

# BALA GOPAL REDDY PEDDIREDDY

✉ [balagopalreddypeddireddy@gmail.com](mailto:balagopalreddypeddireddy@gmail.com) ☎ +1 513-306-3253 🔗 [linkedin.com/in/bala-gopal-reddy-peddireddy/](https://www.linkedin.com/in/bala-gopal-reddy-peddireddy/)



📄 [balupeddireddy08.github.io/myResume/](https://balupeddireddy08.github.io/myResume/) 📖 [medium.com/@balupeddireddy08/](https://medium.com/@balupeddireddy08/) 🏠 [github.com/balupeddireddy08/](https://github.com/balupeddireddy08/) 📺 [www.youtube.com/@unleashingtomorrowstech](https://www.youtube.com/@unleashingtomorrowstech)

## SUMMARY

Data Scientist with 2+ years of experience in machine learning, deep learning, NLP, and GenAI solutions. Skilled in statistical modeling, time series forecasting, and developing AI-driven models for real-world applications.

## EDUCATION

**University of Cincinnati**, Master of Engineering in Computer Science **CGPA: 3.62** Aug 2023 – Apr 2025  
*Coursework: Cloud Computing, Advanced Machine Learning, Data Encoding, Deep Learning*  
**VNR Vignana Jyothi Institute of Technology**, BTech in Electronics and Communication Engineering **CGPA: 3.64** Aug 2017 – May 2021  
*Coursework: Artificial Neural Networks and Fuzzy Logic, Data Structures, DBMS, Web Development*

## SKILLS

- **Programming:** Python, SQL, C++, C, Java
- **Frameworks:** TensorFlow, PyTorch, Keras, Scikit-learn, Hugging Face, LangChain, Transformers, Apache MXNet, FAISS, OpenAI
- **Tools:** Jupyter Notebook, GitHub, Docker, Tableau, Power BI, FastAPI, MLflow, Data Version Control, DataRobot, Streamlit, Microsoft Office
- **Databases & Cloud Technologies:** MySQL, MongoDB, Azure (Machine Learning, Blob, Data Factory), AWS (S3, Lambda, EC2)

## EXPERIENCE

**Tata Consultancy Services Ltd** Oct 2021 – Sep 2023

**Data Scientist** [Python, Machine Learning, Data Science, Statistics, Matplotlib, Scikit-Learn, Statsmodels, Plotly, TensorFlow, OpenCV]

- Developed a machine learning pipeline utilizing advanced deep learning and time-series analysis techniques to assess gearbox health from sensor data with a sampling rate of **3000**.
- Created a predictive forecasting model leveraging the Auto\_Arima, TimeGPT, and Temporal Fusion Transformers enabling **78%** accurate prediction of future plant stoppages based on historical data patterns and trends.
- Achieved **>85%** accuracy in particle size analysis for cone crushers using a novel image processing system with a modified DexiNed network. This system leverages edge detection, contour refinement, and mass estimation techniques.

**Cognizant**

Mar 2021 – July 2021

**Programmer Analyst Intern** [Python, Apache Spark, SQL, Tableau, Hadoop, HDFS]

- Created and executed ETL processes with SSIS, developed SQL transactional queries for enhancing data reliability.
- Combined Apache Spark with Hadoop to manage large-scale data processing more efficiently. Produced actionable, client-facing reports using Tableau, meeting intricate business needs and aiding stakeholder decision-making.

## PROJECTS

**Advanced RAG Question & Answering System** [Python, LangChain, LLM, RAG, FAISS, Llama, HuggingFaceEmbeddings, AI]

- Built a question-answering system using LangChain, RAG, and Llama **8B** Instruct Model, boosting accuracy by **40%** and reducing manual effort by **50%**.
- Implemented data preprocessing with WebBaseLoader, CharacterTextSplitter, and PdfReader, and developed a Streamlit frontend, enhancing productivity by **45%**.

**IntelliNews AI Agent** [CrewAI, Google Gemini Pro LLM, Python, NLP, AI]

- Built News AI agents using CrewAI and Google Gemini Pro LLM models to deliver personalized news content, increasing recommendation accuracy by **30%**.
- Utilized advanced NLP techniques to enhance user interactions, leading to a **20%** boost in user engagement.

**AI-Powered Multimodal Voice Assistant** [Python, Llava, Whisper, Multimodal LLM, AI]

- Developed a multimodal AI voice assistant app using Llava and Whisper, enhancing voice recognition and natural language processing capabilities.
- Integrated multimodal LLM for seamless voice and text interactions, improving user engagement and response accuracy by **50%**.



**Fake News Classification Using LSTM** [Python, TensorFlow, Keras, NLTK, Scikit-learn, Seaborn, Pandas, NumPy, LSTM, Pytorch]

- Executed comprehensive data cleaning and preprocessing, including stop-word removal, lemmatization, and n-gram analysis, improving text data quality by **30%**.
- Implemented an LSTM model with one-hot encoding and pre-padding, achieving **92%** classification accuracy in detecting fake news.

**Detection Of Alzheimer's Disease** [Flask, Random Forest, Gradient Boosting, LightGBM, Logistic Regression, Decision Tree]

- Performed data preprocessing, and applied various machine learning models (Random Forest, Gradient Boosting, Adaptive Boosting, Logistic Regression, Decision Tree), achieving a precision of **92%** and recall of **89%** through RandomizedSearchCV hyperparameter tuning.
- Crafted and integrated an interactive website using HTML, CSS, and Bootstrap with the machine learning model via the Flask framework, enhancing user accessibility and engagement.

## CERTIFICATIONS

- Machine Learning Certificate - Course authorized by Stanford University. 
- Tableau Desktop Certified Associate - Course offered by Edureka. 

## ACHIEVEMENTS

- **Received** - Graduate Incentive Award from the University of Cincinnati, valued at **\$9,226**, in recognition of academic excellence. *Aug 2023*
- **Achieved** - Elite and Silver Medals in Programming, Data Structures, and Algorithms Using Python through NPTEL Exam. *Sep 2019*