KARTHIK METTU

| Houston, TX  |  [karthikrm202002@gmail.com](mailto:kmettu@cougarnet.uh.edu,%20mettukarthik123@gmail.com,%20karthikrm202002@gmail.com)  |  +1 713 202 0004  | [www.linkedin.com/in/karthik-mettu](https://www.linkedin.com/in/karthik-mettu/) |

**education**

**Masters in Statistics and Data Science** December 2024

University of Houston, College of Natural Sciences and Mathematics Houston, Texas

* Cumulative GPA: 3.7/4.0
* Coursework: Machine Learning, Deep Learning, Probability, Statistics, Data Visualization, Big Data Analytics

**Bachelor’s in Mechanical Engineering, Minors in Computer Science and Engineering** May 2023

Vasireddy Venkatadri Institute of Technology Guntur, India

* Cumulative GPA: 3.5/4.0, Computer Science CGPA: 3.5/4.0, Dean’s List

**experience**

**Instructional Assistant** August 2023 – December 2024

University of Houston Law Center - Information Technology Houston, Texas

* Conducted data reconciliation and integrity checks by comparing academic and administrative records, utilizing advanced Excel and database systems to ensure report accuracy.
* Develop and oversee instructional and administrative reports, leveraging Power BI dashboards and data analytics to enhance strategic decision-making.
* Collaborate with internal departments and external vendors to streamline data management processes and support educational and HR systems like PeopleSoft, ensuring compliance and maintaining data security.

**GenAI Engineer Volunteer** May 2024 – Present

American Red Cross Houston, Texas

* Developed a predictive model for receipt classification using Azure AI Studio, PromptFlow, and the RAG framework, achieving 90% accuracy in purchase classification, resulting in significant cost savings.
* Enhanced CPR manual content with profession-specific GenAI images using Azure AI Studio and Machine Learning, improving reader comprehension by 30% and enhancing training effectiveness.

**Data Science Intern** January 2024 – May 2024

University of Houston - Department of Biology and Biochemistry Houston, Texas

* Analyzed over 1,000 tRNA sequences from the UCSC Genome Browser and processed 100GB of genomic data, increasing detection of high-abundance genomes by 20% and ensuring data integrity.
* Applied machine learning algorithms (Random Forest, SVM) with scikit-learn to tRNA sequences, achieving 85% accuracy and 80% recall, uncovering key patterns for scientific research.

**Data Science Research Intern** May 2024 – August 2024

University of Houston - Department of Chemistry Houston, Texas

* Utilized PyMol, ChimeraX, and AlphaFold 2 to visualize and modify over 8,000 protein structures, resulting in a 25% increase in protein modeling accuracy and efficiently handling PDB files of 50GB data.
* Collaborated with faculty to visualize interactions among 2,000 distinct proteins using Python scripting, enhancing the comprehension of protein-protein interaction structures by 30% and improving research outcome.

**skills**

* **Programming & Data Science:** Python, R, Java, C++, C, MATLAB, Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, SpaCy, NLTK, Seaborn, NLP, Transformers, FAISS, Knowledge Graphs, SPSS, SAS
* **Big Data & Cloud Computing:** PySpark, MySQL, AWS (EC2, S3, SageMaker, Athena, Glue), Azure (AI Studio, Prompt Flow, SQL Server, Synapse Analytics), GCP (Vertex AI), Docker, Git, Kubernetes, Data Wrangling, Data Mining, A/B Testing, SQL, NoSQL, Data Lakes, ETL, Microsoft SQL Server Suite (TSQL, SSRS, SSAS)
* **Data Visualization and Bioinformatics:** Tableau, Power BI, Excel, PyMol, ChimeraX, AlphaFold 2

**PRojects**

**Comparative Analysis of Bidirectional LSTM and Attention Models for NLP Tasks** March 2024

* Conducted Twitter sentiment analysis on 13 emotions in 420,000 tweets, showing Transformer achieved 99.8% accuracy in 14 minutes, outperforming LSTM's 99.2% in 60 minutes, providing insights for NLP selection.

**Multivariate Analysis of Global Suicide Rates data from 1990 to 2022**  February 2024

* Performed time series analysis on over 10,000 global suicide records from 1990 to 2022, using ARIMA and LSTM models to identify trends and develop forecasts with a MAPE of 2.5%, providing actionable insights for public health strategies.

**Visualization of Microplastic Pollution and Health Impacts Using Tableau** January 2024

* Executed Tableau project analyzing wealth, health, and pollution correlations using data from the World Bank, WHO, and World Inequality Database, creating dashboards that improved insights and decision-making by 30%.