

BALA GOPAL REDDY PEDDIREDDY

📍 Cincinnati, Ohio 45220 ✉ 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://medium.com/@balupeddireddy08> 🐙 [github.com/balupeddireddy08/](https://github.com/balupeddireddy08)

CAREER OBJECTIVE

Experienced Data Scientist with a proven track record in machine learning, Python, and statistical analysis. Skilled in data preprocessing, visualization, and analytical acumen, driving success in interdisciplinary teams. I excel in optimizing data processes and enhancing business outcomes. My adaptability and versatility make me an asset in the realm of Computer and Data Science.

EDUCATION

Master of Engineering in Computer Science	Aug 2023 – Dec 2024 (Expected)
University of Cincinnati, Cincinnati, Ohio	3.83 GPA
Bachelor of Technology in Electronics and Communication Engineering	Aug 2017 – Aug 2021
VNR Vignana Jyothi Institute of Technology, Hyderabad, India	3.64 GPA

PROFESSIONAL EXPERIENCE

Data Scientist, Tata Consultancy Services Ltd - Hyderabad, India	Oct 2021 - Aug 2023
(High Frequency Vibration Analysis – Holcim Client)	

- Preprocessed sensor data, including time-series formatting, exploratory data analysis (EDA), seasonal decomposition, and stationarity checks.
- Generated insightful visualizations, including time-waveforms, frequency spectra, and overall trends, through time resampling, data smoothing with Exponentially Weighted Moving Average (EWMA), and DBScan clustering.
- Developed a centralized Python code to compute statistical parameters like RMS, peak, and kurtosis for the sensor data.
- Implemented an ARIMA forecasting model to predict future plant stoppages.
- Utilized libraries such as Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, Statsmodels, Plotly, and TensorFlow.

Analyst – Intern, Cognizant, Hyderabad, India	Mar 2021 - July 2021
<ul style="list-style-type: none">• Acquired proficiency in Java, HTML, CSS, and Bootstrap through practical hands-on experience and training.• Enhanced interpersonal skills through active participation in collaborative team activities.	

TECHNICAL SKILLS

- **Languages:** Python, C Language, C++, Java
- **Web Programming Languages:** CSS 3, HTML 5, Bootstrap
- **Database:** MySQL
- **IDE & Tools:** Jupyter Notebook, Visual Studio Code, GitHub, Matlab, Tableau, Excel, Word, PowerPoint
- **Cloud Platform:** Azure, AWS (S3, EC2, RDS)
- **Technologies:** Data Science, Statistics, Machine Learning, Time Series Analysis, Deep Learning, NLP

ACADEMIC PROJECTS

Detection Of Alzheimer's Disease Using Machine Learning

- Performed data preprocessing, addressing missing values, label encoding, and standardization. Implemented dataset split and hyperparameter tuning with RandomizedSearchCV.
- Utilized a variety of machine learning models, including ensemble algorithms (Random Forest, Gradient Boosting, Adaptive Boosting), Logistic Regression, and Decision Tree models and evaluated model performance by considering precision and recall as key evaluation.
- Designed an interactive website with HTML, CSS, and Bootstrap and integrated it with the machine learning model using the Flask framework.

Automatic Number Plate Detection Using Yolo and Pytesseract

- Used TensorFlow for image tasks, including processing, resizing, and converting to arrays, followed by normalization.
- Located number plate center using bounding box diagonal coordinates, improving YOLOv5 model accuracy with pre-trained weights.
- Set confidence and probability thresholds, executed non-maximum suppression, and extracted Region of Interest from detected bounding boxes.
- Implemented pre-trained PyTesseract OCR model to extract text info from Number Plate Detection project's Region of Interest (ROI).

Netflix Movie Recommendation System

- Calculated User-User and Movie-Movie Similarity matrices using cosine similarity and engineered new features, such as similar user ratings for the same movie, similar movie ratings by the same user.
- Utilized machine learning models like XGBRegressor, Surprise BaselineModel, and KNNBaseline, along with Matrix Factorization (SVD, SVD++), for recommendations. Implemented feature stacking to boost accuracy in recommendation systems.
- Considered RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error) as evaluation metrics to assess the model's performance.

CERTIFICATIONS

- PCAP: Certified Associate in Python Programming Certification - Course offered by Python Institute. [↗](#)
- Machine Learning Certificate - Course authorized by Stanford University and offered through Coursera. [↗](#)
- Python for Time Series Data Analysis - Course offered by Udemy. [↗](#)
- The Complete SQL Bootcamp: Go from Zero to Hero - Course offered by Udemy. [↗](#)
- Tableau Desktop Certified Associate - Course offered by Edureka. [↗](#)
- Microsoft Certified: Azure Data Fundamentals Certification (DP-900) - Course offered by Microsoft. [↗](#)
- Microsoft Certified: Azure Fundamentals Certification (AZ-900) - Course offered by Microsoft. [↗](#)
- Microsoft Certified: Azure AI Fundamentals Certification (AI-900) - Course offered by Microsoft. [↗](#)
- Microsoft Certified: Azure Data Scientist Associate Certification (DP-100) - Course offered by Microsoft. [↗](#)
- HTML, CSS3 & Bootstrap Certification – Course offered by Udemy. [↗](#)

PUBLICATIONS

- A Short Journey to Deep Learning [↗](#)
- Forecast The Future with Time Series Analysis [↗](#)
- Deploy a Static Website using GitHub [↗](#)
- Get Rid of Web Development's Phobia – Part1 [↗](#)
- Create a Virtual Environment in Windows OS and run a flask application Locally [↗](#)

PRESENTATIONS

Peddireddy B. 2022. Detection & Handling of Outliers Presentation. Presenter. ScaleAI-100 Meeting (Virtual). 05 January

- I have presented about outliers, covering their definition, causes, and detection techniques, such as Hypothesis Testing, Z-score, Robust Z-score, I.Q.R, Winsorization, DBSCAN clustering, and Isolation Forest.
- I have discussed strategies for handling outliers, including deleting observations, transforming values, imputation, and separately treating outliers to ensure effective data management.