Project Initialization and Planning Phase

Date	12 June 2025
Team ID	SWTID1749653449
Project Name	**Economic Growth: A Machine Learning Approach to GDP per Capita Prediction**
Maximum Marks	3 Marks

Define Problem Statements (Customer Problem Statement Template):

Accurate prediction of GDP per capita is critical for policymakers, economists, and financial institutions to understand economic growth, assess living standards, and guide investment decisions. Traditional econometric models often struggle to capture complex nonlinear relationships between economic indicators. This project proposes a machine learning-based approach to predict GDP per capita using multidimensional socioeconomic, demographic, and geopolitical data. By leveraging algorithms such as Random Forest, Gradient Boosting, and Neural Networks, the model aims to improve prediction accuracy and uncover key contributing factors. The project will also explore feature importance and model interpretability to support data-driven economic policy formulation.

Customer Problem Statement Template

	I am	I'm trying to	But	Because	Which makes me feel
	Economist or GDP Analyst	Forecast GDP based on the previous performances of World Economies	Existing forecasting tools are often outdated	They rely heavily on traditional econometric models	uncertain and constrained in my descision making
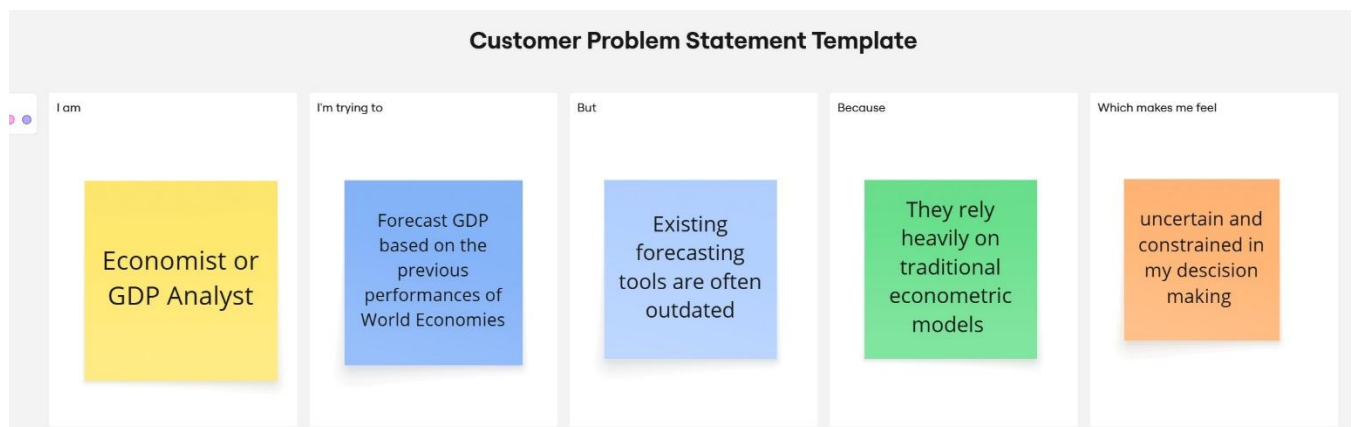
Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	a government policymaker/economic analyst who needs reliable, data-	understand and forecast GDP per capita more accurately to	existing forecasting tools are often outdated,	they rely heavily on traditional econometric models that	uncertain and constrained in my decision-making, knowing that critical economic policies

	driven insights to design effective economic policies and resource allocation strategies.	improve economic planning, reduce inequality, and promote sustainable development.	oversimplified, and fail to reflect complex, real-world economic interactions.	can't easily incorporate or analyze large, multidimensional datasets with nonlinear relationships.	may be based on incomplete or inaccurate information.