Bhargav Sri Sai Nama

Boulder, CO | 720-334-3495 | bhna2906@colorado.edu| linkedin.com/in/bhargav-nama | github.com/bhargavn | bhargavnama.dev

Experience

Lead Software Engineering Intern - ML/NLP — Denver, CO
Mode Civil Services — Sep 2024 - Present

- Engineered an **ML-powered Autodesk automation system** with **C#** and **.NET8**, streamlining **CAD workflows** and reducing design cycle time by **20%** for **500+ civil engineers**.
- Developed and integrated a **sandbox environment for LLM-driven simulations**, enabling seamless integration of **natural language processing workflows** and interactive AI-driven user interactions.
- Deployed **LLM-powered automation** for contextual analysis and user interaction simulation, enhancing decision-support systems and increasing **workflow efficiency by 40%**.
- Architected **scalable ML pipelines** with **CI/CD automation (Azure DevOps, Git)**, reducing deployment cycles by **30%** and ensuring robust, production-ready models.

Research Intern — College Station, TX
Texas A&M University — Jun 2022 - Apr 2023

- Designed and launched **computer vision & ML models (TensorFlow, OpenCV, Apache Kafka)** for real-time monitoring, improving operational efficiency by **70%**
- Leveraged **AWS** for scalable storage and computation, and implemented **Apache Spark** to process large-scale streaming data, enabling faster insights and supporting critical decision-making.
- Synthesized and optimized **predictive algorithms** with **Scikit-learn** and **Python**, improving **process control accuracy by 40%** through advanced pattern recognition and trend analysis.
- Led a **cross-functional team** optimizing workflows, improving manufacturing precision by **60%** and reducing operational bottlenecks.

Projects & Publications

ChatGPT Hallucination Checker [GitHub] — Jan 2025 - Feb 2025

- Built a **Chrome extension** for **AI hallucination detection** using **Wikipedia** and **DuckDuckGo API**, integrating a **BERT-based NLI model** for fact verification, reducing hallucinated responses by **40%**.
- Created and deployed a **Flask backend** for LLM inference, reducing response latency from **800ms to 300ms**, and released an interactive **frontend UI** for real-time analysis.

ClauseRAG: Local Retrieval-Augmented Legal Document Analysis [GitHub] — Dec 2024 - Jan 2025

- Assembled a **RAG pipeline** for clause retrieval with **FAISS** and **SentenceTransformers**, increasing accuracy by **10%**, and enhanced **vector search** with optimized **FAISS indexing** for low-latency retrieval.
- Created a **Streamlit UI** for interactive document analysis and implemented an **Airflow-based pipeline** for automated clause extraction from **1,000+ legal documents**, processing **5,000+ clauses**.

LLM-Enhanced Teaching Assistant for Summarization and Knowledge Retrieval [GitHub] — Jan 2024 - May 2024

- Fine-tuned **BART, BERT, T5** for lecture summarization, improving comprehension by **30%** (**ROUGE-L: 0.45**) and optimizing NLP pipelines for real-time scalability.
- Prototyped a **context-aware Q&A system** with **Llama-2-70B**, achieving **85%** accuracy and reducing latency by **40%**.

Predictive Modeling of Unemployment and Crime Trends [GitHub] — Aug 2023 - Dec 2023

- Conducted **EDA** on **5M+** data points using **R**, identifying key **socioeconomic trends** affecting crime and employment, and improving **data accessibility** for decision-makers.
- Trained and fine-tuned **ML models (XGBoost, Random Forest)** to enhance crime forecasting accuracy and designed **interactive dashboards (Tableau, Matplotlib)** for real-time policy insights.

AI-Driven In-Situ Defect Detection and Quality Assurance in WAAM — Nov 2021 - May 2023

- Constructed a **computer vision pipeline** for real-time 3D point cloud analysis, boosting defect detection accuracy by **80%**, and applied **deep learning-based anomaly detection** (**TensorFlow, PyTorch**) to achieve **95%** classification precision.
- Implemented **MLflow** for experiment tracking and hyperparameter tuning, optimizing high-speed **sensor data streaming** (**Apache Spark, AWS Kinesis**) at **200 MB/sec**.
- **Publication:** Published in **Progress in Additive Manufacturing**, 2025. DOI: 10.1007/s40964-025-01004-9.

Technical Skills

Languages: Python, R, C#, MATLAB, SQL
Frameworks & Development: .NET8, ASP.NET, FastAPI, Flask
Machine Learning & AI: TensorFlow, PyTorch, Scikit-Learn, XGBoost, LightGBM, Hugging Face Transformers, LangChain
LLMs & NLP: GPT-4, BERT, RAG (Retrieval-Augmented Generation)
Data Engineering & Cloud: ETL Pipelines, Apache Spark, Apache Kafka, Databricks, AWS (EC2, S3, RDS, Lambda), Azure
MLOps & Deployment: MLflow, Docker, Kubernetes, Azure DevOps, CI/CD Pipelines
Data Visualization & BI: Tableau, Microsoft Excel, Plotly, Matplotlib, Seaborn

Education

University of Colorado Boulder — Boulder, CO
Master of Science in Data Science, CGPA - 4 — Aug 2023 - May 2025 (exp.)
Coursework: DSA, Computer Vision, Data Mining, STAT 1 & 2, Machine Learning, Information Visualization

Indian Institute of Technology Tirupati — Tirupati, India
Bachelor of Technology in Mechanical Engineering, CGPA - 8.64 — Jul 2019 - May 2023