

SRIPHANI VARDHAN BELLAMKONDA

☎ 3477448482

✉ SRIPHANI345V@GMAIL.COM

🏠 1143, 67TH BROOKLYN, NEW YORK, 11219

🌐 BELLBYTES.NETLIFY.APP

Education

New York University

Master of Science in Computer Science; GPA: 4.0/4.0

Sept. 2023 - Present

Manhattan, NY

National Institute of Technology Warangal (NITW)

Bachelor of Technology in Computer Science and Engineering; GPA: 7.2/10.0

Aug. 2018 – May 2022

Telangana, India

Technical Skills & Relevant Courses

Languages: C, C++ , Python , Java ,Javascript, Typescript

AI ML Frameworks: numpy, pandas, scikit-learn, PyTorch, Tensorflow, AWS Sagemaker

Databases: MySQL, SQL, AWS DynamoDB, SQLAlchemy, MongoDB, NoSQL

Big Data & MLOps: Hadoop, Spark, Kubernetes, Docker, CI/CD, HDFS, MapReduce, AWS

Web Technologies: HTML, CSS, Django, Flask, Spring Boot MVC (Java), FastAPI, Node, React, Angular

Certifications: AWS Solutions Architect SAA-C03, Azure Fundamentals AZ900

Coursework: Mathematics, Linear Algebra, Machine Learning, Statistics, Software Engineering, Databases, Big Data

Work Experience

AI/ML Research Assistant — New York University — Manhattan, New York

March 2024 - Present

- Part of the DARPA funded Computational Cultural Understanding (CCU) project to identify cross-cultural misunderstandings in dialog
- Conducted Multilingual Speech Recognition and Diarization using Whisper, a state-of-the-art Speech Language Model, implemented with PyTorch and Hugging Face Transformers, processing over 9 terabytes of multilingual audio/video data on Linux High-Performance Computing Clusters.
- Designed and optimized end-to-end ML pipelines for data preprocessing, model inference, and evaluation, utilizing Python, Numpy, and Pandas, orchestrated via SLURM, resulting in a competitive Word Error Rate (WER) of approximately 10%.
- Programmed shell scripts to crawl and backup 10 TB and files and directories with deduplication strategy.

Software Engineer — Wells Fargo — Hyderabad, India

July 2022 - July 2023

- Designed scalable architecture for an enterprise system with 10+ modules, employing Java Spring Boot and Angular Typescript.
- Attained a code coverage of over 95% through unit-testing with JUnit and Mockito, ensuring software robustness and reliability
- Migrated from MS Access to SQL Server and algorithmized in python a query grammar translator reducing manual efforts by 98%
- Developed diverse backend algorithms, including server side pagination, filtering, search, CRUD operations, and SQL queries for Oracle database, achieving a performance enhancement of 80%.

Software Engineer Intern — Wells Fargo — Hyderabad, India

May 2021 - July 2021

- Architected a scalable full-stack Deep Learning solution for credit risk modelling with 5 classification models and performed Exploratory Data Analysis and Preprocessing on the Dataset.
- Developed dashboard monitoring with Flask (Python) & Firebase(NoSQL) optimizing and retrieval speed less than 500ms
- Implemented user authentication using OAuth 2.0, token-based authorization mechanisms with JWT, and caching which boosted performance by 70%.

Undergraduate Researcher — NIT, Warangal — Warangal, India

Sept. 2021 - May 2022

- Introduced a novel mini batch sampling method which utilizes feedback from the previous minibatch for active sampling.
- Developed custom Machine Learning training scripts in Tensorflow and scikit-learn achieving 1-4% increase in accuracy.
- Experimented with Logistic Regression, Neural Networks, and Convolutional Neural Networks on 4 different datasets on classification tasks.

Projects & Publications

Paper: NLP Transformer models on Low Resource Language | *PyTorch, Web Scraping*

- Curated a Telugu Humor Dataset and **fine-tuned Multilingual BERT models** on A100 GPU for sentiment analysis.
- Published in **ACL 2022** Workshop with a F1 test score of 0.82 using XLM-RoBERTa.

Machine Learning Pipeline for Text Summarization with Sagemaker (MLOPS) | *Python, Transformers, HuggingFace, AWS*

- Orchestrated the data ingestion, validation, transformation, model training, and evaluation by 2 methods.
- Method 1: Containerized the app with Docker and AWS EBS, served the inference and training endpoints on AWS EC2 with the backend by FastAPI and NGNIX, performed Automated deployments and CI/CD with Github Actions.
- Method 2: Preprocessed raw data and stored in AWS S3 bucket, applied Lambda to trigger a training job using Hugging Face Deep Learning Container and EC2, then served via AWS Sagemaker endpoint while managing model versions with model registry.

MongoDB Flask app with Kubernetes | *AWS EKS, Kubernetes, Flask, MongoDB*

- Constructed a Flask(Python) app with MongoDB in a Linux environment using **minikube** and onto **AWS using EKS service**.
- Parameterized the **replication controller feature** with 3 replicas and the **rolling update** strategy.

Lightweight ResNet Architecture | *HPC, SLURM, PyTorch*

- Experimented with ResNet Architecture by altering the block sizes, kernels, skip connections etc. with less than 5×10^6 params
- Achieved Accuracy of 96.2% on the CIFA-10 dataset while resnet18 had achieved only 86%

Big Data Analytics on NYC Taxi Data | *PySpark, Hadoop Map Reduce*

- Processed NYC Taxi Data on HDFS with **Hadoop Map Reduce** and PySpark (Python)
- Performed analytics by joining multiple datasets and writing custom queries in a parallel manner boosting efficiency over 85%.