KASHISH GOEL

(929) 786-8039 kg3044@columbia.edu linkedin.com/in/kashish-goel-930 github.com/kashish-goel

EDUCATION

Columbia University New York, NY
Master of Science in Computer Science, Machine Learning, GPA: 3.85 Dec 2023
Coursework - Applied Machine Learning, Cloud Computing & Big Data, Databases, AI, Natural Language Processing (TA)

Banasthali University Rajasthan, IN
Bachelor of Technology in Computer Science and Engineering, GPA: 9.15
Publication - *Hybrid Quantum Algorithm for 1-D Array: Integer Sort Using Quantum Gates in O(N logN)* time complexity

SKILLS AND TECHNOLOGIES

Programming	Python, C++, C, Java, Qiskit, SQL, BigQuery, MATLAB, Flask, React, REST, HTML, CSS
Packages	NumPy, Pandas, Scikit-Learn, TensorFlow, Keras, NLTK, Matplotlib, PyTorch, Jupyter
Other	AWS, Google Cloud (GCP), LaTeX, Pennylane, D3JS, PostgreSQL, MySQL, OpenCV, Kubernetes, Docker

WORK EXPERIENCE

RadicalX New York, NY
Machine Learning/AI Intern – Generative AI Jun 2023 – Sep 2023
- Developing ML models for pattern and anomaly detection focusing on plagiarism detection and code style analysis for the AI manager of their EdTech learning platform
- Exploring MinHash and Rabin-Karp algorithms, LSTM and GPT models for detecting code similarity using Hugging Face and JavaScript source code dataset

Columbia University New York, NY
Research Assistant – Large Language Modeling Jun 2023 – Present
- Fine-tuning GPT-2 model using LoRa to build a question-answering task model pipeline for Llama using webnlg dataset

Research Assistant – Quantum Machine Learning Feb 2023 - Present
- Building ML pipelines to predict neuro-immuno disorder biomarkers using 700+ T1 image scans. Generated over 1k features with radiomics and implementing Autoencoder networks for anomaly detection and condition monitoring

Nanyang Technological University Singapore, SG
Natural Language Processing Intern Aug 2021 – Feb 2022
- Utilized 20k+ tweets and SenticNet data to generate pipeline for conducting polarity analysis for personality detection
- Trained BERT and RNN classifier models on GPU for sentiment analysis and obtained an accuracy of approx 71.9%
- Set up a server to collect twitter tweets and comment data using REST API and populated it in SenticNet database

Indian Institute of Technology Bombay Mumbai, IN
Software Development Engineer Intern Jul 2021 – Nov 2021
- Implemented SVM and KNN model with 89% accuracy to classify large and noisy waste dataset in recyclable categories
- Designed the schema and engineered backend database hosted on PostgreSQL. Created triggers for data ingestion and collection, cutting down manual intervention significantly by half
- Conceptualized and hosted a React dashboard to visualize waste collection and integrated UI components for user query reduction, being used by 20k+ users registered under Municipal corporation

Artificial Brain Delaware, DE
Software Development Intern - Artificial Intelligence and Quantum Computing Jan 2022 – Jul 2022
- Contributed to the development of python-based quantum library, enabling writing cross-platform compatible code
- Developed a Hybrid Quantum CNN model for image dataset and optimized the circuit to an accuracy of 65.7%
- Increased the computational speed for Snake AI game by 19% using a novel Quantum Reinforcement Learning model
- Developed and published a quantum wheel android application (5k+ downloads) providing truly random event selection

PROJECTS

Dining Concierge Assistant Feb 2023 – Apr 2023
- Built a serverless, microservice-driven web app, hosting a dining concierge chatbot by scraping 5000 restaurants data from Yelp API and store it in DynamoDB database. Utilized AWS S3, API Gateway, Lambda functions to handle responses
- Containerized and deployed the application on local Kubernetes cluster using Docker, implanting scalable features, replication controllers, health monitoring, rolling updates and alerts

Self-Driving Car Control ***Link*** *Jan 2023 – Apr 2023*
- Trained a self-driving car using CARLA simulator using DQN based approach for reinforcement learning algorithms
- Improved the car control and safe navigation with Imitation learning approach and implementing SARSA with 5% less training loss than baseline