

# PAWAN S BANT

Data Scientist

☎ +91-8762070702 ✉ [bantpawan@gmail.com](mailto:bantpawan@gmail.com) 🌐 [GitHub](#) 📁 [Portfolio](#) 🔗 [LinkedIn](#)

## Professional Summary

Data Science enthusiast with a strong foundation in analytical modeling, programming, and statistical analysis. Skilled in Python and SQL, focused on extracting insights and automating processes. Eager to apply my knowledge and continue learning in a data-driven environment.

## TECHNICAL SKILLS

**Programming:** Python (NumPy, Pandas, Scikit-learn), SQL, VBA

**Data Science:** Supervised/Unsupervised Learning, Deep Learning, Statistical Analysis

**MLOps:** Azure ML, Docker, CI/CD Pipelines, MLflow

**Big Data Tools:** PySpark, Databricks, Delta Lake

**Visualization:** Power BI (DAX, ETL), Excel (Pivot Tables, Macros), Streamlit

**Cloud/Databases:** Azure, MySQL, PostgreSQL

## PROJECTS

### 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%.
- Designed a personalized property recommendation system with Streamlit, boosting user engagement by 20% and reducing load time by 30%.

### Taxi Fare Prediction Application 🌐 | PySpark, Machine Learning, Streamlit, Azure 2023-24

- Developed an end-to-end ML application for NYC taxi fare prediction on a 35M-row dataset, leveraging advanced data cleaning and feature engineering to advanced 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%.
- Built a personalized prediction system with interactive visuals, boosting engagement by 25%. Fine-tuned regression models on Databricks, cutting response time by 30%, and deployed via Streamlit and Azure CI/CD pipelines.

### Enterprise Sales Dashboard 🌐 | Power BI, Advanced Excel, MySQL 2023-24

- Conducted in-depth analysis of sales trends using MySQL, enabling a 12% increase in operational efficiency and building an intuitive Power BI dashboard featuring KPIs and heat maps that reduced reporting time by 20%.
- Created an Excel dashboard for loan metrics, boosting tracking efficiency by 18% with advanced visualizations.
- Optimized decision-making by 10% and reduced non-performing loans by 10% through portfolio analysis.

## WORK EXPERIENCE

### Data Science Fellow, Besant Technologies 2023 – 2024

- Trained 100+ hours in ML algorithms, Power BI, and Azure deployment.
- Led 3 end-to-end projects from data collection to production deployment.

## EDUCATION

### B.Tech in Engineering, KLE Institute of Technology Hubballi, Karnataka

Relevant Coursework: Data science, Machine Learning, Statistics, Cloud Computing. 2015 – 2019

## Key Achievements

- Improved taxi fare prediction accuracy by 30% ( $R^2$ : 0.84) using Gradient Boosting, built a Streamlit app, and explored MLOps pipelines with Docker for scalable deployment.
- Streamlined Azure training expenses by implementing spot instances, cutting costs by 40%.
- Increased real estate price forecasting accuracy from 62% to 89% with Random Forest Regression.