(979) 985-7012
College Station, Texas
bhanu@tamu.edu

Venkata **Bhanu** Teja Pallakonda

https://bhanu.cyou

GitHub: pvbhanuteja
LinkedIn: pvbhanuteja

Education

Master of Science in Computer Science, *Texas A&M University, College Station* — Aug 2021 - Present
Bachelor of Technology in Electrical Engineering, *Indian Institute of Technology (IIT) Tirupati* — Aug 2015 - May 2019

Skills & Relevant Specialization

Programmming — Python [Pytorch, TensorFlow, OpenCV, Sklearn, Rasa, FastAPI], Javascript, Reactjs, SQL, Docker.
Specializations — Deep Learning, Pattern Recognition and Machine Learning, Analog Circuits, Computer Vision, Complex Variables, Artificial Intelligence, Calculus, Image Processing, Linear Algebra, Optimization Techniques

Work Experience

Software Engineer, Machine Learning — **Oct. 2020 - Jul. 2021**
Legato Health Technologies (Anthem Inc.) — *Hyderabad, India*

- Developed NLP systems to understand and structure biomedical information, using a combination of pre-trained models like Jasper, GPT-2, and BERT on custom datasets for meeting minutes generation from video recordings.

Machine Learning Engineer — **Jun. 2019 - Oct. 2020**
Fincare Small Finance Bank — *Bangalore, India*

- Engineered an NLP-based banking chatbot using Hugging Face transformer models for intent classification and entity extraction, working with a variety of structured and unstructured data sources.
- Designed and improved core NLP components for ID card detection, field extraction, and field masking to ensure privacy, incorporating the latest research, technologies, and techniques directly into a production environment.

Machine Learning Internship — **May 2022 - Aug 2022**
Productiv — *Seattle, Washington*

- Designed an automated pipeline to parse critical fields from customer contracts, including a document type classifier, a fine-tuned LayoutLMv3 model, and a Label Studio-based annotation pipeline for training data and success evaluation.
- Imagined and implemented creative data-acquisition and labeling systems, employing scalable and novel machine learning pipelines using Airflow on Kubernetes.

Projects

Any-to-Any Voice Conversion using Transformers Link to presentation — **Feb. 2022 - Present**
Texas A&M University — *College Station, Texas*

- Extracted linguistic features and voice identity from utterances, utilizing them independently to achieve versatile voice conversions with transformers for training, resulting in clear synthesized speech and successful voice conversions.

Repaper - Python Package Link to Github — **Oct. 2022 - Nov. 2022**
Open-source Contribution

- Developed a Python package to generate editable PDF forms or online forms from sample form images using LayoutLM and easy-ocr for key-value pair identification and text information extraction.

MixRNet Link to arXiv — **Sep. 2021 - Nov. 2021**
Texas A&M University — *College Station, Texas*

- Applied mixup data augmentation technique for regularization and enhanced ResNet50 architecture accuracy in image classification tasks, achieving an error of **4.87%** on the CIFAR-10 dataset.

UpMail.info Link to Website — **Jun. 2021 - Jul. 2021**
Founder — *Hyderabad, India*

- A real-time email validation API to check deliverability of the email by contacting the SMTP server. Techstack used FASTApi, Docker, NextJs, Supabase, and Stripe.

Cricket Classification Link to Github — **Apr. 2023**
Texas A&M University — *College Station, Texas*

- Fine-tuned a large Audio Spectrogram Transformer model for cricket songs classification using 2*A100 GPUs with PyTorch DDP approach, achieving 97.5% and 94.5% accuracy on 5 and 8 genus classification tasks.
- Given the small dataset, employed audio data augmentation and Voice Activity Detection (VAD) techniques to improve data quality and model performance.

Research Assistant (NSF Funded), Teaching Assistant — **Jan. 2022 - Dec 2022**
Texas A&M University — *College Station, Texas*

- Enhanced pancreatic cancer prediction by reducing features and increasing recall using advanced machine learning techniques, incorporating state-of-the-art techniques like Mixed Precision on Transformer networks.