

# EASHAN JOSHI

Software Engineer

Rochester, NY, USA

☎ +1-585-202-4547 ✉ ej2637@rit.edu 🌐 <https://www.linkedin.com/in/eashanjoshi> 🐙 <https://github.com/EashanJoshi>

## Education

### Rochester Institute of Technology

Aug. 2022 – April 2024

Master of Science, Data Science

Rochester, NY, USA

- Coursework: Fundamentals of Data Science & Analytics, Software Construction, Applied Statistics, Business Intelligence, Data Warehousing

### Pune Institute of Computer Technology

Aug. 2016 – Jun. 2020

Bachelor of Engineering, Computer Engineering

Pune, Maharashtra, India

- Coursework: Engineering Mathematics, Data Structures, Software Engineering, Database, Theory of Computation, Computer Networks, Soft Computing & Optimization, Data Mining, Cybersecurity, AI, ML

## Technical Skills

**Languages and Databases:** Python, JAVA, VB.NET, C#, C++, R, SQL, MongoDB, OLTP, OLAP

**Professional software and tools:** .NET Framework, JMP Pro, VS Code, Google Colab, Tableau, MySQL Workbench

**Libraries:** Scikit-learn, Numpy, Pandas, Matplotlib, Fuzzy, Seaborn, Keras, NLTK, Tensorflow, OpenCV

**Web Technologies:** HTML, XML, CSS, PHP, JavaScript

**Management:** Lean Six Sigma

## Work Experience

### Full Stack Software Engineer | .NET, JavaScript, SQL

Sept. 2020 – Sept. 2021

Yardi Software India Pvt. Ltd.

Pune, Maharashtra, India

- Implemented client-driven enhancements to Oil and Gas and Tax and Insurance modules, resulting in streamlined processes and increased operational effectiveness.
- Solved **30+ developmental defects** and **50+ client configuration issues** and debugged the code to identify errors and fixes.
- Updated the OLTP and OLAP SQL queries which reduced the time to display client data by **10%**.
- Engineered a cutting-edge financial Custom Programming Reports (CPR) using Yardi Spreadsheet Reporting (YSR), employing intricate SQL queries; streamlined financial analysis procedures, saving an average of **5 hours per week** and enabling real-time data-driven decision-making.
- Collaborated with key stakeholders to establish a robust communication framework, fostering clear and efficient information flow during release planning, resulting in a **25% reduction** in project delays.

## Projects

### Duplicate Bug Report Detection using Siamese Convolutional Neural Network | Python, nltk.

May. 2023

- Managed to pre-process primary Eclipse Bug Report Dataset from Bugzilla & Constructed 2 bug categories using TF-IDF & Cosine Similarity and, N-grams & quartiles to separate textually similar & dissimilar bug reports.
- Generated an embedding matrix of the description column and trained the training data having **0.001** learning rate, **12** epochs & **256** batch size.
- Evaluated the model on test data which gave an accuracy of **93%** & accuracy of **54%** & **71%** for textually similar and dissimilar data respectively.

### Multiclass Image Classification using Deep Convolutional Neural Networks | Python, tensorflow.

Nov. 2022

- Engineered a base model which was a Deep Convolutional Neural Network with 6 layers on Tiny Imagenet dataset that gave a Validation Accuracy of **57.15%** in **100** epochs.
- Researched and Selected 2 pre-trained models, VGG16 and Xception for comparison.
- Concluded by ranking models according to validation accuracy, VGG16: **80%**, Base: **57.15%**, Xception: **35%**.

### Exoplanet Discovery through Machine Learning Algorithms | Python, scikit-learn.

Nov. 2022

- Conducted data cleaning on the NASA-provided Kepler Objects of Interest dataset, ensuring accuracy and completeness, resulting in a refined and reliable dataset for analysis.
- Implemented Logistic Regression, K-Nearest Neighbor, Decision Tree, Random Forest algorithms & achieved highest accuracy of **96%** through Random Forest algorithm.

### Early Prediction of Heart Attack using Machine Learning Algorithms | Python, scikit-learn.

May. 2020

- Selected 13 important clinical features and implemented Random Forest algorithm for classifying heart disease based on these clinical features and achieved an **accuracy of prediction is 84%**.
- Led the development of a state-of-the-art Data Visualization system that utilizes patient clinical data to predict heart disease, saving **50%** of medical professional time, ultimately saving lives and improving patient outcomes.