HARSHITH ROOPA MANJUNATH

harshu98@gmail.com | LinkedIn | GitHub | 230 W Hunter Ln., Fremont, CA - 94539-6338 | +1 (510)-770-4996

Professional Summary

- **I am a passionate software developer** skilled in transforming **business requirements** into **precise functional specifications** and translating **complex concepts** into **clean**, **maintainable code**.
- **My commitment** to consistent **innovation** and **growth** stems from the **sense of progress** and **fulfillment it fosters**, which drives my passion for writing and refining code. I am currently building large language models (**LLMs**) from **scratch**, using insights from **Sebastian Raschka's book** to deepen my **expertise in AI/ML**.
- **Certified in Machine Learning**, with expertise in concepts such as **ridge/lasso regression** (e.g., capacity planning, resource allocation), **kNN** and **Bayesian** methods for recommendation and sentiment analysis (e.g., ticket prioritization, incident classification), and **clustering** techniques like **k-means** (e.g., customer segmentation).
- **Proficient in dataset preprocessing**, including **feature engineering**, **normalization**, **sparse data handling**, and **partitioning datasets** for training, validation, and testing. During my OPT period at Key Solutions Inc., I applied these techniques to **large-scale datasets** from **grants.gov** and **USDA.gov**, ensuring high-quality data readiness.
- **Strong foundation in Java** and Object-Oriented Programming, **SOLID** principles, enterprise design patterns, and **MVC architecture** and following **best practices** in **web development** for scalable, maintainable solutions.
- **Experienced in Spring Boot microservices** with **RESTful APIs**, leveraging caching, interceptor-based handling and **Redis** for performance optimization. Proficient in **JPA** with **Hibernate**, utilizing **indexing** and **lazy loading**.

Skills

Operating Systems: Linux (distributions - Ubuntu, Linux Mint, Fedora), Windows (7 - 11)

Prog. & Scripting: Bash, Perl, C++, Java, Go, Python, SQL, Typescript, Javascript, HTML5, CSS3

Data Ecosystems: MySQL, PostgreSQL, MongoDB, Redis, Amazon S3, ElastiCache, Amazon Aurora & RDS

ML & Data Analytics: Keras, TensorFlow, PyTorch, Ollama, Transformers Architecture, RAG Framework, RLlib

Frameworks & Dev Tools: Spring, Maven, Gradle, Thymeleaf, React.js, Vue.js, Redux, GraphQL, Jest, Cypress

Cloud Infra. & DevOps: AWS EC2, GCP, Azure, Terraform, Ansible, Docker, Kubernetes, Git, Jenkins, Jira

Experience

Software Engineer Associate - Key Solutions Inc. **Fremont, CA**

Protocol Search Engine & Resource Utilization Reports - Worked in a Team of 4 Mar 2024 - Sep 2024

- Worked on implementing the architectural design of a protocol search engine, developing a classification framework to distinguish between recommendation and search-based user queries within the eProtocol application.
- Designed a distributed data pipeline using Apache Spark, utilizing Spark's Structured Streaming API to sync operational databases with Delta lakes, ensuring consistent data synchronization for training datasets.
- Built a dynamic reporting system to generate real-time and historical reports based on user-defined criteria providing insights into active/inactive protocols and storage utilization across IRB, IACUC, IBC, and CS modules.
- Designed and implemented an intuitive UI, streamlining data input and retrieval, reducing report generation time by 40%, and enabling seamless access to module-specific metrics with just a few functional input parameters.
- Significantly enhanced the operational database performance by strategically implementing indexing and utilizing advanced caching techniques, resulting in a 35% reduction in query execution time within a period of 2 months.
- Implemented report scheduling on a monthly and quarterly basis by leveraging Maven for build management and Jest for rigorous Test-Driven Development, attaining a 99% test pass rate before production deployment.
- Drove Agile development, led 250+ SCRUM meetings and used Open Project to track roadmaps and deliverables ensuring seamless CI/CD deployment with Jenkins and Git across distributed computing environments.

