

# Devansi Mahatar

Email: devansimahatar@gmail.com | LinkedIn | GitHub | Phone: 8249537214

## EDUCATION

VIT Bhopal University, B.Tech, CSE (Specialization in E-Commerce Technology) 2022-2026	CGPA: 8.85/10
DAV Public School, Pokhariput, CBSE XIIth (June 2022)	82%
Blossoms School, Bhubaneshwar, CBSE Xth (April 2020)	97.2%

## WORK EXPERIENCE

<b>Data Science Intern   Sabudh Foundation (Remote)</b>	Jan 2025 - June 2025
<ul style="list-style-type: none"><li>Designed and deployed LLM-powered document analysis system, reducing manual insight extraction time by 60%.</li><li>Implemented Retrieval-Augmented Generation (RAG) pipeline using Pinecone and LangChain for semantic search across multi-document datasets.</li><li>Engineered preprocessing workflows including chunking, embeddings, ranking optimization, and structured validation.</li><li>Created SQL-based agent and Streamlit interface enabling automated querying and real-time insight visualization.</li></ul>	

## PROJECTS

<b>1. InvoiceShield</b> - Invoice Anomaly Detection & ML Pipeline ( <a href="#">LINK</a> )	
<ul style="list-style-type: none"><li>Architected <b>end-to-end invoice processing pipeline</b> using OCR (88% field extraction accuracy) and engineered features for anomaly detection.</li><li>Applied <b>DBSCAN</b> clustering to detect pricing patterns and duplicates, improving anomaly detection precision by 25%.</li><li>Integrated <b>FastAPI</b>-based ML inference pipeline integrated with <b>PostgreSQL</b> for near real-time anomaly scoring.</li><li>Established automated data ingestion, preprocessing, and validation workflows to ensure structured consistency and reliable anomaly analysis.</li></ul>	
<b>2. ChurnX</b> - Customer Churn Prediction & Retention Modeling ( <a href="#">LINK</a> )	
<ul style="list-style-type: none"><li>Analyzed 7,043 customer records using <b>SQL</b> to calculate churn rates across tenure, contract type, and monthly charges.</li><li>Performed <b>data cleaning, EDA and feature engineering</b> in Python to identify <b>key churn drivers</b> and <b>high-risk segments</b>.</li><li>Optimized <b>Logistic Regression</b> and <b>Random Forest</b> models using GridSearchCV, achieving ROC-AUC score of 0.86.</li><li>Built <b>Power BI dashboard</b> to visualize churn trends, 30-35% revenue at risk, and retention-focused customer segments.</li></ul>	

## TECHNICAL SKILLS

<b>Machine Learning:</b> Feature Engineering, Cross-Validation, Hyperparameter Tuning, Logistic Regression, Random Forest, Decision Trees, DBSCAN, Linear Regression, K-means	
<b>Generative AI:</b> LangChain, Pinecone (Vector Database), Semantic Search, Prompt Engineering	
<b>Programming:</b> Python (Pandas, NumPy, Scikit-learn), SQL (Joins, Subqueries, Window Functions)	
<b>Tools &amp; Platforms:</b> PostgreSQL, MySQL, Streamlit, Git, Jupyter Notebook, VS Code, GCP	

## ACHIEVEMENTS & LEADERSHIP

<ul style="list-style-type: none"><li>National Semi-Finalist in Flipkart GRID 7.0</li><li>Advanced to the 3rd round of Arthneeti 4.0 business case study held by IIM Nagpur</li><li>Completed 2nd round (Technical interview), Google STEP internship</li></ul>	July 2024 - May 2025
---	----------------------

## LEADERSHIP EXPERIENCE

<b>Meraki - The Fine Arts Club   Decor Team Lead</b>	July 2024 - May 2025
Led a team of 10 members in planning and executing décor operations across 5+ college events, coordinating timelines and responsibilities to ensure smooth execution.	

## CERTIFICATIONS

<ul style="list-style-type: none"><li>GenAI Academy (Google Cloud x Hack2Skill) - Vertex AI, Gemini APIs, Imagen, Streamlit, Multimodal RAG</li><li>Dataiku Machine Learning Practitioner</li><li>Dataiku Generative AI Practitioner</li><li>Statistics Foundations Professional - Wolfram Research</li><li>HackerRank SQL (Intermediate)</li></ul>
---

**Interests:** Reading, Mandala Art, Travel and Cultural exploration

**Languages:** Fluent in Hindi, English, Odia; Conversational Proficiency in Tamil