DEVIDUTTA PARIDA

AI Engineer
+91-9876543210 — deviduttaparida07@gmail.com — GitHub — LinkedIn

Summary

Aspiring AI Engineer and Data Science student with strong foundations in Statistics, Machine Learning, and Deep Learning . Experienced in building end-to-end ML pipelines, handling real-world data challenges, and deploying production-ready AI systems using FastAPI and Streamlit. Currently exploring Generative AI concepts including RAG, Vector Databases, and Semantic Search. Passionate about solving real-world problems using data-driven intelligence.

Education

B.Tech in Computer Science & Engineering 2023 – 2027
GIET University
CGPA: 8.0

Intermediate (12th) 2021 – 2023
APEX English Medium School

Secondary (10th) 2019 – 2021
Odisha Adarsha Vidyalaya

Technical Skills

Programming & OS:	Python, Java, Linux
Data Science & ML:	Pandas, NumPy, Scikit-learn, XGBoost, TensorFlow, SpaCy
ML Techniques:	Feature Engineering, Encoding, Standardization/Normalization, SMOTE, Cross Validation, GridSearchCV, RandomSearchCV, Overfitting & Underfitting Handling
Algorithms:	Linear & Logistic Regression, KNN, Decision Tree, Naive Bayes, SVM, Random Forest, XGBoost, Stacking, KMeans, DBSCAN, PCA
Generative AI:	RAG, FAISS (Testing), Quadrant (Production), Semantic Search, LLM API Integration
Visualization:	Matplotlib, Seaborn, Power BI, Excel

Projects

LoanLens AI (Universal Loan Approval Prediction System)
Tech Stack: Python, Pandas, Scikit-learn, XGBoost, FastAPI, Streamlit

- Aggregated and harmonized multiple public loan datasets to build a unified credit risk prediction system.
- Resolved inconsistent target encoding, structural mismatches, duplicate records, and outlier-sensitive imputation.
- Applied SMOTE and class-weighted XGBoost to handle imbalanced classification.
- Optimized model using Cross Validation and Hyperparameter Tuning.
- Achieved **F1 Score: 94.92%**.
- Deployed production-style pipeline using FastAPI and Streamlit

PdfSenseI (RAG-Based PDF Semantic Search System)
Tech Stack: Python, LLM API, Embeddings, RAG , Vector Database

- Built end-to-end Retrieval-Augmented Generation (RAG) pipeline for querying unstructured PDF documents.
- Implemented document ingestion, chunking, embedding generation, and vector similarity search.
- Designed semantic search system to retrieve the most relevant contextual information using vector.
- Integrated LLM API for accurate, context-aware answer generation.

BuySmart (PriceIntelligent Price Comparison System)
Tech Stack: Python, Pandas, Numpy, Selenium

- Built automated product data extraction pipeline using Selenium.
- Scraped pricing data across multiple e-commerce platforms.
- Designed comparison logic to identify lowest available product price.
- Structured extracted data for analysis and decision support.

Internship Experience

Machine Learning Intern Feb 2024 – Mar 2024
GIET University

- Developed Breast Cancer Prediction model using supervised ML.
- Performed preprocessing, feature engineering, and evaluation.
- Improved classification performance through ensemble techniques

Certifications

Data Analytics with Python – NPTEL
Data Science – Udemy
Industry Exposure Program – Hebbale Academy

Leadership & Achievements

Management and Operation Co-Head, Data Science Club
Runner-Up, Smart India Hackathon (SIH) 2024
Gold Medalist in Swimming

Additional Information

Languages: Hindi, Odia, English
Hobbies: Swimming, Reading Books, Online Strategy Games