Dhanya Bahadur

Data Scientist

+91-932-650-1398 /Dhanyabahadur dhanyabahadur2000@gmail.com
Bihar, India Dhanya Bahadur /in/dhanya-bahadur

Experienced data scientist skilled in advanced analytics, machine learning, and predictive modeling. Proficient in Python, creating accurate models within tight timelines. Passionate about problem-solving, optimizing outcomes, and fostering a collaborative learning environment by sharing knowledge and expertise in ML algorithms and Python coding.

Education

IIT Bombay 2017 – 2021
Bachelor of Science, Chemistry *Mumbai, India*

Chinmaya Vidyalaya 2015 – 2017
Intermediate *Bokaro, Jharkhand*

Work Experience

AllCargo Logistics Oct'22 – Nov'23
Assistant Manager - Data Science *Mumbai*

» **Warehouse Booking Demand Forecasting**
› Implemented automated Demand Forecasting solution using advanced Machine Learning algorithms at multiple granularity levels
› Engineered insightful features by leveraging COVID wave timelines, monthly/weekly/daily patterns, and public/company holidays
› Deployed monthly models on 441 lanes, achieving a remarkable 92% accuracy and optimizing forecasting while improving efficiency

» **Email Automation**
› Leading a LLM-based initiative to automate responses, streamlining email management and enhancing customer experience

Guavus Jul'21 – Sep'22
Data Scientist *Gurgaon*

» **Manufacturing Site Production Ready Date Prediction**
› Attained comprehensive understanding of business context and requirements through interviews and user story development
› Performed data cleansing and engineered features by employing techniques such as target-guided encoding, binning, and others
› Developed robust three separate models for each constituent process after identifying key trends and factors impacting lead time
› Reduced lead time inaccuracies from an average of 25 days to 12 days by leveraging ML techniques, improving prediction accuracy

» **Customer Churn Prediction Analysis**
› Constructed a dimensionality-reduced Random Forest classification model using Autoencoders for enhanced accuracy

» **Participated in the Guavus Ideathon - Supply Chain Suite**
› Collaborated within a dynamic team to develop and implement a strategy addressing the growth challenges in the Supply Chain
› Engineered a CNN model for defect/non-defect forecasting in finished goods, achieving 93% accuracy, high precision, and recall
› Enhanced order execution by creating a ML model to estimate shipping time between ports, enabling smarter decision-making

Deterministic Algorithms Lab May'20 – Jul'20
Data Analyst *Remote*

» **Auto-Dubbing Research Paper Implementation Team**
› Generated talking faces for auto-dubbing by synchronizing images with a lip-sync rate of 3.89, ensuring seamless alignment
› Performed analysis on outcomes and conducted literature review on synthesizing speech from lip movements, contributing insights

Projects

Academical: Self Driving Cars | Research Project May'19 – Aug'19
» Department of System and Controls, IIT Bombay
› Acquired deep expertise in Q-Learning and leveraged it to develop an Advanced Overtaking Policy, seamlessly integrated with a robust Behaviour-Based Architecture, resulting in an impressive success rate for a complete driver on the TORCS simulator
› Delivered comprehensive technical briefings, documenting progress, accomplishments, and challenges affecting the task

Personal: Projects Sep'22 – Oct'22
› Utilized Machine Learning to develop a system for detecting financial fraud, achieving a remarkable 98% AUPRC score
› Evaluated and compared machine learning algorithms for Android malware detection, employing techniques to enhance performance and engineer top 10 features. Achieved 92% accuracy with high precision and recall, demonstrating effectiveness

Skills

Software Packages/Library Numpy, Pandas, Matplotlib, Scipy, Sklearn, HyperOpt, Optuna, TensorFlow, PyTorch, Word2Vec, ChatGPT
Programming Languages Python, C++
Miscellaneous Feature Engineering (Encoding of rare categorical labels, Count encoding, Ordered Integer encoding, Mean Encoding), Feature Selection (Recursive Feature Elimination), Hyperparameter Tuning (GridSearchCV, RandomizedSearchCV, Bayesian Optimization), Machine Learning Algorithms (Linear Regression, Random Forest, K-means clustering, Logistic Regression, XGBoost, Light GBM, Neural Networks), Time Series Forecasting, Prompt Engineering