

ANUSHRI BHANSALI

anushribhansali.14@gmail.com | <https://www.linkedin.com/in/anushri-bhansali/> | <https://github.com/AnushriBhansali>

EDUCATION

North Carolina State University, Raleigh

August 2023 to May 2025

Master of Computer Science

GPA:3.81/4

Database Management Systems, Automated Learning and Data Analysis, Software Engineering, Efficient Deep Learning, Neural Networks, Real Time Machine Learning & AI

Charotar University of Science & Technology

August 2019 to June 2023

Bachelor of Technology in Computer Science

GPA:9.58/10

TECHNICAL SKILLS

Programming Languages: Python, C++, TypeScript, Java, **Frameworks/Tools:** Flask, React, Node.js, Azure DevOps, AWS, Databricks

Database: SQL, NoSQL (MongoDB) **AI/ML:** Pytorch, Hugging Face Transformers, NumPy, Matplotlib, NumPy, Scikit-learn, CNNs, Apache TVM, Machine Learning Compilation Code, TensorFlow

WORK EXPERIENCE

Software Engineering and AI Intern, Lexis Nexis, Raleigh, NC, United States

June 2024 to August 2024

- Performed fine-tuning of Large Language Models (Meta Llama, Mixtral), achieving a 60% improvement in search testing framework performance. Adopted LLMs as judge to validate LLM reasonings and outputs.
- Created APIs and integrated AI models using Databricks workflows and finally saved 15 manhours weekly.

Software Engineering and AI Intern, Netweb Software, India

January 2023 to June 2023

- Built a real-time updating feature in the resume parsing platform, enabling recruiters to process 50+ resumes daily, leading to a 30% increase in interview scheduling efficiency.
- Engineered a vector-based semantic ranking pipeline using embeddings and RAG for context-aware candidate-job alignment, boosting shortlisting precision by 45%.
- Developed Java REST APIs and integrated a responsive React frontend with PostgreSQL backend, reducing recruiter workload by 15+ hours per week and accelerating hiring through AI-assisted shortlisting.

Machine Learning Intern, BISAG-N, Government of India

May 2022 to July 2022

- Built a malware detection model with 96% accuracy using classical ML and vectorization techniques.
- Enhanced model interpretability using tools like SHAP and feature importance plots, enabling security analysts to derive actionable threat classification insights and accelerate incident response times by up to 30%.

PROJECTS

License Plate Recognition Model Optimization

- Optimized LPRNet for efficient license plate recognition, reducing inference time from 0.25s to 0.028s and model size from 1.11MB to 0.51MB. Applied unstructured pruning, achieving 82.3% accuracy with a 0.71MB model.
- Implemented post-training quantization, further reducing inference time by 20% but maintaining competitive accuracy.
- Employed Apache TVM AutoScheduler, improving inference speed to 0.047s with 89.4% accuracy. Leveraged PyTorch, CUDA, TVM, ONNX, NumPy, and compiler-level optimizations (vectorization, parallelization, loop unrolling).

LLM-Driven Healthcare Support System

- Created a healthcare platform integrating Disease Prediction and Q&A models using RoBERTa, Python, and Flask.
- Designed a React frontend and a PostgreSQL database to manage patient information.
- Facilitated efficient interaction between application components through the implementation of RESTful APIs, resulting in reduced data retrieval times by 20%.

Data Driven Movie Recommendation System

- Engineered a high-accuracy recommendation model on the Netflix Prize dataset using matrix factorization and SVD, achieving an 18% improvement in NDCG through grid search and Bayesian parameter optimization.
- Conducted exploratory data analysis to extract user behavior signals, driving model refinements and enhancing personalization strategies.

RESEARCH & PUBLICATIONS

Interference Classification for Open RAN (O-RAN) xApp.

- Prepared a deep learning xApp for jammer detection in O-RAN using spectrograms from OTA testbed data on the Near-RT RIC, along with employing embedding space augmentation techniques (E-Mixup, E-Stitchup) to enhance generalization under noisy conditions.
- Trained ResNet, DenseNet, MobileNetV2, ViT, and ConvNeXt, performing comprehensive evaluation and t-SNE latent space analysis to optimize performance.

Language Identification Using Machine Learning & Vectorization Techniques

- Proposed multi-language detection system trained on vectorized phonetic inputs.

PUBLISHED under State Government of Gujarat SSIP research initiative.