# Modeling Economic Resilience and Systemic Vulnerability

# A Machine Learning Exploration of Economic Stability and the Limits of Growth

Laura Rojas

---

## Justification

The modern economic discourse has long been dominated by neoliberal currents that artificially separated the market from its political, social, and ecological contexts. Under this paradigm, state intervention was reduced to the management of monetary flows—mainly through interest rate adjustments—while the economy was idealized as an autonomous, self-regulating entity. This depoliticization of the economy obscured its embeddedness in material and ecological realities.

At the same time, the reliance on Gross Domestic Product (GDP) as the central measure of progress has been widely criticized. GDP fails to account for social well-being, ecological degradation, unpaid labor, and systemic vulnerability. Yet, it remains the most consistent and comparable indicator available in official datasets. This project uses GDP-based data as a starting point—not to reaffirm its centrality, but to demonstrate empirically its limitations in capturing economic resilience.

By analyzing the capacity of national economies to withstand and recover from global shocks, this study seeks to reveal the structural fragilities of high-growth, neoliberal economies and open the discussion toward alternative indicators of resilience—those grounded in sustainability, equality, and socio-ecological balance. The goal is to bridge empirical modeling and critical political economy, showing that even the most powerful Global North economies that pursued hypergrowth lack the systemic resilience to protect their societies during global crises.

Thus, the project's contribution is twofold:

Empirical: Develop predictive tools to assess economic vulnerability using historical and contemporary data.

Theoretical: Demonstrate how resilience—understood as stability, adaptability, and sustainability—cannot be achieved within the logic of perpetual growth, inviting a shift toward degrowth and post-growth policy frameworks.

## General Objective

To analyze and predict the resilience of national economies to global shocks through a machine learning approach, in order to evaluate the structural limits of growth-based economic systems and contribute to the development of alternative frameworks for economic stability grounded in post-growth and degrowth principles.

## Specific Objectives

1. Empirically assess the capacity of different economies (Global North and Global South) to maintain stability during major global crises using historical and contemporary macroeconomic indicators.

2. Identify the main structural and policy factors that correlate with economic resilience and vulnerability, paying particular attention to those linked with neoliberal policy regimes (financial openness, debt levels, fiscal capacity, etc.).

3. Evaluate the predictive potential of machine learning models in identifying early-warning signals of systemic fragility.

4. Expose the conceptual limits of GDP-based measures in capturing social and ecological resilience, using the modeling results as evidence of the inadequacy of growth-centric indicators.

5. Propose alternative pathways for measuring and fostering resilience inspired by degrowth economics—highlighting well-being, resource balance, and democratic governance as critical dimensions.

## Research Questions

- What structural economic factors most strongly predict a country's ability to maintain stability during global shocks?
- To what extent do GDP-based metrics obscure underlying social and ecological vulnerabilities?
- Can machine learning approaches reveal new patterns of resilience that challenge the assumption that growth equals stability?
- How might this empirical evidence support a transition toward post-growth economic governance frameworks?

## From Growth Resilience to Systemic Resilience: A Paradigm Shift

Traditional economic policy treats resilience as the capacity to return to growth after a shock. This conception assumes that stability equals the resumption of accumulation. However, from a degrowth and ecological economics perspective, resilience must be redefined as the capacity of societies to maintain well-being and ecological balance under conditions of reduced throughput.

In this project, resilience is treated not as a return to "normal growth," but as a measure of systemic adaptability—how economic structures, institutions, and policies enable societies to sustain livelihoods without over-dependence on continuous expansion. The empirical models developed here will serve to identify the limits of the current growth regime and to open space for designing alternative, regenerative economic indicators.

## Stakeholders

Primary Stakeholders:

- Government Policy Makers & Central Banks: Need early warning systems for economic vulnerability assessment. In general the markets, organizations, and the private sector are always tracking risks of recession but it's not common to measure levels of preparedness for such economical shocks. It's like if you know that something is gonna happen but you don't know how vulnerable/prepared you are for that situation.

