Ashwin Sateesh Kumar

Dallas, TX | 925-445-6494 | ashwinsateesh5@gmail.com | linkedin.com/in/ashwins24 | github.com/ashwin-sateesh

Experience

AI Software Engineer, AVP Dec. 2024 – Present
Citibank *Dallas, TX*

- Reduced false positives by 30% and improved detection accuracy by 25%, by building a dependency-aware deep code analyzer with source-to-sink reasoning into a ReAct-based multi-agent vulnerability detection system
- Cut manual penetration testing lead times by over 50% by integrating a report generation agent that delivers structured vulnerability analysis, exploit simulation, and remediation guidance into the ReAct pipeline
- Scaled ML-driven threat assessment to 8,000+ analysts by productionizing a multi-tool security platform backed by a custom MCP server for threat modeling, architecture diagram, and cloud log analysis
- Improved retrieval precision by 25% and reduced redundant vector lookups by 40% by architecting a microservices RAG system with async chat history, BM25 + cross-encoder reranking, RAGAS evaluation, and prompt caching

Machine Learning Engineer Jan. 2024 – Dec. 2024
Abecedarian *Boston, MA*

- Architected a multimodal Yoga Assistant, achieving 2x faster image inference while preserving CLIP alignment by distilling Stable Diffusion via UNet pruning and fine-tuning GPT-3.5 via a queue-based continual learning pipeline
- Led the design of a multimodal agricultural policy recommendation system combining hyperspectral crop modeling (ViT + PCA), LSTM-based risk forecasting, and BERT-derived market analysis into a fine-tuned GPT engine
- Delivered grounded policy recommendations to stakeholders via LangChain orchestration and Streamlit UI

Research Assistant July 2023 – Nov. 2023
Northeastern University *Boston, MA*

- Uncovered visual and linguistic patterns in social media by training a disentangled multimodal β-VAE, enabling precise control over image attributes; deployed on GCP with bfloat16 quantization, improving inference by 30%

Machine Learning Engineer Intern, R&D June 2022 – Dec. 2022
Signify (Phillips Lighting) *Boston, MA*

- Modeled adaptive lighting behavior across 18 smart homes using SARIMAX and XGBoost, achieving 97% accuracy for energy-efficient automation
- Engineered a user re-identification system for personalized home automation, achieving 0.95 mean average precision (mAP) in identity matching across multi-camera views using an omni-scale feature learning architecture

Trainee Software Engineer July 2019 – Nov. 2020
KPIT Technologies Ltd. *Bengaluru, India*

- Developed an autonomous driving perception system using U-Net traffic scene segmentation with 0.89 Intersection over Union (IoU) and vision–radar sensor fusion with kinematic features for real-time object detection
- Improved perception model performance on 1M+ images by enhancing annotation quality through transfer learning and human-in-the-loop validation

Technical Skills

Languages: Python, SQL (PostgreSQL), R
Frameworks: PyTorch, TensorFlow, Scikit-Learn, LangChain, Hugging Face, FastAPI, Flask
Developer Tools: Git, Bash, Docker, GCP (Vertex AI, Cloud Storage), AWS (S3, SageMaker), Red Hat OpenShift
Libraries: NumPy, pandas, SciPy, OpenCV, NLTK, XGBoost, FAISS

Projects

HealthBot | *Python, PyTorch, Flask, HTML, Docker* Sep. 2023 – Dec. 2023

- Developed a chatbot generating medically relevant responses (0.75 semantic similarity) by fine-tuning FLAN-T5 with LoRA and RLHF (PPO), integrating BERT named entity recognition and disease classification (0.96 F1)

Education

Northeastern University Boston, MA
Master of Science in Data Science ***(GPA: 3.8/4.0)*** *Sep. 2021 – Dec. 2023*

PES University Bengaluru, India
Bachelor of Engineering in Electronics and Communications *Aug. 2015 – May 2019*