

MRITYUNJAY ABHIJEET BHANJA

☎ +1 (518)-930-6153 📍 New York, NY ✉ mb9348@nyu.edu 🌐 [/mewtyunjay/](https://mewtyunjay/) 📺 [/mewtyunjay](https://mewtyunjay/)

Education

New York University | Masters of Science in Computer Science

Sep. 2023 – Present

Coursework: Artificial Intelligence, Cloud Computing, Big Data, Machine Learning, Algorithms, Data Science

Amity University Haryana | Bachelors of Technology in Computer Science

Aug. 2017 – May 2021

Experience

New York University

April 2024 - Present

Software and AI Developer

New York, NY

- Developing an AI-driven tool leveraging Large Language Models for facilitating better data discovery, enhancing critical data literacy and facilitating better research.
- Engineering web and mobile applications integrating custom AI solutions to facilitate educational data analysis, promoting nuanced understanding of data biases and power dynamics.

Maersk

Aug. 2021 – Aug 2023

Associate Data Scientist

Bengaluru, India

- Contributed to DataLab, an open-source platform for data analytics and ML with JupyterLab, Spark on Kubernetes. Integrated Spark Monitor, Spark Lineage, and Config-driven DAG tool, improving data process visibility and saving \$4 million annually.
- Researched voice analytics systems using various open-source STT models, such as OpenAI's Whisper, for real-time transcription of orders over phone calls, enabling insights and reduced turnaround time. Achieved a WER of 30% on transcribed calls.
- Conducted comprehensive EDA for a large-scale freight estimation model, enhancing accuracy on freight cost estimation by 20% over existing system.
- Independently engineered a backend system for comprehensive Data Lineage across Maersk's commercial products using Python, MongoDB, and Azure Cloud.
- Innovated a Cost-Forecasting dashboard, powered by Python and Streamlit, providing comprehensive insights into resource utilization and enabling precise future cost predictions per resource group for Maersk's Azure subscriptions.

Hewlett Packard Enterprise

April 2021 – July 2021

AI Engineer Intern

Bengaluru, India

- Developed and deployed computer vision modules, leveraging YOLOv3 and SSDMobilenetV2 in a Smart-City POC. Achieved a Mean Average Precision (mAP) of 0.86, enabling real-time Custom Object Detection of defaulters on bikes without helmets.
- Utilized HoloGAN to create a 3D image generator capable of reconstructing 3D facial images from a single 2D photograph.

Megthink Solutions

June 2020 – Aug. 2020

Data Science Intern

Gurugram, India

- Conducted early-stage Parkinson's Disease Detection, achieving a 98% accuracy rate, and optimizing classification algorithms on a PD dataset using specialized metrics, EDA, and feature refinement to prevent severe motor symptoms.
- Executed Deployment strategy by operationalizing the XGBoost prediction model through a Flask (Python) web application, successfully hosting the same on the Heroku platform.

Projects

Serverless Dining Concierge Chatbot on AWS | *API Gateway, LexV2, Lambda, S3, SQS, SES, DynamoDB, Opensearch, S3*

- Developed a chatbot using Lex and Lambda to provide restaurant email recommendations to users based on their preferences utilizing AWS services like DynamoDB to store user preferences and Amazon OpenSearch for restaurant data indexing, Amazon SES for seamless email communication and Amazon S3 for reliable static website delivery.

AI Culinary Assistant | *LLM, RAG, GPT, Pinecone, Langchain*

- Built a RAG, utilized OpenAI ada embeddings and Langchain with GPT-4 for similarity searches on Pinecone, vectorizing recipes and fridge inventory to tailor personalized recipe recommendations based on user preferences and ingredient availability.

Prostate Cancer Detection | *Computer Vision, Image Processing, Transfer Learning, EfficientNet*

- Engineered a Prostate Cancer Detection system with PyTorch and EfficientNetV2, achieving a 0.71 QWK through meticulous image preprocessing from 383GB to 3GB, optimized training with Adam and BCEWithLogitsLoss, and enhanced data augmentation for robust early-stage identification.

Technical Skills & Certifications

Programming: Python (scikit-learn, numpy, pandas, matplotlib, seaborn, nltk, tensorflow, keras, PyTorch, OpenCV, OpenAI, Streamlit, Flask, Django), PySpark, C, C++, HTML, CSS, Javascript, Typescript, SQL, bash, zsh

Technologies/Frameworks: LLMs, Langchain, Flutterflow, PyTorch, Tensorflow, Flask, Keras, OpenAI, Streamlit, Linux, Hadoop, Git, Azure, AWS, DevOps, ETL, Data Science, Machine Learning, Computer Vision (Classification, Object Detection and Recognition, Semantic Segmentation, Face Recognition, Pose Estimation, Feature Extraction, Optical Character Recognition), NLP (Text Classification, Summarization, Question-Answering, Chatbots, Named Entity Recognition, Automatic Speech Recognition), Large Language Models, HuggingFace, Time Series Forecasting, REST APIs, Docker, Kubernetes, Apache Spark, OOP, Asynchronous Programming, APIs, Powerpoint, AWS (Lambda, S3, API Gateway, EC2, SES, SQS, SNS, CloudWatch, Athena, OpenSearch), Azure (Data Factory, Data Lake, Databricks, Synapse, Cognitive, ML Studio), H2O, Dremio, Node.js, React.js, REST API, GraphQL.

Certifications: Microsoft Certified Azure Data Scientist, Microsoft Certified Trainer, Microsoft Certified Trainer

Publications

Model Selection for Parkinson's Disease Classification Using Vocal Features [Springer] – *Bhanja, M. Chaudhary, S.*

Mar. 2021