**Vighneshwar Reddy Katipally**
**Mobile No.:** +91-9333344488
**E-mail: vighnesh.katipally@gmail.com**
**LinkedIn:** https://www.linkedin.com/in/vighneshwar-reddy-katipally-26083815b
**GitHub:** https://github.com/katipallyvig8899
**Portfolio website:** https://main--luxury-daifuku-c79bdd.netlify.app

## Career Objective:

I am a masters graduate in data science looking for an opportunity to contribute in AI/ML domain as a Data Scientist, Data analyst or Machine Learning engineer by developing cutting-edge products.

## Professional Summary:

- Data Science Intern : 4 months

## Technical Skills:

| | |
|---|---|
| **Programming languages** | Python, SQL, C, Java |
| **Technologies** | Machine Learning, Deep Learning, Natural Language processing, Computer vision, Reinforcement learning, Generative AI, Large Language Models, MLOps, Data Analytics |
| **Tools** | Matlab/Simulink, Kubernetes, Jenkins, Docker, Grafana, Kibana, Key cloak, Microsoft Power BI, Excel, Huggingface, Streamlit, Arduino tools |
| **Python Libraries** | Numpy, Pandas, Matplotlib, Scipy, SciKit-Learn, NLTK, Pytorch, Tensorflow, Keras, Opencv, Kivy |
| **Version Control** | Git/Gerrit/DVC |
| **Project management method** | Agile (Jira) |
| **Unit testing** | Pytest |
| **Front end** | HTML/CSS/JS, React |
| **Backend** | Flask, Django, Rest API development |
| **Cloud** | AWS Sagemaker |

## Professional Experience

1. **Innomatics Research Labs**
   - **Location** : Hyderabad
   - **Role** : Data science intern – full time
   - **Duration** : 4 months (Jan 2021 – Apr 2021)
   - **Tools** : streamlit, colab, pandas, sklearn, matplotlib, data visualization

- **Project** **:** Bulding end to end ML models

## Projects:

1. **Project** : Multi Modal Large Language Model (MMLLM) pre-training and fine tuning
   **Languages/ Tools** : Python, Google Colab for training, Huggingface for deployment, Gradio
   **Hugging face link** : https://huggingface.co/spaces/vigraj/MultiMODEL-LLM

   **Project Summary**:
   **Goal** : Building a Multi-Modal Large Language Model for Text generation
   **Inputs** : text, images, and audio
   **Output** : generate textual data
   **Dataset** : COCO-2017 dataset for images, Instruct 150 K dataset
   **Models** : Foundation model (Microsoft Phi-2 LLM), Clip Model, Whisper model
   **Fine tuning** : Qlora

   **Project Description**:

   - This project involves a three-stage process to create a Multi-Modal Large Language Model.
   - **Stage 0**: **Pretraining**
     - Foundation Model: Microsoft Phi-2 LLM (Text input and Text output)
     - Clip model as Image encoder for getting embeddings from input images (COCO-2017 dataset). Then we trained the projection layer and projection model to convert above embeddings to be compatible with the Microsoft Phi-2 model.
   - **Stage 1**: **Finetuning**
     - Performed fine-tuning on projection layer, projection model and Phi-2 model using Instruct 150 K dataset, enabling the model to understand conversations from images.
     - Adopted the Qlora fine-tuning strategy, we optimize the phi-2 model for multi-modal tasks, enhancing its ability to process text, images and audio.
   - **Stage 2: Deployment**
     - Utilized whisper model for audio data
     - Deployed in Hugging face

2. **Project** : Microsoft Phi-2 Based AI Assistant using Qlora strategy
   **Languages/Tools** : Python, Google Colab for training, Huggingface for deployment, Gradio
   **Huggingface link**: https://huggingface.co/spaces/vigraj/ChatbotwithQlora

   **Project Summary:**
   **Goal** : AI Assistant / Chatbot
   **Inputs** : text
   **Output** : generate textual data
   **Dataset** : Open Assistant dataset

   **Models** : Foundation model (Microsoft Phi-2 LLM)

**Fine tuning** : Qlora

3. **Project** : Handwritten Character Recognition using Deep Learning
   **Languages/Tools** : Python, Tensorflow, Keras

**Project Summary:**

**Goal** : Recognizing the hand-written characters
**Inputs** : hand-written text
**Output** : generated text
**Dataset** : IAM Dataset
**Models** : CNN's, Bi-LSTM

## Technical Publications

**Publication Name : 2023 3rd International Conference on Intelligent Technologies (CONIT)**

**Title : Sentiment Analysis at Document Level of Telugu data from Multi-Domains**

**Publication Link: https://ieeexplore.ieee.org/document/10205691**

## Academics

| Degree | Specialization | College | Duration | Percentage/CGPA |
|---|---|---|---|---|
| Master of Technology | Data Science | Center of Excellence in Networking (CEN), Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India. 641112 | 2021-2023 | 7.0/10 |
| PG Diploma | Artificial Intelligence and ML | E & ICT Academy NIT Warangal | 2020-2021 | 3.43/4.0 |
| Bachelor of Technology | Computer Science Engineering | Vardhaman College of Engineering,Shamshabad, Hyderabad,Telangana,India. | 2016-2020 | 7.1/10 |

| | | | | |
|---|---|---|---|---|
| Intermediate | MPC | Sri gayatri junior college, Hyderabad, Telangana, India. | 2014-2016 | 74% |
| High School | NA | Gautham High School, Hyderabad, Telangana, India | 2014 | 8.2/10 |

## Continuous Learning / Certifications

1. iNeuron – Full Stack Data Science Program (1 Year)
2. NPTEL - CERTIFICATION ON MACHINE LEARNING (2 months)
3. Data camp - Python Developer (2 months)
4. The school of AI – Extensive & Reimagined AI (ERA-V1) Program 2023 (9 months)

## Personal Details:

Full Name : K a t i p a l l y V i g h n e s h w a r R e d d y
Date of birth : 3 0 / 08/ 1999
Gender : M a l e
Blood group : O +
Languages Known : English, Telugu, Hindi