SHUBHAM KUMAR PANDAY

AI Engineer | Data Scientist | Machine Learning Engineer
 +91 6360198698 | ravan5980@gmail.com | LinkedIn | GitHub

PROFESSIONAL SUMMARY

AI Engineer and Data Scientist with hands-on experience building **LLM applications, enterprise RAG systems, and production-ready ML pipelines**. Strong in Python, SQL, MLOps, and cloud-native deployments. Proven ability to convert business problems into scalable AI solutions using LangChain, vector databases, MLflow, Docker, CI/CD, and GCP. Passionate about designing systems that are **fast, reliable, and cost-efficient**.

WORK EXPERIENCE

AI Intern | Incanto Dynamics, Bengaluru
Mar 2025 – Jun 2025 - Built and optimized ML models for classification and regression use cases, improving model accuracy and recall.
- Designed automated data pipelines for feature engineering, reducing data preparation time by 25%.
- Performed exploratory data analysis (EDA) and developed interactive visual dashboards for stakeholders.
- Automated reporting workflows using Google Cloud APIs, enhancing data accessibility by 30%.

EDUCATION

Chanakya University
MCA (Data Science Specialization) | CGPA: 8.3/10 | 2023 – 2025

Bangalore North University
BCA (Computer Applications) | CGPA: 7.6/10 | 2020 – 2023

TECHNICAL SKILLS

Programming & Analytics: Python, SQL, Pandas, NumPy, Scipy, Scikit-learn
Machine Learning: Regression, Classification, Clustering, Feature Engineering, Model Evaluation, Ensemble Methods, Hyperparameter Tuning
Deep Learning: Neural Networks, CNNs, RNNs, Transformers, TensorFlow
Generative AI & LLMs: LangChain, RAG Pipelines, Vector Databases (FAISS, Pinecone), OpenAI APIs, Prompt Engineering
Data Visualization: Matplotlib, Seaborn, Plotly, Tableau, Power BI, Streamlit
Databases & Data Engineering: MySQL, PostgreSQL, ETL, Data Cleaning, Airflow (Basics), Data Pipelines
Statistics & Analytics: Hypothesis Testing, Predictive Modeling, A/B Testing, Correlation & Regression Analysis
MLOps & Cloud: MLflow, DVC, Docker, Kubernetes, CI/CD, GitHub Actions, GCP
Other Tools: Jupyter, VS Code, Git, Google Cloud APIs

PROJECTS

1. Production-Ready RAG System (LLM Application)

- Built an enterprise-level RAG pipeline using LangChain + FAISS with chunking, embeddings, and caching optimization.
- Integrated GPT-4/GPT-3.5 for reliable contextual answers; reduced latency by 18% via vector search optimizations.
- Deployed as a containerized microservice API using FastAPI + Docker.
- Added observability with logging, prompt tracking, and model-response analytics.

2. Hotel Reservation Cancellation Prediction (Analytics + ML)

- Built predictive ML model using Scikit-learn, improving accuracy by 20% using feature engineering & grid search tuning.
- Extracted insights on customer behaviour to reduce cancellation rates.
- Developed Tableau dashboards for management insights.

3. Vehicle Insurance Risk Prediction (End-to-End MLOps Pipeline)

- Designed a full ML lifecycle: ingestion → preprocessing → model training → deployment → monitoring.
- Used MLflow, DVC, GitHub Actions, Docker for reproducibility and CI/CD.
- Deployed on AWS with automated re-training & model versioning.
- Implemented real-time performance monitoring with alerts.

4. Automated ML Workflow (CI/CD + Model Registry)

- Built a system integrating MLflow, DVC, Kubernetes, and GitHub Actions.
- Reduced deployment cycle time and ensured experiment traceability & reproducibility.

ADDITIONAL DETAILS

- **Languages:** English, Hindi, Kannada, Telugu
- **Interests:** AI Research, LLM Optimization.