DUVVURI DURGA PRASAD

Data Scientist | ML Engineer | Deep Learning Engineer |• Python • MLOps • AWS • MLflow

Address: Hyderabad, Telangana, India | +91 88858 02840 | satyadurgaprasad28@gmail.com

GitHub | LinkedIn

PROFESSIONAL SUMMARY

Data Scientist with a PG Diploma in AI & Data Science from the University of Liverpool, specializing in building and deploying production-grade ML systems. Experienced in designing scalable pipelines using MLflow, DVC, Docker, and AWS EC2, with robust MLOps integration. Hands-on across the entire ML lifecycle from data wrangling and feature engineering to deployment and real-time monitoring. Well-versed in CI/CD automation, model tuning, and experiment tracking for seamless reproducibility. Passionate about driving impact through AI in fintech, edtech, healthtech, and cloud-native environments.

EDUCATION

University of Liverpool, Liverpool, UK.
PG Diploma in Data Science and Artificial Intelligence – 09/2020 – 12/2024

- **Dissertation**: Designed and developed a Python-based 2D Sprouts game, a mathematical strategy game utilizing graph theory, with a focus on computational logic and algorithm design.
- **Relevant Coursework**: Machine Learning, Deep Learning, Data Visualization, Artificial Intelligence, Data Mining, Statistical Modeling

Avanthi Institute of Engineering and Technology, Hyderabad, India.
Bachelor of Engineering in Mechanical Engineering – 06/2013 – 05/2017

- **Final Year Project**: Designed and manufactured a working solar-powered car prototype, applying mechanical engineering principles and energy optimization techniques.
- **Relevant Coursework**: Thermodynamics, Engineering Mechanics, Computer-Aided Design (CAD), Control Systems, Applied Mathematics

PERSONAL PROJECTS

Credit Score Classification (FinTech, MLOps & Deployment)

- Built and optimized an end-to-end ML pipeline with 82% accuracy using Python and XGBoost. Applied feature selection and model evaluation techniques, with hyperparameter tuning via GridSearchCV.
- Used **MLflow** for **experiment tracking** and **model lifecycle management** and **DVC** for dataset versioning.
- Deployed the model via a **Flask API on AWS EC2** and **containerized the app using Docker** for scalable production use.

Student Performance Predictor (EdTech, ML & Deployment)

- Built a regression model for predictive modeling that achieved 90%+ accuracy using Scikit-learn, Pandas, and feature engineering techniques.
- Designed a robust **data pipeline** covering **preprocessing**, training, and performance monitoring using R^2 and MAE.
- Deployed a real-time prediction app using **Streamlit**, enabling accessibility for educators.

Movie Recommender System (NLP, MLOps & Streamlit Deployment)

- Developed a content-based recommendation engine using **cosine similarity** and TMDB metadata with NLP techniques.
- Tracked experiments using **MLflow** and versioned data with **DVC** to ensure reproducibility **and manage the complete model lifecycle**.
- Deployed the app via **Streamlit** and enhanced user experience by integrating with dynamic movie posters via the TMDB API.

Heart Disease Prediction (HealthTech, ML & Deployment)

- Built a clinical data-driven model (age, cholesterol, blood pressure) using multiple classifiers (logistic regression, random forest, XGBoost), achieving up to 86.6% accuracy.
- Performed EDA and feature engineering; deployed a real-time prediction app via Flask.

Portfolio: github.com/Duvvuridurgaprasad28

TECHNICAL SKILLS

- **Programming Languages:** Python (Advanced), SQL, HTML, CSS, Excel
- **Machine Learning & Deep Learning Libraries:** Scikit-learn, XGBoost, Pandas, NumPy, SciPy, TensorFlow, Keras
- **Deployment & Web Tools:** Flask, Streamlit, Render, FastAPI (Basics)
- **MLOps & DevOps Tools:** MLflow, DVC, Git, GitHub Actions, Docker, AWS EC2, Linux CLI
- **Data Visualization Tools:** Matplotlib, Seaborn, Plotly, Power BI
- **Databases:** MySQL, SQLite
- **Cloud Platforms:** AWS EC2 (Deployment), Azure (Learning), GCP (Learning)
- **Version Control:** Git, GitHub
- **Specialized Domains & Techniques:** Natural Language Processing (NLP), Large Language Models (LLMs), Exploratory Data Analysis (EDA), Feature Engineering, Data Cleaning, Data Preprocessing, and Data Engineering.
- **Mathematics & Theoretical Skills**: Statistics, Linear Algebra, Algorithms, Applied Mathematics.

CERTIFICATIONS

- Jovian – *Data Analysis with Python: Zero to Pandas – 05/2022 – 08/2022***.**
- *Jovian – Machine Learning with Python:* Zero to GBMs – 08/2023 – 02/2024.

SOFT SKILLS

- **Clear Communicator:** Comfortable explaining technical concepts to non-technical stakeholders.
- **Self-Starter**: Independently built and deployed full-stack ML applications, applying data insights to business decisions.
- **Creative Problem Solver:** Uses innovative approaches to tackle real-world challenges.
- **Curious & Adaptive Learner:** Actively exploring MLOps, cloud, and applied AI technologies.
- **Collaborative:** Actively contributed to team-based academic projects and group initiatives.