- International Development Organizations (World Bank, IMF, UN): Require risk assessment tools for aid allocation and policy recommendations

Secondary Stakeholders:

- Civil society organizations, sustainability policymakers, and researchers, Academic economic analysts

## Research and Policy Objectives

1. develop a framework to assess resilience beyond growth metrics
2. Predictive Accuracy: Develop models that can forecast economic stability with >85% accuracy during shock periods
3. Early Warning System: Create lead indicators that identify vulnerability 2-3 years before major shocks
4. Policy Insights: Generate actionable recommendations for building economic resilience
5. Risk Quantification: Provide numerical resilience scores for comparative country analysis
6. Resource Optimization: Enable targeted interventions and more efficient allocation of development resources

### Machine Learning Justification

A machine learning approach is essential to address some of the limitations of traditional methods:

- Complex Non-linear Relationships: Economic resilience involves intricate interactions between dozens of macroeconomic indicators that traditional econometric models struggle to capture
- Pattern Recognition: ML can identify subtle patterns in historical shock responses that human analysts might miss
- Multi-dimensional Analysis: The ability to simultaneously process 70+ economic indicators across multiple time periods
- Adaptability: ML models can learn from new shock events and adapt predictions accordingly

- Scale: Traditional analysis cannot efficiently process the volume of cross-country, multi-decade data required.

## Dataset Description

Primary Data Sources:

1. Maddison Project Database: Historical GDP and population data (1870-2018)

    - 169 countries, 12,337 observations
    - Founded on Angus Maddison's pioneering work in historical national accounts
    - Provides the longest-running, most comprehensive economic development dataset
2. World Bank Open Data: Contemporary economic indicators (1960-2023)

    - 24 key macroeconomic indicators
    - 9,935 observations across multiple countries

Final Engineered Dataset (Filtered using quality metrics):

- 1,292 country-year observations
- 78 sophisticated features including:
    - Core economic fundamentals (GDP, growth rates, development levels)
    - Investment and savings patterns
    - Trade and openness metrics
    - Financial development indicators
    - Government fiscal capacity
    - Labor and human capital measures
    - Innovation and technology adoption

### The Maddison Project Legacy

The Maddison Project, building on Angus Maddison's groundbreaking work, represents gold standard in historical economic data. Maddison pioneered the reconstruction of long-term economic development patterns, creating internationally comparable GDP series that enable researchers to study economic growth patterns across centuries. This dataset's unique value lies in its ability to capture long-term structural changes and provide context for modern economic resilience patterns.

### World Bank

The World Bank has provided high-quality, internationally standardized economic indicators with broad country coverage and consistent time series since 1960. Its rigorous data validation, frequent updates, and global comparability make it an essential source for analyzing contemporary macroeconomic trends and benchmarking resilience across countries and time.

### Economic Shocks Analyzed

Our analysis focuses on five major global economic shocks:

- Asian Financial Crisis (1997-1999)
- Dotcom Recession (2001-2002)
- Global Financial Crisis (2008-2010)
- European Debt Crisis (2010-2013)
- COVID-19 Pandemic (2020-2022)

### Success Metrics

Technical Metrics:

- Accuracy: >85% classification accuracy for stability prediction
- Precision/Recall: Balanced F1-score >0.80 for vulnerability detection
- Cross-validation: Consistent performance across time periods and regions
- Feature Importance: Clear identification of top predictive indicators

Additional Metrics:

- Policy Relevance: Actionable insights for economic planning
- Predictive Lead Time: 2-3 year advance warning capability
- Geographic Coverage: Model effectiveness across different development levels
- Interpretability: Clear explanation of vulnerability factors

---

## Problem Solving Process

### 1. Data Acquisition and Understanding

Data Collection Strategy:

Multi-source integration approach
Sources:
Maddison Project Database (Historical foundation)

World Bank Open Data (Contemporary indicators)
Economic shock identification (Previous Literature review)
Regional classifications (World Bank regions)

Data Quality Assessment:

