Bala Gopal Reddy Peddireddy

Cincinnati, Ohio 45220 balagopalreddypeddireddy@gmail.com +1 513-306-3253
linkedin.com/in/bala-gopal-reddy-peddireddy/ balupeddireddy08.github.io/myResume/
medium.com/@balupeddireddy08/ github.com/balupeddireddy08/

Summary

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.

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
- **Cloud Platform:** Azure, AWS (S3, EC2, RDS)
- **Technologies:** Data Science, Statistics, Machine Learning, Time Series Analysis, Deep Learning **(Work Exp)**

Work Experience

TATA Consultancy Services Ltd– Hyderabad, India 10/2021 – 08/2023
Project: High Frequency Vibrational Analysis – Holcim
Data Scientist

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

Cognizant – Hyderabad, India 03/2021 – 07/2021
Analyst – Intern

- 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.

Academic Projects

Detection Of Alzheimer's Disease Using Machine Learning

- Developed a machine learning model for early-stage Alzheimer's disease detection using the Oasis Longitudinal dataset. Conducted exploratory data analysis (EDA) to understand the dataset.
- Performed data preprocessing, which involved handling missing values, label encoding, and standardization to prepare it for model training.
- Split the dataset into training and testing sets and conducted hyperparameter tuning for each model using RandomizedSearchCV to optimize model performance.
- Utilized a variety of machine learning models, including ensemble algorithms (Random Forest, Gradient Boosting, Adaptive Boosting), Logistic Regression, and Decision Tree models to enhance predictive accuracy.

- Evaluated model performance by considering precision and recall as key evaluation metrics to ensure robust Alzheimer's disease detection.
- Designed an interactive website with HTML, CSS, and Bootstrap to provide a user-friendly interface for users.
- Integrated the website with the machine learning model using the Flask framework, allowing users to interactively test the Alzheimer's disease prediction model in a web-based environment.

Netflix Movie Recommendation System

- Developed a movie recommendation system for Netflix using various data preprocessing and machine learning techniques.
- Merged multiple data files to create a comprehensive dataset, handled missing values (NaN), duplicates, and computed the distribution of ratings.
- Calculated User-User Similarity matrix and Movie-Movie Similarity matrix using cosine similarity.
- Engineered new features, such as similar user ratings for the same movie, similar movie ratings by the same user, average movie ratings, and average user ratings.
- Modified the data to work seamlessly with Surprise-based algorithms from Surprise library.
- Employed machine learning models, including XGBRegressor, Surprise BaselineModel, and Surprise KNNBaseline predictor, to find the best recommendation model.
- Utilized Matrix Factorization-based algorithms, such as SVD and SVD++, for recommendation.
- Implemented feature stacking by using the output of previous models as features for subsequent models, enhancing recommendation accuracy.
- Considered RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error) as evaluation metrics to assess the model's performance.

Education

Master of Engineering, Computer Science Graduating: December 2024

University of Cincinnati, Cincinnati, Ohio

Bachelor of Technology, Electronics & Communication Engineering July 2017 – June 2021

VNR Vignana Jyothi Institute of Engineering & Technology — Hyderabad, India CGPA – 8.88

Certifications

- PCAP: Certified Associate in Python Programming Certification - Course offered by Python Institute.
- Machine Learning Certification - 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.

Articles

- 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

Accolades

- Achieved Graduate Incentive Award (GIA) from University of Cincinnati worth $9,226 based on academic excellence.
- Presented a seminar on “Detection & Handling of Outliers” for the associates at TCS.
- Achieved an Individual Performance Factor (IPF) greater than 4 out of 5, indicating outstanding performance, in two consecutive appraisal cycles at TCS.
- Achieved Gold Medal for academic excellence from VNR Vignana Jyothi Institute of Engineering & Technology for batch 2017 - 2021.
- Secured an Elite and Silver Medal in Programming, Data Structures and Algorithms Using Python by NPTEL Online Certification.