Software Engineer Intern - Key Solutions Inc. **Fremont, CA**

Support Tools Framework - Individual Contributor Sep 2023 - Feb 2024

- Developed support tools for the customer support team in eProtocol using the Spring Web Framework, deploying the microservices on Kubernetes and Docker, boosting operational efficiency by 30% through automation.
- Gathered and documented business requirements, creating 10+ user personas and mapping functional specifications to use cases and unit tests, and implemented RESTful interfaces for each support tool.
- Worked alongside senior developers to implement cron jobs for automating user-heavy tasks, pagination for efficient data retrieval, reducing the response time for each tool by 25%, and effectively clearing roadblocks.

Software Engineer Intern - NTTF Electronics & IT — **Bangalore, India**
Smart Smoke Detector System - Project Lead — Jun 2019 – Aug 2019

- This internship aimed to formulate a problem task and solve the potential risks in daily life. Led a team of five members in developing a solution for hazardous smoke and fire-related activity in apartments.
- Developed a real-time sensor network simulator using 4 types of sensors comprising Temperature (DHT11), Smoke, Humidity, and Air Quality Index (AQI) sensors, all connected to an Arduino for reading the input.
- Interfaced the Arduino input readings with a Raspberry Pi microprocessor with the help of an HC05 Bluetooth module, which functioned as a local server logging in the data onto a spreadsheet.
- After training three ML models on the smoke detection IoT dataset, the logistic regression model achieved the best performance with 98.36% accuracy, while feature evaluation yielded an F1 score of 96.7%.

Education

Stevens Institute of Technology — **Hoboken, NJ**
Master's in Computer Science - Graduate Certificate in Machine Learning — Aug 2021 - May 2023

BNM Institute of Technology — **Bangalore, India**
Bachelor in Computer Science and Engineering — Aug 2016 – Nov 2020

Projects

NASA Launch Website - MERN Stack — Jun 2023 – Aug 2023

- Developed the NASA Launch Website using the MERN stack, implementing the MVC pattern to manage the front and back end. The UI allows users to create launch missions by entering mission details, such as rocket, dates, and destination, using the Arwes web framework for futuristic designs, animations, and sound effects.
- The React.js front end interacts with the back end through Express API routes to handle user input, display historical data, and display upcoming launches. Redux manages the state across components.
- Middleware is used for error handling, authentication, and security, optimizing routing and ensuring efficient front and back-end communication.
- The Express API processes requests and connects to a cloud-hosted MongoDB database via Mongoose for reliable data persistence and querying.

Toxic Spans Detection - Sequence Labeling Task using TensorFlow — Feb 2023 –May 2023

- The intuition behind this Sem-Eval 2021 Task was that of Named Entity Recognition (NER). The task involved classifying word tokens into two categories (toxic or not) and returning the indices of the toxic words in a sentence.
- This study reveals that the data type and labeling criteria are more crucial for malicious speech identification than the model alone. Models trained on expert annotations from the civil comments dataset outperformed those trained on amateur ones.
- Experimented on the dataset using a batch size 16 for training on LSTM CRF and CRF layers with RoBERTa as the embedding layer, executing for 2 epochs, and yielded F1 scores of 64.73% and 65.96%.
- This task utilizes the Hugging-Faces' RoBERTa transformer and Adam optimizer with a learning rate of 3e-5.

Playing Atari with Deep RL - Dueling DQN using Keras and Gym — Oct 2022 – Dec 2022

- Led a team of four members in developing a CNN model with a DQN algorithm to train an agent to play three Atari:2600 games from the openAI gym environment.
- Employed Markov Decision Process framework and performed Image preprocessing on the observations, storing them in a replay buffer. Built a fully connected 2D Convolutional Neural Network with ReLU activation.
- Trained the DQNAgent with Linear Annealed Policy to implement the Epsilon greedy action selection on the model.

Rice Grain Quality Determination - Classification Task using PNN — Mar 2020 – Jun 2020

- Using a radial basis function for network activation, we developed a Probabilistic Neural Network model to determine the quality of rice grains. The model achieved high precision and recall scores of 96.83 and 97.32.
- Collaborated with a team of 4 members and published a paper in the Springer Journal at the International Conference on Sustainable Communication Networks and Application (ICSCN 2020).

Certifications & Awards

First Place in Innovative Project Lab - Summer Competition | Winter Competition

Machine Learning A-Z | SQL for Data Science | Cisco Prog. in Python | Entrepreneurship | Internet of Things