

Srikar Sarma, Parnandi, V.

srikarsarma@gmail.com | Ph.no: 214-228-1231 | <http://in.linkedin.com/in/srikarpv> | github.com/srikarpv

Summary

- Software Developer with 5+ years experience working on ML, AI and web applications.
- Skills: Python, Java, C++, C#, JavaScript, Bash, Scala, MATLAB, Perl; OS & DB: Linux, OS X, Windows, UNIX, MySQL, MongoDB, Postgres, Hive-QL; ML & Cloud: AWS, Azure, Git, Docker, Jenkins, Tableau, Kafka, Triton, RAG, and PyTorch tools.

Employment

Software Engineer-Machine Learning, Iron Mountain

Sunnyvale, CA

June 2022 - Present

- Designing, Implementing and Deploying Deep Learning models for Intelligent Document Processing, integrating Natural Language Processing, Computer Vision, multimodal LLMs like LLaT, and GenAI.
- Implemented complete machine learning lifecycle, from conducting research and development to implementing production-ready solutions. Built reusable AutoML pipelines.
Tech stack: Kubeflow, MLflow, ONNX, Triton, RAG, Open source LLM, langchain, Pytorch.
- Part of a team that developed Iron Mountain's Lending AI vertical, delivering scalable machine learning solutions for package splitting, document identification, and entity validation, now serving three banking clients. The AI system reduced loan processing costs from \$3 per package to just 1 cent, becoming a significant income stream for Iron Mountain.

Software Development Engineer-Data Lab, PricewaterhouseCoopers(PwC)

Tampa, FL

Sept 2019 - June 2022

- Designed and implemented a web app for SAP financial data Anomaly Detection, identifying fraudulent transactions for a Fortune100 client.
- Conversational AI ChatBot using seq2seq, TensorFlow; training and report generation with Kubernetes CronJobs.
- Built dashboards (MEAN) to fetch relevant data to monitor team activities and intelligent diagnostics for PwC leadership.
- Developed automation scripts, microservices for VM automation to reduce repetitive tasks, save man hours and expenditure for the VDI usage using PowerShell, Python, and Bash. Built desktop applications using WPF, C# and Applescript.

Machine Learning Intern(AI/NLP), Verizon Wireless (Point of Sale)

Piscataway, NJ

May 2018 - August 2018

- Test case recommender system and classification based on Java/JS source code commits from Git and AccuRev to speedup Regression Testing using TF-IDF, word2vec, Stochastic Gradient Descent classifier, Pandas. Patent: 10810115

Research Assistant, Institute for Simulation and Training, CASS

Nov 2017 - May 2019

University of Central Florida, FL

- Deep feature based on a convolutional auto-encoder for compact semantic hashing. (Submitted for the ACM SIGGRAPH / Eurographics Symposium on Computer Animation (SCA))
- Predicting Block I/O from system traces using the seq2seq TensorFlow model for prefetching I/O addresses.
- Worked on Modeling and Simulation of user behavior in a social network (Github) using Repast Symphony library.

Data Scientist, MetalOT Technologies

Aug 2016 - July 2017

Bangalore, India

- Built forecasting models using accelerometer data to detect potholes, speed bumps, and assess driver performance, generating ratings and trip predictions.

Education

Master of Science, Computer Science

Aug 2017 - May 2019

University of Central Florida, Orlando, FL

GPA: 3.8

Bachelor of Engineering, Computer Science

June 2012 - Aug 2016

Visvesvaraya Technological University, Bangalore, India

GPA: 3.3 (top 15% of the class)

Coursework: Natural Language Processing, Machine Learning, Algorithms, Computer Vision, Data Structures in C++, Data-Intensive Computing, Intelligent Systems, Artificial Intelligence, Computer Security and Forensics.

MOOC and Certifications: *Machine Learning- Andrew Ng, Deep Learning Specialization-* Coursera, *Python for Data Science and ML Bootcamp-* Udemy, *Algorithms - Princeton University, Data Scientist's Toolbox-* Coursera, *Parallel Programming with CUDA-* Udacity

Projects:

- Artistic Style Transfer with Convolutional Neural Networks. Used *Tensorflow, PIL, OpenCV, Seaborn*.
- Fake News detection with Keras. Used *sklearn, pandas, beautifulsoup*.
- Real-time Intelligent system for Driver Fatigue detection. Used *Keras, PIL, OpenCV, Numpy, Scikit, RESNET, CNN's*
- Q learning, DQN, DRQN(Deep Recurrent Q-network), A2C based AI that plays atari games. Used *gym, Tensorflow, AWS*

Awards and Achievements

- Winners Infosys Aspiration 2020 Programming contest held at Infosys DC, Bangalore.
- "Project of the Year Award" Pratyush-2016, Won the project of the year award in the Department Fest for final year Project.