Pawan S Bant

Data Scientist

☎ +91-8762070702 ✉ bantpawan@gmail.com ○ GitHub in LinkedIn

COURSEWORK / SKILLS

- Data Science and Machine Learning (Supervised, Unsupervised, and Deep Learning)
- Big Data Processing (PySpark, Delta Lake)
- MLOps (Azure ML, Docker, CI/CD Pipelines)
- Database Management System (DBMS) - SQL
- Power BI
- Object-Oriented Programming (OOP)
- Python
- Advanced Excel & VBA
- Cloud Platforms (Azure)

WORK EXPERIENCE/ PROJECTS

Taxi Fare Prediction Application ⌇ | PySpark, Machine Learning, Streamlit, Azure   **2023-24**

- Built an end-to-end ML application for NYC taxi fare prediction on a 35M-row dataset, leveraging advanced data cleaning and feature engineering to enhanced performance by 13%.
- Applied advanced feature engineering and selection (e.g., Random Forest, Gradient Boosted Trees, SHAP), assessed various regression models, and fine-tuned to improve accuracy by 15%.
- Developed a personalized prediction system with interactive visualizations, boosting user engagement by 25%.
- Fine-tuned regression models with advanced hyperparameter optimization on Databricks, reducing response time by 30%, and deployed via Streamlit and Azure CI/CD pipelines.

Real Estate Recommendation System ⌇ | Python, Machine Learning, Streamlit   **2023-24**

- Engineered a robust application with advanced data cleaning, feature engineering, and selection, and analysis.
- Implemented and evaluated multiple regression models, enhancing property price prediction accuracy by 10%.
- Built a personalized recommendation system for tailored property insights, increasing user engagement by 20%.
- Deployed the application using Streamlit for real-time recommendations, reducing load time by 30%.
- Analytical mindset for innovative solutions, simplifying concepts, and presenting results via visual reports.

Shop Sales Analysis & Dashboard ⌇ | MySQL, Power BI, Data Analysis   **2023-24**

- Analyzed sales trends and product performance using MySQL, improving efficiency by 12%.
- Designed a Power BI dashboard with KPIs and heat maps, reducing reporting time by 20%.
- Boosted decision-making by 10% with interactive slicers and visualizations.

Bank Loan Report Dashboard ⌇ | Advanced Excel   **2023-24**

- Created a dashboard to track key loan metrics—applications, funded amounts, payments, and interest rates—improving tracking efficiency by 18% with visualizations such as line charts, filled maps, and bar charts.
- Reduced non-performing loans by 10% through portfolio monitoring and analysis of good vs. bad loans.

EDUCATION

KLE Institute of Technology   **2015 – 2019**
B.Tech Engineering   *Hubli, Karnataka*

TECHNICAL EXPERTISE & CORE COMPETENCIES

Programming: Proficient in Python (NumPy, Pandas, Scikit-learn, TensorFlow, PySpark) and SQL.
Data Science & Machine Learning: Proficient in building predictive models and applying advanced ML algorithms, statistical methods, and deep learning to solve real-world problems.
Data Analysis & Visualization: Advanced in Excel (Pivot Tables, Lookups, Macros) and Power BI (DAX, ETL), specializing in crafting interactive dashboards and reports to deliver data-driven insights.
Data Transformation & Big Data: Skilled in data cleaning, transformation, and modeling with Python, PySpark, and Delta Lake. Proficient in managing large datasets and optimizing big data workflows.
SQL & Database Management: Advanced proficiency in MySQL and PostgreSQL (Joins, Subqueries, Optimization). Skilled in querying, database design, and performance optimization.
MLOps & Deployment: Hands-on experience with MLOps tools like Azure ML, Docker and CI/CD pipelines for efficient machine learning deployment and monitoring.
Business Analytics: Trend analysis, analyzing data, and delivering insights to drive business decisions.
Tools & Platforms: VS Code, Jupyter Notebook, Azure ML Studio, Power BI and GitHub(version control)

CERTIFICATIONS

Data Science Master Program – *Besant Technology, Rajajinagar, Bangalore*
Completed training in Data Science, SQL, Power BI, Excel, Python, and Machine Learning.