PRABHAT DHAR

Data Analyst | Python | SQL | Power BI | Machine Learning | MLOps | DevOps

London, UK prabhatdhar32@gmail.com LinkedIn GitHub

Portfolio

PROFESSIONAL SUMMARY

Data Analyst with expertise in transforming raw data into actionable insights through Python, SQL, and advanced visualization tools. Strong background in implementing end-to-end ML pipelines, CI/CD automation, and containerized deployments. Committed to leveraging data science, MLOps, and DevOps practices to drive data-driven decision-making.

EDUCATION

PG Diploma in Data Science and its Application *November 2022*
University of Essex, Colchester, United Kingdom 165 Credits

Bachelor's in Physics *August 2021*
University of Calcutta, West Bengal, India 7.240 CGPA

I.S.C (12th) *April 2017*
St. Xavier's Institution, Kolkata, India 85%

I.C.S.E (10th) *April 2015*
St. Xavier's Institution, Kolkata, India 93.8%

TECHNICAL SKILLS

Programming: Python, SQL, JavaScript, HTML, CSS, LaTeX

Data Analysis: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Power BI, Tableau, Excel

Machine Learning: Scikit-learn, TensorFlow, PyTorch, PyCaret, Statsmodels, OpenAI GPT-4

DevOps & MLOps: Docker, Kubernetes, Jenkins, CI/CD, Git, GitHub

Web Development: FastAPI, Streamlit, Flask, Uvicorn

Tools: Jupyter, Google Colab, VSCode, Linux/Unix

KEY PROJECTS

Health and Fitness AI App

GitHub: Xclipxz07/Health-and-Fitness-AI-App

- Developed AI-powered health and fitness application providing personalized recommendations, fitness tracking, and wellness insights based on user health data and behavioral patterns
- Integrated ML algorithms to analyze workout history, nutrition, sleep patterns, and goals, delivering customized fitness plans tailored to individual needs
- Utilized neural networks and recommendation systems to predict fitness outcomes, identify health trends, and suggest optimal exercise routines
- Built backend with Python leveraging Scikit-learn, TensorFlow, and PyTorch for robust predictive modeling and pattern recognition
- Created interactive dashboards with Matplotlib, Seaborn, and Plotly displaying user progress, fitness metrics, and personalized insights
- Implemented Streamlit web framework for real-time health data input, recommendations, and AI-powered fitness guidance
- **Tech Stack:** Python, Scikit-learn, TensorFlow, PyTorch, Pandas, NumPy, Streamlit, Matplotlib, Seaborn, Plotly

Rainfall Prediction Using ML Models (Dissertation Project)

GitHub: Xclipxz07/Rainfall-prediction

- Developed ML dissertation project forecasting rainfall based on historical Australian weather patterns
- Conducted comprehensive EDA to understand weather patterns, correlations, seasonal trends; performed data cleaning including handling missing values and outliers
- Implemented multiple algorithms: Logistic Regression, Random Forest, Gradient Boosting, Neural Networks to identify most effective predictor
- Leveraged PyCaret AutoML library to automate workflow including feature engineering, hyperparameter tuning, and model selection
- Performed comprehensive evaluation using accuracy, precision, recall, F1-score, ROC-AUC curves, and confusion matrices
- Created interactive visualizations using Matplotlib, Seaborn, and Plotly to illustrate patterns, model performance, and predictions
- Developed detailed dissertation using LaTeX with literature review, methodology, results analysis, and conclusions
- **Tech Stack:** Python, PyCaret, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Plotly, Statsmodels, LaTeX

Data Analysis for DevOps, MLOps & GenAI

GitHub: Xclipxz07/DATA-ANALYSIS-FOR-DEVOPS-MLOPS-GENAI

- Built production-ready sales forecasting system combining ML, generative AI, and DevOps practices
- Implemented end-to-end ML pipeline with SQLite database, Jupyter Notebook, Linear Regression, and OpenAI GPT-4 integration
- Developed scalable REST API using FastAPI with endpoints (`/sales`, `/predict`) and auto-generated documentation

- Containerized application using Docker enabling consistent deployment across environments and production scalability
- Orchestrated deployment using Kubernetes with configurations supporting horizontal scaling, load balancing, and high availability
- Implemented complete CI/CD pipeline using Jenkins for automated testing, building Docker images, and Kubernetes deployment
- Created interactive visualizations using Plotly for sales trend analysis and Power BI dashboards for business intelligence
- Developed technical documentation using LaTeX covering architecture, ML methodology, API specs, deployment procedures
- **Tech Stack:** Python, Jupyter, SQLite, Scikit-learn, OpenAI GPT-4, FastAPI, Plotly, Power BI, Docker, Kubernetes, Jenkins, LaTeX

CORE COMPETENCIES

- Machine Learning & Predictive Modeling
- Data Cleaning & Preprocessing
- Feature Engineering & Selection
- Exploratory Data Analysis (EDA)
- Data Visualization & Storytelling
- Business Intelligence Dashboards
- SQL Query Optimization
- API Development & Integration
- Containerization (Docker)
- Orchestration (Kubernetes)
- CI/CD Pipeline Automation
- Technical Documentation
- Problem Solving & Critical Thinking
- Version Control (Git/GitHub)

References available upon request