Vamsi Krishna Koppala

☎ (940) 999-8706 ✉ vamsikvk1234@gmail.com in Vamsi-Krishna Vamsi-Krishna63

Professional Summary

Highly skilled **Software Engineer** and **Research Assistant** specializing in **cloud computing**, **DevOps**, **infrastructure management**, and **AI-driven applications**. Proficient in **Microsoft Azure**, **Linux**, and **Docker** with hands-on experience in **CI/CD pipelines** and **automation**. Experienced in developing **full-stack solutions** that integrate **Large Language Models (LLMs)**, advanced **NLP techniques**, and **sensitive data detection** frameworks using technologies like **Meta-Llama-3**, **BERT**, and **Whisper**. Skilled in real-time data processing, **semantic search**, and web development using **Flask**, **React**, and **Node.js**. Certified in **Azure Developer Associate (AZ-204)**, **AWS Cloud Practitioner**, and **Oracle Cloud Infrastructure Foundation Associate**, with a passion for building secure, scalable, and intelligent cloud and AI solutions.

Education

Texas Tech University, Lubbock **Jan 2024 – May 2025**
Master of Sciences in Computer Science
GPA: 3.9/4.0

Annamacharya Institute of Technology and Sciences, Tirupati **Jun 2018 – Jun 2022**
Bachelor of Technology in Computer Science
GPA: 8.23/10.0

Experience

Research Assistant, Texas Tech University *April 2024 - May 2025*

- Conducted applied research in **Sensitive Data Detection**, integrating **Speech-to-Text (STT)** systems with **Large Language Models (LLMs)** to identify sensitive information in unstructured text and audio.
- Engineered a full-stack **Python** web application combining **Flask**, **HTML5**, **CSS3**, and **Bootstrap** for real-time processing of audio and text data with sensitive information detection.
- Developed a dual-model sensitive data detection framework using **traditional pattern matching** (regex, Word2Vec similarity) and **LLM-based contextual analysis** powered by **Meta-Llama-3-8B-Instruct.Q8_0.**
- Implemented **speech recognition** pipelines utilizing **OpenAI Whisper** with **GPU acceleration** via **PyTorch**, enabling real-time transcription from microphone inputs and audio file uploads.
- Built a **Qdrant vector databas**e for efficient semantic similarity search of sensitive terms based on **Word2Vec** embeddings, enhancing approximate matching and classification performance.
- Designed and deployed an **adaptive feedback mechanism** that dynamically refines detection models based on user corrections, ensuring continuous learning and accuracy improvement.Integrated text **pre-processing techniques** such as **tokenization**, **stemming**, **lemmatization**, stop-word removal, and TF-IDF vectorization to enhance feature extraction.
- Integrated **Transformer-based NER** models like **BERT fine-tuned** and **DistilBERT model** for entity recognition and **Zero-shot classification** models (**Facebook/BART-large-MNLI**) to augment semantic understanding and context analysis.
- Conducted comparative analysis between traditional methods and LLM-based methods, improving detection accuracy and **reducing false positives** through **confidence scoring** and **overlap analysis**.
- Developed interactive **data visualization dashboards** and **downloadable analysis** reports to streamline user **interpretation** of sensitive data detection results.

Associate Professional Software Engineer, DXC Technology *May 2023 - Dec 2023*

- The team **executes VM (virtual machine)** and **physical server OS patching operations** for security purposes while maintaining **stability and industry compliance.**
- The software developer has **expertise** in **Azure VM OS patching** and **troubleshooting** while also **deploying automated deployment** of patches to enhance system performance.
- A total of **95%** of system complex issues on **CentOS** and **RHEL** systems were successfully resolved by my **troubleshooting efforts**.The system **performance** and **reliability** gain improvement through **automated patch deployment** methods.
- The professional maintains a specialization in **infrastructure management** where they excel at **complex infrastructure** design work and maintenance tasks in **Linux**, **VMware**, and **AIX** environments.
- My **expertise** includes detailed understanding of **virtualization** along with **Microsoft Azure** and other services such as **VMware ESXi**, **Hyper-V**, and **IBM AIX virtualization.**
- Experienced in **Azure Virtual Network infrastructure** creation as well as **Network Security Group** deployment, VM Scaling (**VMSS**), and **Load Balancer configurations** to **optimize** cloud performance.

- Security policies are implemented with three core components: **SSH key management**, **role-based access control** (**RBAC**), and **access control** methods.
- Worked with **DevOps tools** such as **Jenkins**, **Docker**, and **Git**, contributing to **CI/CD pipeline development**, containerized application deployment, and **version control** best practices.

Junior Software Intern, Sola Info IT Solutions Pvt Ltd *Apr 2022 - Mar 2023*

- Hands-on experience in **Microsoft Azure**, specializing in **deploying, configuring, and managing Azure Virtual Machines (VMs), Azure Blob Storage, Azure Virtual Networks (VNet),** and **Azure Resource Manager (ARM)** templates to optimize cloud infrastructure.
- **Proficient in Linux system administration**, particularly with **Red Hat Enterprise Linux (RHEL)**, focusing on **user management**, **file system handling and process automation to ensure system stability and security.**
- **Configured and optimized networking resources**, including **Azure Load Balancer**, **Network Security Groups (NSG)**, **Virtual Private Network (VPN) connections**, and **DNS settings**, to improve system connectivity and security.
- **Developed and automated infrastructure solutions** using **Ansible**, and **PowerShell scripting**, **enabling Infrastructure as Code (IaC)** for efficient **cloud resource provisioning and configuration management**.

