YASH PANDEY

yashpandey1626@gmail.com | P: +91 9821946530 | PORTFOLIO

EDUCATION

GGSIPU, September 2022 - Present Bachelor of Technology in Artificial Intelligence & Data Science
CGPA: 8.8

EXPERIENCE

AI/ML Intern
Dategain | March 2025 – Present | Delhi/Remote (Remote)

- Built CI/CD pipelines and a system to extract and capture information from screenshots and images using OCR and computer vision.
- Conducted EDA, data cleaning, feature engineering on large-scale datasets, and designed Power BI dashboards to visualize hiring trends.
- Developed AI-powered BIO and prompt generators using NLP and deep learning.

PROJECTS

Rooftop Solar Energy Potential Map | Github Skills used: Python, Deep Learning, Machine Learning, OpenCV

- Achieved 97.01% accuracy using YOLOv8-seg and ResNet50-U-Net models for building footprint extraction from satellite imagery
- Designed the full prediction pipeline, improving IoU from 0.85 to 0.92 using post-processing.
- Calculated rooftop areas for solar energy estimation, enabling strategic solar panel placement.

Network Security | Github Skills: Python, FastAPI, Docker, AWS CLI, S3, MongoDB, MLflow

- Designed and implemented a full machine learning pipeline for phishing detection, covering data ingestion, validation, transformation, model training, and deployment.
- Built modular, reusable components for preprocessing, feature engineering, and model training, and automated the pipeline with logging, exception handling, and MLflow tracking.
- Developed a **CI/CD** pipeline for automated model deployment, managed artifacts in AWS S3, and built a FastAPI backend for real-time phishing detection.

Quant(Trading Project)| Github Skills:Python, XGBoost, Logistic Regression, Technical Analysis, Bayesian Optimization

- Developed a machine learning model to predict short-term price movements (>0.5% within 5 minutes) with 81.33% accuracy by optimizing a feature set of 20+ technical indicators (RSI, MACD, Bollinger Bands) and engineered interaction features.
- Built an ensemble model combining XGBoost and Logistic Regression, using Bayesian optimization to fine-tune thresholds (0.4–0.65) and ensemble weights (0.5 XGBoost + 0.5 LR) for maximum predictive performance.
- Fine-tuned hyperparameters (e.g., `scale_pos_weight` between 4.0–4.5) through iterative experimentation, aligning model outputs closely with real-world high-frequency trading objectives.

SKILLS

Languages: Python | PostgreSQL | JavaScript
Technical Skills: Tensorflow| Pytorch | Scikit-learn| NLP | Deep Learning | Machine Learning | OpenCV
Tool & Platforms : Git & Github | AWS| Docker | Power BI | VS Code | Firebase

Achievements

• **ISRO Bhartiya Antariksh Hackathon – 2024 |** View LinkedIn Post with photos of the certificate and award Secured Top 12 nationwide among 3,500+ teams and 34,000 students.

• **Smart India Hackathon 2024:** SMART INDIA HACKATHON 2024 FINALIST **|** CERTIFICATE