KASHISH GOEL

(929) 786-8039 kg3044@columbia.edu linkedin.com/in/kashish-goel-930 https://github.com/kashishgoel9

EDUCATION

Columbia University New York, NY
Master of Science in Computer Science, Machine Learning, GPA: 3.80/4.0 Dec 2023
Coursework: Applied Machine Learning (TA), Cloud Computing & Big Data, Databases, AI, NLP (TA), Data-driven Design
Thesis: **Multimodal** Story Generation from multiple Images: Context through Captions

Banasthali University Rajasthan, IN
Bachelor of Technology in Computer Science and Engineering, GPA: 9.15/10
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, MATLAB, HTML, CSS, Hadoop, CI/CD, Computer Vision, Deep Learning, Testing
Packages: Numpy, Pandas, Scikit-Learn, TensorFlow, Keras, NLTK, Matplotlib, PyTorch, PySpark, React, REST
Other: AWS, Google Cloud (GCP), LaTeX, Pennylane, D3JS, PostgreSQL, MySQL, Jupyter, Kubernetes, Docker, Analytics

WORK EXPERIENCE

RadicalX - Roblox New York, NY
Machine Learning/AI Engineer, Intern Jun 2023 – Sep 2023

- Developed ML models for **anomaly detection** focusing on code **style analysis** and similarity with an **accuracy of 75%**
- Engineered AI manager Chatbot for their EdTech learning **platform infrastructure**, making the **chat personalized** by modelling and fine-tuning various LLM models. Benchmarked decoder-only vs **encoder-decoder** models.

Columbia University New York, NY
Graduate Research Assistant, Prof. Ansaf and Dr. Korhan Feb 2023 – Dec 2023

- **Fine-tuning** GPT-2 model using LoRa to build a question-answering task model pipeline vs Llama using webnlg dataset
- Built ML pipelines to **predict biomarkers** for Multiple Sclerosis, using 700+ T1 and MRI image scans. Generated over **1k features** with radiomics and implemented **Autoencoder networks** for anomaly detection and condition monitoring.

SenticNet - NTU Singapore, SG
Natural Language Processing Engineer Aug 2021 – Feb 2022

- **Set up the server** to collect tweets and social media posts data and used **REST API** to populate the SenticNet database
- Implemented **end-to-end BERT and RNN** classifiers for **sentiment analysis**, obtaining an **accuracy of 89.2% and 71.9%**
- Utilized 20k+ tweets to generate **knowledge graph** and conducted **polarity analysis** for personality detection

Indian Institute of Technology Bombay Mumbai, IN
Data Scientist, Software Developemnt Jul 2021 – Nov 2021

- Implemented an **ensemble model** of SVM, KNN and Random Forest with 93.3% accuracy to classify **large and noisy real-time waste dataset** in recyclable categories and integrated it with analytics dashboard using JSON query
- **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, integrating UI components for user query reduction, being used by 9k+ users registered under Municipal corporation

Artificial Brain Delaware, DE
Software Engineer - Artificial Intelligence & Quantum Computing Jan 2022 – Jul 2022

- Contributed to the development of **python-based quantum library,** enabling writing **cross-functional compatible code**
- Developed a Hybrid Quantum-CNN model for geo image dataset and **optimized** the circuit to get an accuracy of 65.7%
- Increased **computational speed** of classical Snake-AI game by 19% using a novel Quantum Reinforcement Learning model
- Published a quantum android application (5k+ AppStore users) providing truly random event selection, utilizing QRNG

PROJECTS

Harmonize - Music Dating Application * Feb 2023 – Apr 2023

- Led a team to build serverless, microservice-driven web application for university students by collecting 500+ user profile data stored in **DynamoDB**. Utilized **AWS** services like S3, API Gateway, Lambda, OpenSearch, SQS to handle responses
- Implemented **BERT model and cosine similarity score** to generate keywords for candidate matching algorithm
- Containerized and deployed the application on local Kubernetes cluster using **Docker**, implanting auto-scalable features, replication controllers, health monitoring, rolling updates and alerts

Self-Driving Car Control * Jan 2023 – Apr 2023

- Trained a self-driving car using CARLA simulator using **DQN based approach** for reinforcement learning algorithms
- Improved car control and safe navigation with **Imitation learning and SARSA** with 5% less training loss than baseline