product measures. **Certified TensorFlow Developer and Azure Data scientist** with a strong background in software Engineering and Large-Language Model. Enthusiastic about open-source contributions, actively participates in Kaggle competitions. Well-versed in LLM & RAG frameworks.

TECHNICAL SKILLS

Technologies: Transformers, LLM, RAG, ML-Ops, Deep Learning, CV, NLP, MLFlow, CNN.

Programming and Scripting: Python, shell-scripting, Bash, SQL, HTML.

Technical & OS: Data Structures and Algorithms, OOPs, Windows, Linux (RHEL, Ubuntu, Debian).

Backend Framework: TensorFlow, Pytorch, keras, Flask, Django, Streamlit, Pyspark, Fast-API.

Devops Orchestration: Docker, Airflow, Git, GitLab CI/CD, Azure Devops.

Analytical & Cloud: PowerBI, Tableau, Advance Excel, AWS, Azure, Heroku, Google Cloud.

Databases: Vector database, Redis, MySQL, Postgres, MongoDB.

WORK EXPERIENCE

Lead Machine Learning Software Engineer

Detect Technologies Pvt Ltd

November 2021 - Present, Toronto, ON, Canada

- Lead the Machine Learning AI software Engineering team of 6 members. And Responsible for the product design and architecture of ML system.
- Executed **ML-OPS** pipeline for the post-production model monitoring, model artifacts, logs and retraining when there is a data drift. Integrated CD pipeline to deploy the updated model.
- Developed and **launched the pilot version** of the research project (Optimize Hydrogen wastage for plant efficiency) convert into production ready product that drives **\$1M revenue every year**.
- Build complex FAQ **Retrieval Augmented Generation (RAG)** system with opensource **Llama-2 model** and vector database using **Llama-index** and **Lang chain Framework**.
- Developed Incident management reporting tool using **OpenAI GPT-4** preview model to estimate the Incident type, Injuries and Environmental etc... based on the user incident description.
- Design and implemented scalable automated application for managing image annotation for model building and **achieve 28% throughput**.

Machine Learning Engineer

Aleafia Health

January 2024 – April 2024, Toronto, ON, Canada

- Implemented **Optical-character recognition** for invoice processing using Azure API. Integrated **RAG based LLM system** to populate the insights from the invoices.
- Developed **Django web framework** for Employee Expense Reports to Automate vendor invoice processing with power automate and seamless integration with existing SQL database.
- Preformed Enhancing PowerBI reporting through DAX and Dynamic dashboard creation under the guidance of **Vice president of Finance**.
- Conducted **training and awareness sessions** to educate Finance team on python and PowerBI.

Data Scientist

Loblaw Companies Limited

May 2023 – August 2023, Brampton, ON, Canada

- Spearheaded the implementation of Bert Transformer model for the weekly-flyers Analysis and

achieved **significant efficiency improvement** by optimizing the weekly flyers comparison.

- Implementation of PowerBI reporting tool to automate data analysis, resulting in **40% reduction in manual reporting**. Build End-End data pipeline to streamline the data from **GCP BigQuery**.
- Developed production reports for promotional-optimization model, comparing total market tonnage to identify opportunities to increase market-share by **15% and drive \$1.5M in revenue**.

Data Analyst

Jean Martin Systems Ltd

December 2019 - October 2021, Tamil Nadu, India

- Researched and developed the codebase of product development prototyping using AI technology (Deep Learning). Analyze dataset to provide insights to experts from various domain.
- Automated client project to reduce human resource effort & allocation, **achieved 40 more in profit**.
- Design and build automated Data Visualization platform and data scrapping tool (Web scrapping).

Graduate Engineer Trainee

Bharat Petroleum Corporation Ltd

November 2018 - October 2019, Bangalore, India

- Assisted sales team to **increase sales** by providing reports and dynamic charts and analyse the data.
- Gathering, Manipulation & Interpretation of data using Excel and Tableau for reporting Purpose.

EDUCATION

Ontario college Post-Graduation in AI and Data Science

Loyalist College • Toronto, ON, Canada • 2022 – 2024 • CGPA 3.74

Bachelor of Technology in Engineering, Mechanical Engineering

SRM College of Engineering • Chennai, Tamil Nadu, India • 2014 – 2018 • CGPA 3.4

PROJECTS

Incident Management Application using RAG and LLM

Detect Technologies Ltd, January 2023

- Designed and deployed a LLM system to estimate Incident type, action cause, object cause, etc... based on the user provided Incident Description. And integrated the FAQ based RAG system.
- Integrated with 9+ external plugins for data source and additional functionality of “Pandas Querying” using LLAMA-Index and Lang chain Framework.
- **Technologies: Python, LLM, RAG, OpenAI, Lang chain, Llama-Index, Vector DB (Pinecone).**

Diesel Hydro-treater Optimization using Multiclass Regression

Detect Technologies Ltd, December 2021

- Build a multiclass regression model to optimize the usage of temperature and hydrogen to refine the crude oil. Implemented wide & Deep custom Neural Network model, setup streaming batch data flow.
- Implemented E2E MLFlow pipeline to store the artifacts and identify the data drift to retrain model.
- **Technologies: Python, Flask, MS-SQL, SQLite, Docker, MLFlow, ML-Ops, Neural Networks.**

AutoPilotML

Open-Source Contributions, September 2023

- Build an open-source Automated ML library, simplifies the end-end ML pipeline for production. It is available for installation via pip at <https://pypi.org/project/autopilotml/>.
- **Technologies: Python, Machine Learning, Clustering, ML-Flow, statistical modelling. XGboost.**

CERTIFICATIONS

TensorFlow Developer Certification from Google	2024
Azure Data Scientist Associate from Microsoft	2024
Azure AI Fundamental from Microsoft	2019
Advance Deep Learning, Computer Vision, NLP from Ineuron.ai	2020