Sean (Eashan) Joshi

Software Engineer

New York, USA

+15852024547 eashanj4@gmail.com linkedin.com/in/eashanjoshi eashanjoshi.github.io

EDUCATION

Master of Science (MS), Data Science (Specialization: Artificial Intelligence) **Aug. 2022 - Dec. 2024**

Rochester Institute of Technology - ***3.74/4.00*** *Rochester, NY, USA*

- Coursework: Machine Learning, Software Construction, Artificial Intelligence, Applied Data Science, Applied Statistics & Calculus, Data Warehousing, Big Data, Generative AI, Visual Analytics

Bachelor of Engineering (BE), Computer Engineering **Aug. 2016 - Jun. 2020**

Pune Institute of Computer Technology *Pune, India*

- Relevant Coursework: Operating Systems, Embedded Systems, Microcontrollers, CPU architecture

TECHNICAL SKILLS & CERTIFICATIONS

Certifications: **AWS Certified Cloud Practitioner** (Amazon Web Services), **Google Project Management** (Google)

Languages & Databases: Python, JAVA, VB.NET, C#, C++, R, SQL, NoSQL, AWS DynamoDB, MongoDB
Data Processing & Management: OLTP, OLAP, ETL, Data Pipelines, A/B testing, AWS Lambda, AWS EC2
Professional software & tools: Linux, SAP, AWS Services, .NET, SAS, Apache Spark, Pentaho, Docker, Git, Tableau
Libraries: LLM, GAN, Scikit-learn, Numpy, Pandas, Seaborn, Keras, NLP, Tensorflow, PyTorch, OpenCV
Web Technologies: HTML, XML, CSS, PHP, JavaScript, Django, JSON

RESEARCH

Debugging Optimization: LLM-Generated Multi-Turn Queries using DevGPT dataset. **Aug. 2023 - Present**

Dr. Mohamed Wiem Mkaouer, Director, Dept. of Software Engineering, RIT *Rochester, NY, USA*

- Executed an analysis engine on DevGPT dataset, uncovering **1,957,830** ChatGPT mentions in GitHub issues, **3,050,232** in commits, and **25,601** in discussion sections.
- Performed sentiment analysis on issues, yielding an average **sentiment score of 0.11**, & found a **correlation coefficient of 0.1906** between debugging complexity and programming languages used.
- Engineered a **GPT-4** based interactive debugging system, generating **multi-turn queries** for precise bug localization.

WORK EXPERIENCE

Full Stack Software Engineer | ***.NET, JavaScript, SQL*** **Sept. 2020 - Sept. 2021**

Yardi *Pune, India*

- Resolved **30+ development defects**, & developed **50+ new client features** in Trust Accounting, a cloud-based financial fiduciary management platform using **.NET & SQL**, and maintained production systems within coding standards.
- Conducted Query Optimization (**OLTP** and **OLAP**) for enhanced efficiency in Trust Accounting, achieving a **10% transparency boost** in asset management, client data visualization and retrieval times.
- Engineered **2 Custom Programming Reports** using **SSRS, SQL**, & integrated automation solutions for loan, tax & insurance management enhancing cloud compliance & reducing **financial data error rate by 15%**.
- Collaborated with key stakeholders to establish a robust communication framework, fostering clear and efficient information flow during release planning, resulting in a **10% reduction** in project delays.

PROJECTS

Advanced Financial Analysis & Risk Assessment of Listed Companies (Startup Project) | *Python, ML.* **LINK.** **Feb. 2024**

- Executed data cleansing on a financial screening dataset with **80 indicators** provided by FinSharpe, a SEBI registered investment advisor, and conducted quantitative risk analysis, identifying **362 high-risk stocks**.
- Performed market cap segmentation using 'FinSharpe Overall Score' and 'Value Score' to isolate **240 and 470 firms** for in-depth performance evaluation, which aided in securing a **$120,000** investment for FinSharpe.

Duplicated Bug Report Detection using Siamese Convolutional Neural Network | *TensorFlow, NLP.* **LINK.** **May 2023**

- Processed Eclipse Bug Report Dataset using TF-IDF, Cosine Similarity, and N-grams for categorization, and crafted an embedding matrix for training the model, achieving an overall **93% accuracy**.
- Achieved **54% and 71% accuracy** for similar and dissimilar bug reports, respectively, aiding future enhancements in Bugzilla's categorization capabilities.

Multiclass Image Classification using Deep Convolutional Neural Network | *Python, TensorFlow.* **LINK.** **Nov. 2022**

- Created a DCNN by modifying VGG16 architecture achieving **80% accuracy** on Tiny ImageNet dataset.

LEADERSHIP

Recreation & Fitness Supervisor: Supervised a team of 10 Fitness Attendants at RIT Wiedman Fitness Center.