

Venkata Harshitha Bathala

+1(682)-406-1948 |  venkataharshithab01@gmail.com |  [Git Hub](#) |  [Portfolio](#)

Arlington, Tx, 76013

Professional Summary

I am a dedicated graduate student pursuing a master’s in computer science at the University of Texas at Arlington, with a strong academic foundation in data structures, the software development lifecycle, and Agile methodologies. My role as a Graduate Research Assistant provided practical experience in web development and UNIX environments, while my position at DXC Technology sharpened my skills in data analysis and process automation. With proficiency in Python, Java, and C/C++, I am committed to delivering efficient and robust technology solutions. Additionally, I have knowledge of DevOps tools like Docker, Kubernetes, and CI/CD pipelines, which I am eager to apply in real-world environments. My certifications as an AWS Certified Cloud Practitioner, Microsoft Azure (AZ-900), and Lean Six Sigma Yellow Belt underscore my dedication to continuous learning and technical excellence.

EDUCATION

UNIVERSITY OF TEXAS AT ARLINGTON - *Arlington, Tx, United states*2024-Present

MASTER'S IN COMPUTER SCIENCE

- CGPA – 3.8/4.0
- Coursework: Data Mining, Web Development, Artificial Intelligence, Machine Learning, Computer Vision, Advance Software Engineering, Distributed Systems

ANNAMACHARYA INSTITUTE OF TECHNOLOGY AND SCIENCES - *Tirupati, AP, India*2018-2022

BACHELOR OF TECHNOLOGY - COMPUTER SCIENCE AND ENGINEERING

- CGPA – 8.34 /10
- Coursework: Web Development, Cybersecurity, Data Structures, Database Management Systems, Software Engineering, Operating Systems, Computer Networks

PROFESSIONAL SKILLS

- Operating System:** Linux, Unix, Windows
- Web Development:** HTML, CSS, JavaScript, PHP
- Programming languages:** C, Java, Python
- Data Analysis and Visualization:** Power Bi, Excel, Python, Tableau
- Cloud:** Microsoft Azure administrator, AWS
- DevOps:** Docker, Kubernetes, CI/CD (Git, Ansible)
- Database:** SQL, MongoDB
- Libraries & Frameworks:** TensorFlow, Keras, PyTorch, Scikit-learn, NLTK, Pandas, NumPy, Matplotlib, Seaborn
- Shell Script:** bash
- Software Development Tools:** JIRA, UCML, STA
- Server:** Apache

WORK EXPERIENCE

University of Texas at Arlington - *Arlington, TX*07/2024-PRESENT

GRADUATE RESEARCH ASSISTANT

- I am responsible for both the **Development** and **Administration Tasks**, prioritizing and executing them based on the specific requirements and needs of the department.
- Developing a Simplified Portal for UTA’s CSE Faculty** to streamline access to essential resources, previously buried under multiple layers of the university’s website. This website will be accessible only by the CSE department.
- Extracted and organized crucial information**, such as inventory details, networking and security policies, supported operating systems and tools, completed projects, important forms, and request submissions for faculty queries.

- The technologies utilized include **HTML, CSS, JavaScript, and PHP** for the programming languages, **Apache** for the web server, and **MySQL** for the database.
- For **Administration Tasks**, I manage the setup and configuration of lab systems for the CSE department.
- **LUKS encryption** was chosen after researching various methodologies, primarily for its wide compatibility across Linux distributions. Its robust encryption features, including multiple keys and secure disk wiping, ensure comprehensive data protection.
- Implemented automated **Bash scripts** for antivirus deployment to maintain system integrity and a user management script to automate student account creation, enforce access restrictions, and manage permissions within lab environments.
- **Optimized virtualization resources**, such as VirtualBox, to ensure students receive enhanced computing capabilities for coursework.
- **Monitored and maintained UTA's two servers**, specifically tracking the percentage increase in resource usage to ensure system health and maintain uptime.
- Operating systems worked on include **Debian-based** (Ubuntu) and **RPM-based** (CentOS, Red Hat, Fedora) systems.

DXC Technology - *Chennai, INDIA*

03/2023 - 12/2023

ASSOCIATE PROFESSIONAL SOFTWARE ENGINEER

- Served as an integral contributor to both Lean and Data Analysis teams, merging process improvements with advanced analytics.
- Led automation initiatives in Lean processes, achieved **5% FTE** savings and enhanced operational efficiency.
- Designed and optimized **Value Stream Mapping (VSM)** strategies to eliminate bottlenecks and improve workflow.
- Spearheaded automation management using **JIRA, UCML, and STA tools**, driving improvements in process reliability and operational efficiency.
- Collaborated directly with clients representing global brands in the automotive and luxury sectors; developed dynamic dashboards in Power BI and Tableau to empower data-driven strategic decisions.
- Integrated machine learning models for predictive analytics, enhancing forecasting accuracy and operational insights.

TECHNICAL PROJECTS

Sentiment Analysis on Song Lyrics (Natural Language Processing): [*Git Repository*](#)

- **Developed a sentiment analysis model** for highly imbalanced song lyric data (Spotify dataset: 5,000
- **TensorFlow**, and **Transformers**. Applied text preprocessing techniques like **tokenization**, **stop-word removal**, and **lemmatization** using **NLTK**, while leveraging **Word2Vec** embeddings with **Gensim** for feature extraction. Balanced the training data using **stratified splitting** and **data augmentation** via the **nlpaug** package to address class imbalance.
- **Fine-tuned the Bidirectional LSTM** with various hyperparameter configurations, achieving **75% accuracy** on the validation set. Further improved the model's classification performance by implementing **BERT** via the **Transformers** library, reaching **77% accuracy**.
- **Evaluated model performance** by building **confusion matrices** using **Scikit-learn** and implemented a **real-time emotion prediction function** capable of analyzing song lyrics and classifying their emotions effectively.

SWIM Protocol: [*Git Repository*](#)

- **Failure Detector Component: Docker containers** (node 1 to node 5) in the membership list ping each other every T' seconds via a gRPC **proto file**. If a node doesn't respond, an indirect ping is initiated by contacting other K-nodes in the membership list.
- **Dissemination Component:** Once a node's failure is detected, the detecting node **multicasts** this failure information to the rest of the membership list. Then, a new container (e.g., node 6) joins the membership, ensuring continuous system updates.
- **Deployment & Tools:** Implemented with 10 containers using a compose.yaml file. All container requirements are installed via a Dockerfile, and gRPC tools (e.g., protocol buffers, gRPC libraries) were used

Smart HealthCare Hub: [*Git Repository*](#)

- Developed a full-stack healthcare management platform connecting patients, providers, administrators, and pharmacists, featuring secure authentication, appointment booking, health record management, symptom checking, and real-time messaging.
- Designed and implemented the solution using HTML, CSS, and JavaScript for the front end, PHP and MySQL for the back end, with Axios integrated for seamless API interactions to enhance healthcare accessibility and efficiency.
- Integrated secure authentication and role-based access control, ensuring data security and compliance.
- **Technologies Used:** HTML, CSS, JavaScript, PHP, MySQL, APIs.

CERTIFICATIONS

- AWS Certified Cloud Practitioner
- Microsoft Azure Certification (AZ900)
- AIGPE LEAN SIX SIGMA YELLOW BELT CERTIFICATION