Sahil Bhagchandani

✉ sahil.bhagchandani520@gmail.com | LinkedIn

Experience

Walmart Global Tech — *United States*

SOFTWARE ENGINEER – AI — *Jan. 2023 - Present*

- Engineered and deployed generative **AI pipelines** using **Python Flask**, **Hugging Face Transformers, Node.js API** and **React UI** to support use cases such as content summarization, automated report generation, and chatbot development
- Developed scalable **data preprocessing** pipelines using **Pandas** and **SQL** to handle multi-million row datasets, ensuring high-quality input for model training and inference
- Designed, fine-tuned, and served **Large Language Models (LLMs) (BERT, LLaMA)** via **Node.js** and **Python Flask** microservices, enabling text classification, entity recognition, and semantic search with 90% accuracy
- Designed and implemented machine learning and deep learning models including **RAG, Gradient Boosting, XGBoost, and Logistic Regression** to solve problems in text classification, entity recognition, semantic search, and predictive analytics
- Built interactive dashboards and performance analytics using **Tableau** and **Looker**, enabling business stakeholders to monitor model performance, data drift, and usage metrics in real-time
- Implemented **A/B testing** and model evaluation frameworks to compare multiple model versions using statistical metrics like **F1-score, precision, recall,** and **BLEU scores**

Intuit — *United States*

SOFTWARE ENGINEER INTERN — *May. 2022 - July. 2022*

- Assisted in developing AI-powered **content recommendation system** using **React, Nodejs, LangChain** and **Pinecone VectorDB**, resulting in a 40% increase in user engagement
- Developed a multi-model AI assistant integrating **OpenAI GPT**, enhancing customer support efficiency by 60%
- Implemented a real-time document analysis pipeline using **LlamaIndex** and **PostgreSQL**, processing over 1M documents daily with 99.9% accuracy
- Performed time series analysis using techniques such as **VAR (Vector AutoRegression) and NNAR (Neural Network AutoRegression)** to model temporal patterns, detect trends, and forecast system metrics and user behavior
- Development of a **Generative AI platform** for creating synthetic **training data, improving model performance** by **30%** across various use cases

Skills

Programming Languages	Python, Java, JavaScript, SQL, Go
Databases and Framework	GraphQL, MySQL, PostgreSQL, Databricks, Redshift, PyTorch, Tensorflow, FastAPI, React, Tableau, Looker
AI/ML Technologies	GenAI, LangChain, RAG, BERT, RoBERTa, GPT, NumPy, Pandas, Transformers, Scikit-learn
Cloud Technologies	Azure, AWS EC2, ECS, ECR, Docker, Kubernetes
Testing Tools	Jasmine, Junit, Langsmith, TruLens, PyTest, Hydra
Source Control Management	Git, TFS, Linux
Certifications	AWS Security Fundamentals, Cisco Python Essentials
Soft Skills	Communication, Collaboration, Problem Solving, Teamwork

Education

San José State University — *San Jose, California*

MASTER OF SCIENCE IN COMPUTER SOFTWARE ENGINEERING — *Jan. 2021 - Dec. 2022*

Indus University — *Gujarat, India*

BACHELORS OF TECHNOLOGY IN COMPUTER ENGINEERING — *Aug. 2016 - Jun. 2020*

Publication

Imbalanced Network Traffic Intrusion Detection Using Deep Learning — *Aug. 2022 - Dec. 2022*

- Applied **Edited Nearest Neighbor (ENN)** and **K-Means** clustering to separate easy and hard samples, reducing redundant majority-class data by up to 99.9% while preserving critical minority patterns
- Leveraged **Python, Scikit-learn, and TensorFlow** to build, train, and benchmark ML/DL pipelines, optimizing hyperparameters, batch sizes, and learning rates to prevent overfitting

Projects

Automated Car Parallel Parking System — *Sep. 2022 - Dec. 2022*

- Implemented Convolutional Neural Network (CNN) algorithms to detect, identify, and locate available parking spaces based on sensor data
- **Technology Stack: Python, C++, Node.js, React, ROS2, Raspberry pi, Breadboard, Sensors, Infrared camera**

COVID-19 Face Mask Detection — *Sep. 2021 - Dec. 2021*

- Employed TensorFlow and NumPy for model development, preprocessing, training, and testing, ensuring robust performance and accuracy
- **Technology Stack: Python, NumPy, TensorFlow, OpenCV, YOLO, CNN**