

# Subhasya Tippareddy

Seattle, WA | (656) 216-5144 | subhasya.tippareddy@gmail.com | [LinkedIn](#)

## WORK EXPERIENCE

### Software Engineer, Experian, USA

Oct 2024 - Present

- Led full software development lifecycle for a centralized workflow automation platform on AWS using Apache Airflow, from single-node to distributed architecture, enabling 15+ teams to host/manage ETL pipelines and large-scale batch processing.
- Designed cloud-native monitoring & alerting frameworks, reducing incident resolution time by 30%.
- Automated the team onboarding workflows, enhancing efficiency by 50% and reducing manual dependencies.
- Collaborated cross-functionally to resolve connectivity and scaling issues, ensuring platform reliability.
- Migrated application from EC2 VMs (Celery) to EKS (K8s) to create an isolated, cost-efficient scaling & match the dynamic workloads.
- Technologies: **Python, Jenkins, Docker, Kubernetes, Helm, Terraform, EMR, EKS, ECR, SQS, Spark, Java**

### Research Assistant, Dept. of Computer Science, University of South Florida

Aug 2024 - Jan 2025

- Created an NLP pipeline to extract structured insights from unstructured uterine cancer pathology reports using OCR.
- Built a transformer-based model for extracting key medical entities, mapping findings to standardized ICD-10 codes.
- Developed a **RAG-based clinical QA system** with vector databases and an interface using React and Vercel AI SDK to allow patients/doctors to ask natural language questions about reports, reducing manual interpretation time by 60%.
- Technologies: **Python, AWS Textract, SpaCy, Transformers, ClinicalBERT, LangChain, Chroma, FastAPI, SageMaker**

### Software Engineer, Experian, India

Sept 2021 - July 2023

- Engineered an automated file scanning service to perform malware detection on external files before transferring them to destinations within the organisation; the service is used globally, processing more than 100 million files across 7 regions.
- Constructed dashboards in Splunk with real-time log analytics to enhance file-tracking visibility and provide seamless user experiences for delivery teams.
- Automated DAG deployment and infrastructure provisioning by implementing CI/CD pipelines, improving deployment efficiency by 45% and reducing manual intervention by 70%.
- Mentored 3 junior engineers, providing hands-on technical guidance, onboarding documentation, and code reviews.
- Technologies: **Python, Jenkins, Terraform, Apache Airflow, RDS, VPC, EFS, SQS, EC2, S3, Cloudwatch**

## TECHNICAL SKILLS

**Programming Languages:** Python, Java, JavaScript

**Frontend Development:** HTML, CSS, React.js, Vercel AI SDK

**Frameworks & Libraries:** Django, Flask, FastAPI, React.js, Node.js, Spring Boot, JUnit 5, Mockito

**Cloud & Technologies:** AWS (EKS, S3, RDS, EFS, EMR, Lambda, EC2, SageMaker, Bedrock), Terraform, Apache Airflow

**Databases:** MySQL, PostgreSQL, MongoDB, Redis, Chroma, FAISS

**Deployment:** Docker, Kubernetes, Jenkins

**ML& NLP:** NumPy, Pandas, Matplotlib, Plotly, Scikit-learn, TensorFlow, Keras, PyTorch, RNNs, LSTMs, NLTK, SpaCy

**Generative AI:** Hugging Face Transformers, LangChain, OpenAI API, Anthropic Claude API, Langfuse

**Developer & Collaboration Tools:** Git, Bitbucket, JIRA, Confluence, Google Colab

## PUBLICATIONS

### Augmenting Legal Decision Support Systems with LLM-based NLI for Analyzing Social Media Evidence

Mar 2025

Developed a framework using **LLMs** and **NLI** to analyze social media evidence for legal decision support. ([arXiv:2410.15990](#))

Technologies: **Python, pytorch, transformers, pandas**

## EDUCATION

### Master of Science, Computer Science, University of South Florida, FL

Aug 2023 - May 2025

Coursework: Artificial Intelligence, Natural Language Processing, Security & Privacy in ML, Social Media Mining

### Bachelor's, Computer Science, Neil Gogte Institute of Technology, India

Jul 2018 - Jun 2022

Coursework: Object-Oriented Programming in Java and Python, Data Structures and Algorithms, Machine Learning, Data Science

## PROJECTS

### Multilingual AI Text Portion Detection [ Transformers | PyTorch | Pandas | Numpy | JSON ]

May 2024

- Developed a multi-lingual tool for detecting machine-generated text, trained with 330,000 samples in 23 languages.
- Achieved ~90% recall in identifying human and AI-generated text, outperforming existing solutions.

### Toxic Comments Classification [ Transformers | PyTorch | Numpy | Pandas | Scikit-learn | Seaborn ]

Dec 2024

- Classifying comments into various categories to determine the toxicity and obscenity using BERT.
- Assessing YouTube's moderation, checking the relevance of the model trained with past data on the current data (collected using the YouTube Data API).

### SnACK, A Food Ordering System [ Python | Django | SQL | HTML | CSS | Javascript ]

June 2021

- A responsive website for students to pre-order food, in order to reduce the waiting time and crowd at the college canteen.
- It also helps the cook to assess the quantity of food