SHRIVATHS SHYAM

San Diego, CA
+1 (619) 953-8113 shrivaths1998@gmail.com LinkedIn GitHub

SKILLS

Programming Languages: Python, JavaScript, SQL, HTML/CSS, Java, C/C++
Frameworks/OS/Tools: Pandas, PyTorch, Jenkins, Git, AWS, MS SQL Server, Apache NiFi, Postman, FastAPI, ReactJS, PM2, NGinx, Bash, Windows, Linux

EXPERIENCE

NLitenData LLC **Aug 2024 – Present**
AI Engineer *San Diego, CA, United States*

- Designed and developed microservices for our product using Python (FastAPI) and Node.js.
- Implemented ETL pipelines using Apache NiFi to ingest data from external sources into AWS RDS (SQL Server).
- Designed optimized SQL database schemas for both application performance and analytical querying, improving data efficiency.
- Developed a FastAPI chatbot using GPT-4o API and the database schema to provide context-aware answers to user queries.
- Built a GPT-4o API based document parser using prompt engineering to extract key details for downstream processing.
- Performed end-to-end testing of product features and external API integrations using Postman, ensuring system robustness.
- Deployed the application on AWS EC2 using Git for version control, with NGINX and PM2 for service management and logging.
- Contributed to the product website design using Miro and developed front-end components with ReactJS, hosted on AWS Amplify.
- Developed predictive models (Logistic Regression, XGBoost) to analyze user engagement trends, informing product design decisions.
- Presented product demos to external clients for validation, effectively communicating value propositions to stakeholders.
- Led Scrum sprints from planning through deployment, managing feature development cycles and cross-functional collaboration.

Talmo Lab, Salk Institute for Biological Studies **Jul 2023 – Jun 2024**
Research Engineer *La Jolla, CA, United States*

- Built new features for SLEAP, an object detection and pose estimation Python library with more than 100k downloads.
- Performed data preprocessing tasks such as frame extraction, normalization, and annotation cleaning using OpenCV and FFmpeg.
- Trained UNet and ResNet models with hyperparameter optimization to improve prediction accuracy and performance.
- Designed and implemented tracking algorithms, such as optical flow and Kalman tracking, to improve inference accuracy.
- Created modular API functions for tasks such as data preprocessing, pose estimation, and visual analytics.
- Built and deployed SLEAP as a Python package compatible with Linux, Windows, and MacOS using git and GitHub Actions for CI/CD.
- Resolved bugs and provided user support via GitHub issues.

Cognizant Technology Solutions **Nov 2020 – May 2022**
Programmer Analyst *Chennai, TN, India*

- Implemented core business functionalities using Pega-PRPC, a business process management (BPM) tool, for Prudential Financial.
- Designed REST API flows and SQL-based logic to streamline data handling in case management systems.
- Developed and hosted full-stack web services using Java, JavaScript, and HTML/CSS via Apache server.
- Managed and oversaw the Continuous Integration/Continuous Deployment (CI/CD) pipeline using Jenkins.

PROJECTS

Fediblend - Fediverse Feed Amalgamator | *HTML, CSS, JavaScript, Git, Agile* | **[Code]** **Oct 2023 – Dec 2023**

- Designed and developed a landing page that pulls trending posts from user-specified instances of the federated social media network.
- Improved the loading time of the page by 30% with parallelization of API calls.
- Implemented test functions in JavaScript with 95% code coverage which ensured that the webpage was robust.

Multi Modal Recommender Systems for Rating Prediction | *Pandas, Vaex, Plotly, PyTorch* **Oct 2023 – Dec 2023**

- Performed web scraping of Google Local Reviews data and executed data engineering tasks using Vaex.
- Utilized Plotly for data visualization and exploratory data analysis (EDA).
- Generated text embeddings with pre-trained *BERT* and image embeddings with pre-trained *VGG16* as multi modal features.
- Developed an Explicit Feedback Recommender System using Deep and Cross Networks, achieving **0.59** MSE through ablation studies.

Uber Data Analysis Using BigQuery | *SQL, Google Cloud Storage, Mage-AI, Python* **Jul 2023 - Aug 2023**

- Designed and executed ETL pipelines to ingest and process Uber trip data using BigQuery and Mage-AI.
- Conducted exploratory data analysis (EDA) and built interactive dashboards for trip pattern analysis.
- Used SQL and Pandas to generate insights into ride demand and pricing trends.

Efficient and Goal-Conditioned Learning in RL using TD-MPC | *PyTorch, Python, Linux, Git* | **[Code]** **Apr 2023 – Jun 2023**

- Improved the performance of existing Temporal Difference-Model Predictive Control (*TD-MPC*) by using offline demos.
- Experimented with conditioned goal generation using a GAN and achieved faster goal completion with 25% lesser training steps (∼150k steps) compared to fixed goal (∼200k steps).

Knowledge Distillation on Transformers | *Python, PyTorch, Transformers* **Apr 2023 – Jun 2023**

- Performed fine-tuning of pre-trained BERT and GPT-2 models using textbrewer on SST-2 and CoLA datasets.
- Reduced the size of 12-layer teacher models by 50-80% while having accuracy drops of only 8-10% and performed comparative analysis on different sizes and their accuracies.

EDUCATION

University of California, San Diego **Sep 2022 – Jun 2024**
MS ECE - Machine Learning and Data Science, GPA: 3.52/4 *La Jolla, CA, United States*
Coursework: Statistical Learning I, Linear Algebra, ML for Physical Applications, Deep Generative Models, Optimization and Acceleration of DL on Various Hardware platforms, Recommender Systems, Software Engineering, Data Systems for Machine Learning

SSN College of Engineering, Anna University, Chennai **Jun 2016 – Apr 2020**
BE - Electronics and Communication Engineering, GPA: 8.88/10 *Chennai, TN, India*