

Varad Patwardhan

Raleigh, NC 27606 | varadpatwardhan.com | (919) 637-1593 | varadp2000@gmail.com

EDUCATION

NC State University, Raleigh, NC

Ph.D. Student in Computer Science

M.S. in Computer Science - (CGPA - 4.0/4.0)

August 2025 – May 2030

August 2023 – July 2025

University of Pune, India

Bachelor of Engineering in Computer Engineering - (CGPA-9.23/10)

August 2019 – May 2023

EXPERIENCE

Teaching Assistant, NC State University

August 2025 – Present

- Support undergraduate CS courses, mentor students, hold office hours, and guide projects in Java and algorithms.
- Assist in grading and labs, ensuring fairness and consistency.

Graduate Research Assistant (Master's Thesis), NC State University

January 2024 – August 2025

- Built Eclipse plugin capturing keystrokes, cursor, and environment data (8M+ events from 36+ students).
- Reconstructed real-time code progress, identifying coding behaviors (burst vs. consistent)
- Applied insights to intelligent tutoring systems, plagiarism detection, and real-time feedback dashboards.

Pentair Graduate Fellow, NC State University Libraries

January 2024 – August 2025

- Developed a staff-facing RAG system (LangChain, GPT-5, ChromaDB) for instant answers on policies and workflows.
- Created AI demos (LLMs, game AI, chatbots) showcased in various pop-ups.
- Delivered 10+ interactive educational AI experiences.

NLP Researcher, L3Cube

June 2022 – May 2023

- Collected/processed Hindi-English code-mixed datasets; optimized HingBERT achieving 95% F1 in NLP tasks.
- Conducted comparative benchmarking of BERT, RoBERTa, HingBERT, HingRoBERTa for low-resource settings.

PROJECTS

StudiosLLM | *LangChain, GPT-5, ChromaDB*

August 2025

- Designed and deployed a Retrieval-Augmented Generation model for NCSU Libraries staff.
- Indexed internal policies, workflows, and IT documentation, enabling staff to query complex queries via natural language.
- Reduced onboarding and support resolution time with instant, context-aware answers.

Bookbot AI | *LangChain, LLaMA, DeepSeek, Python*

December 2024

- Built a retrieval-augmented Q&A system over PDFs with LangChain and LLaMA and DeepSeek.
- Enhanced accuracy and contextual reasoning for diverse queries through custom dataset integration.

Detecting Osteoarthritis in Dogs | *Machine Learning, Computer Vision, Python*

April 2024

- Developed a multi-modal ML model using gait video data and Tekscan pressure data.
- Achieved 93% accuracy via feature engineering, deep learning pipelines, and validation.

TECHNICAL SKILLS

Languages: Python, C++, C, Java, SQL, Bash, JavaScript, C#

AI/ML Frameworks: PyTorch, TensorFlow, Keras, Hugging Face, LangChain, RAGs, ChatGPT API, LLaMA, DeepSeek

Tools: Git, Docker, Jupyter Notebook, Android Studio, Firebase

Expertise: Artificial Intelligence, Machine Learning, Deep Learning, Natural Language Processing (NLP), Large Language Models (LLMs), Data Analytics, Algorithms

PUBLICATIONS

EclipseMonitor: A Real-Time Student Programming Environment Data Collection Tool. - *SIGCSE TS 2025*

- Co-authored a tool enabling large-scale keystroke-level data collection from student programming environments.

Language Augmentation for Code-Mixed Text Classification. - *Elsevier NLP Journal, 2023*

- Improved BERT/HingBERT performance on low-resource Hindi-English datasets using augmentation, achieving up to 95% F1.

Comparative Study of Pre-Trained BERT Models for Code-Mixed Hindi-English Data. - *IEEE I2CT, 2023*

- Benchmarked BERT, RoBERTa, ALBERT, mBERT, HingBERT, and HingRoBERTa, showing code-mixed models outperform vanilla baselines.

Analyzing the Sentiments of Marathi-English Code-Mixed Social Media Data. - *Springer CCIS, 2022*

- Built and benchmarked a dataset of 1,000+ Marathi-English sentences for sentiment analysis with ML/DL models.