

# Vikas Reddy

## Data Scientist – Production ML, Real-Time Analytics & Scalable Systems

(443)-992-2554 | [wh22215@umbc.edu](mailto:wh22215@umbc.edu) | [linkedin.com/in/vvikasreddy](https://www.linkedin.com/in/vvikasreddy) | [github.com/vvikasreddy](https://github.com/vvikasreddy) | [Portfolio](#)

### EDUCATION

#### University of Maryland Baltimore County (UMBC)

*Master's in Data Science; GPA: 4.0*

Baltimore, MD

Aug 2023 – May 2025

#### Institute of Aeronautical Engineering (IARE)

*Bachelor's of Technology in Computer Science and Engineering; GPA: 3.41*

Hyderabad, India

May 2019 – May 2023

### RESEARCH PUBLICATIONS

Deep fusion of neurophysiological and facial features for enhanced emotion detection. (n.d.). IEEE Journals & Magazine — IEEE Xplore. <https://ieeexplore.ieee.org/abstract/document/10945364>

### EXPERIENCE

#### Brain-Machine Interface Lab at UMBC

Baltimore, USA

*ML Researcher, Real-Time Analytics & Scalable Data Products*

Jan. 2024 – Present

- Designed and deployed a real-time ML platform integrating **Transformer** and **ConvLSTM** models, achieving 97% accuracy and enabling actionable insights for end-user decision-making.
- Developed **EmoFormer**, a production-grade pipeline using **SegFormer** and **Scikit-learn**, surpassing emotion detection benchmarks—modeling prescriptive outcomes for adaptive systems.
- Built a high-throughput streaming framework using **Apache Kafka** and multithreaded **CNN inference**, cutting latency by 35% to power signal-driven applications in real time.
- Engineered scalable data ingestion and feature pipelines supporting predictive analytics and robust metric tracking across complex biomedical data streams.

#### Epam Systems

Hyderabad, India

*Software Development Intern – ML Ops & Data Infrastructure*

Jan. 2023 – June 2023

- Built scalable microservices and reusable backend components using **Java**, **Python**, **Kubernetes**, and **Docker**, supporting customer-facing analytics with enterprise reliability.
- Automated release and testing workflows via **CI/CD pipelines** (**Jenkins**, **Git**, **Docker**), reducing product delivery timelines by 35% across multi-team operations.
- Optimized cloud-native data processing on **AWS Lambda**, **S3**, and **Athena**, managing 50GB+ structured/unstructured datasets used for analytical dashboarding and reporting.
- Collaborated with QA and data teams to monitor system KPIs and drive continuous performance improvements across test analytics and pipeline deployment.

#### IndicWiki Project

Hyderabad, India

*Data Analyst Intern – ETL, SQL & NLP Systems*

Mar. 2022 – June 2022

- Designed and deployed ETL pipelines in **Python** and **SQL** for multilingual data processing, enabling clean downstream modeling for business insights and NLP analysis.
- Delivered a high-integrity data processing system with over 98% consistency across unstructured datasets, aligned with operational and reporting needs.
- Created scalable backend scripts and data validation logic for cross-functional team usage—enhancing project delivery velocity and stakeholder alignment.

### PROJECTS

#### Stock Price Forecasting System | *Python, TensorFlow, LSTM, NumPy, Matplotlib*

GitHub

- Developed a time-series forecasting model using **LSTM** to predict stock prices, achieving 84% of predictions within  $\pm 10\%$  error. Applied data normalization, cross-validation, and hyperparameter tuning to boost model accuracy.

#### Lexically Constrained Neural Translation Engine | *Python, PyTorch, Hugging Face Transformers*

GitHub

- Built a custom **beam search decoder** to enforce lexical constraints during transformer-based neural translation. Improved **BLEU scores** on WMT TR-EN data by guiding model outputs across 300+ test scenarios.

### TECHNICAL SKILLS

**Programming Languages:** Python, SQL, Java, C++, Bash, R

**Machine Learning & Analytics:** Scikit-learn, TensorFlow, PyTorch, Transformers, NumPy, Pandas, PySpark

**Data Engineering & Pipelines:** ETL Pipelines, Apache Kafka, Real-Time Streaming, Data Modeling, CI/CD

**Cloud & DevOps:** AWS (Lambda, S3, Athena, Redshift), Docker, Kubernetes, Jenkins, Git

**Visualization & Databases:** Tableau, PostgreSQL, Oracle, SQL Server, Jupyter, Matplotlib