Paritosh Marathe

New York, NY | pmparitoshmarathe@gmail.com | 917-733-1902 | https://www.linkedin.com/in/paritosh-marathe

EDUCATION

Binghamton University, SUNY, Thomas J Watson College of Engineering and Applied Science

Master of Science, Computer Science *August 2022 - May 2024*

Relevant Coursework: Machine Learning, Deep Learning, Design and Analysis of Algorithms, Social Media Data Science and Pipeline, Design and Analysis of Algorithms, Operating Systems

Vishwakarma Institute of Information Technology

Bachelors of Engineering, Computer *August 2016 - May 2020*

TECHNICAL SKILLS

***Programming Languages*:** Python, Java, C, Dart, Bash, Shell Script
Cloud Computing: Azure DevOps Services, Docker, REST
Data/AI Tools: Numpy, Matplotlib, Seaborn, Grafana, Wandb
ML Frameworks: Pytorch, Tensorflow, Keras, Detectron2, Scikit-learn, Hugging Face, LangChain, OpenCV
Machine Learning: CNN, RNN, GAN, Random Forest, Decision Trees, SVM, Clustering, Llama2
***Database**:* PostgreSQL, SQL
***Software and Tools*:** Visual Studio Code, Django, Azure DevOps, Visio, Git, Android Studio, Flutter

PROFESSIONAL EXPERIENCE

Wolfspeed | *Solutions Architect Intern* | Durham, North Carolina August 2023 – May 2024

- Created a comprehensive data dictionary for MES model data, enhancing data management and facilitating clear understanding of data structures and attributes for **300+** database objects within the system.
- Contributing to LaserTec project focused on enhancing data flow by **20%** for archiving purposes and code optimizations.
- Researched application of machine learning for **Predictive Maintenance** of manufacturing equipment under Industry 4.0.
- Demonstrated proof of concept in production for an android based optical character recognition application using Flutter and Java, in manufacturing, facilitating real-time batch tagging and enhancing efficiency by **15%**.

Wolfspeed | *Manufacturing Executive Systems Intern* | Durham, North Carolina May 2023 – August 2023

- Standardized and templated Azure CI/CD pipelines, enabling faster deployment for .NET/Python applications by **30%**, reducing manual configuration and aligning with **Test Driven Development** (TDD) practices for stable releases.
- Integrated pipelines with Sonatype **Nexus** repository for **on-premise** artifactory, fortifying against unauthorized access.
- Implemented SonarQube into build pipeline for static code analysis, ensuring code quality, identifying potential bugs, security vulnerabilities, and code smells early in the development lifecycle, reducing critical vulnerabilities by 12%.
- Devised segmentation model in Python for segregated segment binning from wafer defect maps using watershed and image processing techniques, enhancing defect detection accuracy by 30%
- Explored containerized deployment of GPT4All using Llama model implementing Retrieval-Augmented Generation (RAG) techniques to enhance dynamic content retrieval.

ASN Computer Education | *Instructor* **|** Pune, India August 2021 - July 2022

- Delivered 120+ hours of instruction on IT concepts and cybersecurity skills to a cohort of 50+ students.
- Trained 40+ students in Object-Oriented Programming (OOP) using JAVA, leading to a 90% success rate in application development and empowering students to create functional software solutions.

Anveshak Technologies and Knowledge Solutions | *Associate Developer* | Pune, India September 2020 – June 2021

- Developed a web scraping and PDF extraction pipeline, incorporating data preprocessing and cleaning for efficient contract tender information extraction, increasing data retrieval speed by **35%** in comparison to manual methods.
- Designed and deployed deep learning models for image data analysis using Python, with 95% image recognition accuracy.

PROJECTS

Analyzing Topics, Framing, and Hate Speech in Online Political Engagement

- Developed and deployed a robust ETL pipeline to ingest, process, and store over 200 GB of data from Reddit and 4Chan APIs into a PostgreSQL database, ensuring efficient management and retrieval for large-scale analysis.
- Engineered a natural language processing algorithm leveraging NLTK, optimizing hate speech detection capabilities and enabling sentiment analysis on 10,000+ user comments.
- Built **Grafana** dashboards to visualize key insights, such as online discussion trends, user engagement metrics, and hate speech prevalence, allowing for real-time monitoring and in-depth exploratory data analysis.