Tata Data Analytics Virtual Internship 2025 – Task 2

This repository contains my submission for **Task 2** of the **Tata Data Analytics Virtual Internship 2025**, focused on **predictive modeling** using the **Geldium dataset**. This task builds on insights from Task 1 (EDA) and applies GenAI-powered methods to design a responsible and explainable credit risk model.

📌 Task Objective

To create a structured predictive modeling plan that helps identify customers at risk of **delinquency**, using AI/ML approaches supported by GenAI tools.

Key objectives included:

- ✅ Developing model logic for binary classification (delinquent or not)
- ✅ Selecting and justifying the appropriate model (Logistic Regression)
- ✅ Planning for model evaluation using industry-standard metrics
- ✅ Addressing ethical considerations: bias, fairness, and explainability

---

🛠️ Tools & Platform Used

- **GenAI tools** like ChatGPT & Gemini (for logic generation and refinement)
- **Prompt-driven modeling approach**
- **Microsoft Word** – final plan drafted and formatted for submission
- Dataset provided via **GNI platform**

□ GenAI Prompt Used

"Suggest a predictive modeling approach for credit delinquency using features like income, credit score, and missed payments. Explain model choice and evaluation metrics."

📂 Files Included

- `Geldium_Task2_Model_Plan.docx` – Final model plan including logic, justification, evaluation, and fairness considerations
- `README.md` – Overview of the task, methodology, and prompt used

📄 About the Internship

This task is part of the **Tata Data Analytics Virtual Internship 2025**, designed to equip learners with practical experience in using GenAI to solve real-world business problems, especially in the financial services sector.