

Deepratna Awale

St. John's, NL | +1 709-341-7268 | awale.deep@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

Generative AI and Machine Learning Engineer with 2 years of industry experience and an **IBM Data Science Professional Specialization**, one of the **top certifications in data science**. Strong track record of deploying **full-stack AI applications**, **LLM** and **Diffusion model training and finetuning**, **data engineering**, **ETL**, and **Data Analytics**.

Key Skills: Generative AI, Machine Learning, Data Science, Python, TensorFlow, PyTorch, AWS, ETL, full stack.

WORK EXPERIENCE

AI Trainer: Coding Expert - T3 Outlier [Remote, Freelance] Oct 2024 – Current

- Correct/Improve **LLM Code generation output** by giving a better solution.
- Fixed the **Merge Sort Space complexity** for our custom model which **every** LLM gets wrong.
- Supported code improvement for **Python, JavaScript, Java**.

Generative AI Engineer, Tapestry Video AI, Mountain View, CA, USA Oct 2024 – Dec 2024

- **Created, maintained and optimized ComfyUI AI** workflows by **40%** for **inpainting** on TikTok style videos.
- **Trained and finetuned** custom opensource **diffusion models** on UGC dataset to produce TikTok style content.
- **Built LLM agents** using **LangChain** for analyzing Shopify product pages, and relevant TikTok tags.

Machine Learning Engineer, Axiom Softech Pvt. Ltd., Nagur, MH, India Oct 2021 – Sep 2022

- **Optimized Data Pipelines**, reducing **ETL time** by **35%** using **PySpark** handling data over **4M+** records.
- **Architected PyTorch and TensorFlow NN models**, improving **customer engagement** by **20%**.
- Deployed **end to end AI solutions** on **AWS** using **Firehose, Kinesis Data Stream, Athena, Lamda, S3**, and **Sagemaker**, achieving **99.0%** uptime and cutting costs by **25%**.

Team Lead, IIT Bombay, Mumbai, MH, India Aug 2021 – Sep 2021

- **Led 5 C++ developers** and **5 Python developers**, delivering features in **record** times.
- **Developed Udaan**, a **C++ application** with **QT Framework**, to aid **OCR** output correction for **Indian scripts**.
- **Spearheaded feature introduction** to reduce **manual corrections** by up to **30%** across files.

Junior Data Analyst, Axiom Softech Pvt. Ltd., Nagur, MH, India Jan 2021 – Jul 2021

- Employed **Python** to **create data pipelines**, **eliminate manual cleanup**, achieving **98% data accuracy**.
- Engineered **forecast models** for augmenting **financial decisions** with data (**Scikit-learn, Statsmodels**).
- **Presented interactive dashboards** to stakeholders for decision making, powered by **Tableau** and **PowerBI**.

SKILLS

- **Programming Languages:** Python, Java, C++, JavaScript, SQL.
- **Domains:** Generative AI (LLMs & Diffusion), Machine Learning, Data Science.
- **Frameworks:** TensorFlow, PyTorch, Keras, QT Framework, Flask, Django, FastAPI, PySpark, NextJS, AuthJS.
- **Libraries:** Pandas, NumPy, Matplotlib, LangChain, Sci-kit Learn, OpenCV, Pillow, Selenium, Beautiful Soup, Gradio.
- **Platforms:** Linux, Windows, MacOS, AWS.
- **Tools:** Tableau, Power BI, Docker, Jenkins, Git, Jira, Stable Diffusion.

EDUCATION

MASc, Computer Engineering, Memorial University of Newfoundland, Canada 2024

BEng, Information Technology, RGCEER, India 2021

PROJECTS

Auto Express [Gen AI, Active Users, Live Site]



Jan 2024 - May 2024

- Launched an **open-source application** that **renders 28 facial expressions** on any face using **Stable Diffusion**.
- Created **inpainting** workflows to perform **localized feature changes** guided by **finetuned YOLOv8**.
- Result is a **one-click solution** for **chatbot emotion display**, **optimizing** the process by **91.67%**.

SD-Parsers [Gen AI, Live Site]



Jan 2024 - May 2024

- Launched an **open-source Typescript node package** to extract **AI generation metadata** from images.
- Deployed a live site for seeing the package and api in action.
- The parser is now **available online as api, and as a package**, and supports **multiple generation source**.

3T Chat [Gen AI, Live Site]



Jul 2025 - Jul 2025

- Launched a serverless **LLM chat service** using **Grok AI** with **my personality** in under **2 hours** on **Vercel**!
- Implemented **file upload functionality** using blob storage, **managed PostgreSQL** instance for user Auth.
- Added periodic pulls for **GitHub repo api** and Resume to keep the chatbot updated on my latest contributions.

Fraud Detection [ML]



May 2025 - June 2025

- Built a **PySpark ETL pipeline** to process a **6M+** transactions dataset, **cleaned**, and **feature engineered** new fields.
- Designed and **deployed** a **Neo4j graph schema** to enable **multi-hop transaction analysis**, and train classifier.
- Developed and evaluated **fraud classifiers** using **scikit-learn** on **combined tabular and graph features**, reaching **89% precision and 84% recall** on synthetic data.

Ocean Wave Height Estimation [Deep Learning, Neural Networks, CV]



Jan 2024 - April 2024

- Developed a custom **parser** for **Polar files** from the **Wamos II radar system** using **NumPy**, and **matplotlib**.
- Enhanced **accessibility** by **encoding data** into **PNG** format and **embedding radar data** using **Pillow**.
- Designed a **Hybrid Neural Network** consisting of **CNN and MLP** in **sci-kit learn** to estimate wave height.

3D Surface Reconstruction [CV]



Jan 2023 - April 2023

- Employed **COLMAP's Multi-View Stereo and Multi-View Fusion pipelines** for **3D reconstruction** from images.
- Developed **Python** scripts using **OpenCV**, **NumPy**, and **SciPy** to preprocess the Brutus rig's data.
- Performed frame extraction, **distortion removal**, **camera intrinsic information extraction** and **embedding to DB**.

Theoretical Answer Evaluation System [NLP]



Dec 2020 - May 2021

- TAES is an **NLP** system that evaluates textual answers using **tokenization**, **DAN**, **Rake NLTK**, and **TensorFlow**.
- Implemented features for **plagiarism detection**, **grammar-based penalties**, and **automated score allocation**.
- Achieved alignment with human-given scores 8/10 times, scoring **78%** on the semantic similarity **STS Benchmark**.

CERTIFICATIONS

IBM Data Science Professional Specialization

2019

IIT Madras Programming, DSA Using Python

2019