ADITYA RATAN JANNALI

+19167767065 jannali.a@northeastern.edu https://www.linkedin.com/in/aditya-ratan-j/ https://github.com/rjaditya-2702

EDUCATION

MS in Artificial Intelligence, September 2023 – May 2025
Northeastern University (GPA: 4.00), Boston, MA
Courses: Machine Learning, Foundations of Artificial Intelligence, Programming Design Paradigms, Algorithms, NLP, Masters Project (Kolmogorov Arnold Networks & Reinforcement Learning)

Bachelor of Technology in Electronics and Communications Engineering July 2017 – June 2021
Vellore Institute of Technology (GPA: 9.11) Chennai, India
Courses: Statistics, Calculus, Linear Algebra, Probability Theory, Data Structures, Machine Learning, Digital Image Processing.

EXPERIENCE

Institute of Experiential AI Jan 2025 – present
Data Scientist Co-op, Generative AI *Portland, ME*

- Implement a custom RAG system framework for internal team to answer user queries.
- Implementing and Testing a Semantic Search system to retrieve relevant wiki documents for user search queries. This search system is aimed to reduce the Developers involvement into answering user queries on consumer forums
- Part of ongoing project for North Cross Group to implement an AI took those analyses clauses from legal contracts and performs generative and classification tasks. This project reduces the manual effort required by the clients to manually analyze hundreds of contract documents to generate the structured results.

Amazon August 2021 – August 2023
Application Engineer III Chennai, India

- Provided front line support to the AEE product and MENA Data Engineering and Science team, owning up automation projects to reduce incoming tickets by 35%, saving 20hrs/month of engineering time. Designed and built a common analysis dashboard featuring easy onboarding, customizable filters, and insights, enabling interactive views for engineers and leadership.
- Took full ownership of designing and building end-to-end plugins and ETLM pipelines utilizing Python, PostgreSQL, and AWS to generate insightful reports of pipeline health metrics, usage patterns, and cost utilization of Redshift clusters.
- Setup interactive dashboards for these plugins, providing leadership with real-time access to actionable insights during weekly and monthly business reviews, enabling data-driven decisions on load assessment and resource optimization.

Amazon January 2021 – July 2021
Software Support Engineer II Intern Chennai, India

- Designed, developed, and deployed a utility that collects service level raw data and distributes aggregated metrics to relevant service owners and stakeholders. This utility streamlines data management, enhances information accessibility, and facilitates decision making for key personnel.
- Assumed the role of release engineer, taking ownership of production deployment for multiple pipelines while ensuring their smooth operation.

Antpod April 2020 – December 2020
System Development Engineer Intern Chennai, India

- Involved in Research and Development of a proof of concept for an unmanned vehicle in ‘Land Stress Identification and Remote Sensing’.
- Prototyped an algorithm using Deep Fully Connected Convolution Network using Keras and TensorFlow to segment images from the dataset and perform classification to identify plant disease.

SKILLS

Languages:	Python, Java, C++, SQL, HTML, CSS, JS, shell.
Libraries:	OpenCV, sklearn, pytorch, TensorFlow, gym, transformers, Numpy, Pandas.
Frameworks:	git, huggingface, AWS – Redshift, IAM, S3, Secret Manager, IAM, EC2, ECS, VPC, KMS Sagemaker, Secret Manager, QuickSight, Athena, DynamoDB.

CERTIFICATIONS

- (Coursera) Machine Learning,
- (Coursera) Deep learning using TF – CNN and NLP,
- (Coursera) Natural Language Processing – Classification And Vector Spaces, Probabilistic Models, Sequential Models.
- (Coursera) Digital Image Processing,
-

PROJECTS

1. **Research Project: Evaluating KANs for Reinforcement Learning Applications** Sept 2024 - present
   Implemented a DDQN using KANs to evaluate performance on Atari game environments in OpenAI Gymnasium. An ongoing study of comparative analysis between KAN-based and traditional MLP-based reinforcement learning architectures and training algorithms. Investigating KANs' potential as an alternative to Multilayer Perceptron models in Deep RL frameworks.
2. **Language Model Interpretability** [GitHub] Oct 2024 – Dec 2024
   A study where the attention head(s) of a finetuned GPT-2 small SequenceClassification transformer are masked completely to understand which region is responsible for driving the prediction of a certain class. The project also expands into replacing important tokens in the document before passing it to the model's layers.
3. **Transformer Encoder** Oct 2024 – Oct 2024
   Built a Transformer LM Encoder block from scratch that uses Multi-Head Attention. The encoder is built and trained for sequence level classification task (letter counting). The trained encoder achieved a perplexity of 6.62.
4. **RL Tic-tac-toe** [GitHub] May 2024 – June 2024
   Implemented Q-Learning to train two agents with different reward schemes to play against each other. Each valid board configuration is a state in this environment which is modelled as a 9-digit ternary sequence eliminating the need to precompute the observation space.
5. **LLM generated vs Human text classification** [GitHub] April 2024 – April 2024
   Finetuned and evaluated two transformer models, ALBERT and DeBERTa-XS on DAIGT Proper Train dataset to detect LLM generated essays. Achieved an accuracy of 96.57% for ALBERT, and 99.43% for DeBERTa.
6. **Image Processor** [GitHub] November 2023 – December 2023
   Designed and MVC, developed and wrote unit and integration tests for an image processing application that performs various image manipulations on custom images. The controller uses command design pattern to listen and process user interaction, and JSwing for the view.
7. **S.H.A.P.E.R** [GitHub] November 2023 – December 2024
   Collaborated in developing the collision handler and environment setup of an AI controlled hand that deflects incoming object from hitting the target. This project makes use of genetic algorithm to train a neural network that generates the movement of the arm at each time step.