

CHAKRADHAR REDDY PEDDAVENKATAGARI

Buffalo, NY • +1 (716) 292-2583 • chakradhar3724@gmail.com • pchakradhar.com • [Linkedin](#) • [Github](#)

EXPERIENCE

Research Assistant

Jan 2026 – Present

University at Buffalo SUNY

- Conducted applied research in multimodal AI for accessibility under Prof. Shamsad Parvin, focusing on real-world assistive applications
- Developed vision-language models and built data pipelines to support accessibility-driven AI systems
- Evaluated model performance using quantitative metrics, driving iterative improvements in accuracy and reliability
- Co-authored and presented a research abstract at the Stony Brook AI Symposium (selected through competitive review)

AI Engineer Intern

Jan 2025 – Jun 2025

Excl Solutions Pvt. Ltd.

- Built and optimized end-to-end ML pipelines for large-scale TEM image analysis, improving image processing throughput by up to 80% through optimized preprocessing and model inference workflows.
- Developed an NLP-based semantic retrieval chatbot using embedding-driven search for efficient dataset and document access.
- Designed modular Python tooling for scalable experimentation and maintainable ML workflows.

Received a **Letter of Recommendation from the CEO** for technical ownership and impact.

PROJECTS

End-to-End CAPTCHA Recognition using CNN-BiLSTM (PyTorch)

- Built an end-to-end CAPTCHA recognition system using deep learning, achieving **99% character accuracy** and **95% full CAPTCHA accuracy** without explicit character segmentation.
- Designed a CNN-BiLSTM model in **PyTorch**, using convolutional layers for feature extraction and bidirectional LSTMs to model sequential character dependencies in noisy images.

RAG-Based Document Question Answering System (LLM + Vector Search)

- Built an end-to-end Retrieval-Augmented Generation (RAG) system for answering questions over PDF documents by integrating semantic search with LLM-based response generation for context-aware outputs.
- Designed a full RAG pipeline using **LangChain**, **FAISS**, and LLM APIs, implementing document parsing, chunking, embedding generation, and vector-based retrieval to provide relevant context for generation.
- Improved response accuracy by optimizing retrieval quality using embedding tuning and reranking techniques, reducing irrelevant context and enhancing answer grounding.

Robust Image Classification under Spurious Correlations

- Identified shortcut learning in the Waterbirds dataset by analyzing group-wise performance and visualizing model attention with Grad-CAM
- Evaluated model robustness by applying ERM, Core-Aligned Regularization, and Group DRO on ViT and DINOv2 features
- Improved worst-group accuracy by 3–4% over ERM, with Group DRO achieving the strongest overall performance

PUBLICATIONS

Neural Sequence-to-Sequence Modeling with Attention for Abstractive Text Summarization

- Designed a neural sequence-to-sequence model with attention for abstractive text summarization. This approach enhances contextual understanding, improving the accuracy of generated summaries.

View Paper: arxiv.org/abs/2404.08685

TECHNICAL SKILLS

- **Languages:** Python, SQL, C++, C, JavaScript
- **AI / ML:** Deep Learning (CNNs, LSTMs, Transformers), NLP, Computer Vision, Predictive Modeling, Model Evaluation, Statistical Analysis
- **Frameworks:** PyTorch, TensorFlow, Keras, OpenCV, Scikit-learn, NumPy, Pandas
- **Tools:** Git, GitHub, Jupyter Notebook, Google Colab
- **Cloud & Systems:** AWS, Firebase, API-driven architectures

EDUCATION

Master of Science in Computer Science and Engineering

Aug 2025 – May 2027

University at Buffalo SUNY, New York

CGPA: 3.917 / 4

Bachelor of Technology in Computer Science with Specialization in Cloud Computing

Jun 2021 – May 2025

SRM University, Chennai, India

CGPA: 9.64 / 10

ACCOMPLISHMENTS

- Merit-Based Scholarships totaling ₹1.2L (2022–2024)
- Stanford Machine Learning Specialization; Google Data Analytics Certification
- HackerRank: 5-Star ratings in Python, C, and C++