Projects

On Demand Professor Q&A Bot

- **Deployed and configured** the **Qdrant vector database** via **Docker**, establishing a **highly scalable** and **efficient vector storage system** for seamless integration with the **LLM-powered Q&A bot**.
- **Integrated GPT4ALL as the primary AI engine**, enabling **localized model training** on knowledge documents and **real-time internet-based query expansion** for comprehensive response generation.
- **Designed an optimized document retrieval pipeline** using **SentenceTransformers**, ensuring **accurate semantic embedding, indexing, and page-specific query resolution** to enhance user experience.
- Implemented an **API-driven architecture** to support multi-modal query processing, ensuring efficient retrieval and improved response accuracy for **domain-specific questions**.

Project Shield: Safeguarding Against Deceptive Attacks using Clickjacking

- **Developed and implemented** both **client-side and server-side security** measures to protect web applications from **clickjacking attacks**, enhancing the **safety and integrity of user interactions** online.
- **Conducted comprehensive testing** and vulnerability assessments to identify potential clickjacking risks, ensuring **robust security** measures were in place and effectively **mitigated threats.**
- **Integrated Content Security Policy (CSP) headers** and **X-Frame-Options** to restrict unauthorized iframe embedding, preventing malicious overlay-based attacks and enhancing application security.

Neuro-Symbolic Concept Revision Using Interactive Explanations

- **Developed a pipeline leveraging Neuro-Symbolic Explanatory Interactive Learning (NeSy XIL)** to improve model interpretability and accuracy by addressing **Clever-Hans behavior**, using **CLEVR-Hans datasets for robust evaluations**.
- Conducted extensive implementation and debugging to reproduce and enhance results from state-of-the-art research, **achieving up to 94.96%** accuracy on complex datasets by integrating **symbolic reasoning with neural network models**.
- Optimized feature selection using **attention-based explainability** methods, improving **model generalization** and **reducing overfitting** in vision-language reasoning tasks.

LLM Privacy Evaluation & Defense Framework (LLM-PBE)

- **Implemented four attack vectors** are Data Extraction, Jailbreak, Membership Inference, and Prompt Leakage. These are used to rigorously evaluate privacy vulnerabilities across multiple **LLMs** (**LLaMA2**, **Mistral**, **Gemma**, **Phi**, **Deepseek-R1**) using **Ollama** for efficient local model inference.
- Developed a **Python-based pipeline** to automate attack execution, logging, and accuracy analysis using **prompt engineering**, **semantic similarity metrics**, and **fuzzy matching** for precision measurement.
- Engineered **scrubbing and defensive prompting** modules to mitigate data leakage, achieving **100% mitigation** on select models and **significant accuracy reductions** on others.
- Conducted **comparative benchmarking and visual analytics** using **Matplotlib** and **Pandas**, producing graphs and result tables to assess the effectiveness of privacy-enhancing technologies (PETs).
- **Optimized execution performance** for local LLMs with **GPU offloading**, batch inference strategies, and runtime logging to support reproducibility and scalability on **12GB GPU environments.**

Employee Attrition Prediction

- Leveraged **Machine Learning (ML) algorithms** such as **Logistic Regression, Decision Trees, Random Forest**, and **SVM** to predict employee attrition, optimizing predictive accuracy using transformer-based models.

- Implemented data **preprocessing**, **exploratory data analysis**, **and model evaluation** using **Python, Power BI, and Tableau.**
- Enhanced model performance by applying **feature selection** and **hyperparameter tuning** techniques, resulting in a significant increase in prediction accuracy and interpretability.

Technical Skills

Languages: Python, C, Java, HTML, CSS, Javascript

Database: MySQL, NoSQL, MongoDB, Qdrant

Platforms: Linux, MacOS, Visual Studio, Eclipse, Windows

Web Development: React, Node.js, Flask, Bootstrap 5

Cloud Technologies: Microsoft Azure, AWS

Devops & CI/CD: Docker, Jenkins

NLP Techniques: Named Entity Recognition (NER), Zero-shot Classification, Semantic Similarity Matching

AI/ML Frameworks: Hugging Face Transformers, SentenceTransformers, PyTorch

Version Control & Tools: Git, GitHub, GitLab

Visualization Tools: PowerBI, Tableau, Matplotlib, Microsoft Excel

Speech Recognition: OpenAI Whisper, Vosk

Web Technologies: REST APIs, FormData API

Professional Certifications

Az 204 - Microsoft Certified: Azure Developer Associate

Microsoft Certified: Azure Fundamentals (AZ-900)

AWS Certified Cloud Practitioner

Oracle Cloud Infrastructure 2022 Certified Foundation Associate