- Coverage Analysis: Evaluated 169 countries for data completeness
- Temporal Consistency: Verified alignment across different data sources
- Missing Data Patterns: Systematic analysis of gaps and imputation strategies
- Outlier Detection: Statistical identification of anomalous economic patterns

Preliminary Visualization Strategy:

- Time series analysis of GDP trajectories during shock periods
- Regional comparison of economic resilience patterns
- Correlation matrices for indicator relationships
- Geographic mapping of vulnerability patterns

## 2. Data Preparation and Feature Engineering

Data Cleaning Approach:

Systematic cleaning pipeline
Steps:
1. Standardize country codes and naming conventions
2. Handle missing values using forward-fill and interpolation (used in time series)
3. Remove countries with <50% data coverage
4. Validate economic indicator ranges and relationships
5. Create balanced panel dataset structure

Advanced Feature Engineering: Our feature engineering creates five categories of derived indicators:

Economic Fundamentals:

- Volatility measures (3-year and 5-year rolling)
- Momentum indicators (growth acceleration/deceleration)
- Relative development positions

Resilience Indicators:

- Historical shock performance metrics
- Recovery time and strength measures
- Vulnerability scoring based on past shocks

Structural Indicators:

- Trade openness and balance measures
- Financial development indices
- Investment efficiency ratios

Temporal Features:

- Years since last shock
- Lagged economic indicators
- Trend and acceleration measures

Target Variable Engineering:

- Growth Stability Target: Composite measure combining GDP growth volatility, recovery speed, and sustained performance during shock periods.

Why this target variable? *Growth Stability*: We define economic resilience as the ability to keep annual GDP-growth on a tight leash when the global environment turns hostile. This metric penalises both deep recessions and wild rebounds, providing a forward-looking gauge that starts flashing red up to two years before a formal downturn. It is grounded in OECD business-cycle literature and passes the test for temporal variation.

Scikit-learn Pipeline Implementation:

Modular pipeline architecture
Pipeline Steps:
1. Data validation and cleaning
2. Feature scaling and normalization
3. Feature selection using multiple methods
4. Model-specific preprocessing
5. Cross-validation framework

## 3. Modeling Strategy

Algorithm Evaluation (Minimum 3 + Advanced Approaches):

Traditional ML Algorithms:

1. Random Forest: Excellent for feature importance and non-linear relationships
2. Gradient Boosting (XGBoost): Superior performance for structured data
3. Support Vector Machine: Effective for high-dimensional classification

Advanced Approaches:

4. Ensemble Methods: Voting classifiers combining multiple algorithms
5. Neural Networks: Deep learning for complex pattern recognition
6. Time Series Models: LSTM networks for temporal dependencies (optional)

Cross-Validation Strategy:

- Time Series Cross-Validation: Respects temporal ordering of economic data
- Geographic Cross-Validation: Ensures model generalizes across regions
- Shock-Period Stratification: Balanced representation of different crisis types

Hyperparameter Tuning:

- Grid Search: Systematic parameter exploration
- Bayesian Optimization: Efficient search for optimal configurations
- Feature Selection: Recursive feature elimination and stability selection

Evaluation Metrics:

- Primary: F1-Score (balanced precision/recall for vulnerability detection)
- Secondary: ROC-AUC, Precision-Recall curves
- Business Metrics: Lead time accuracy, regional performance consistency
- Interpretability: SHAP values for feature importance

## 4. Results Interpretation and Communication

Business Insight Translation:

- Convert statistical metrics into economic policy recommendations
- Quantify the economic impact of different resilience factors
- Provide country-specific vulnerability assessments
- Create actionable early warning indicators

Visualization Strategy:

- Interactive Dashboards: Country-level resilience scoring and trends

- Feature Importance Plots: Clear identification of key vulnerability factors
- Scenario Analysis: What-if modeling for policy interventions
- Geographic Mapping: Visual representation of global resilience patterns

Non-Technical Communication:

- Executive summary with key findings and recommendations
- Infographic-style presentation of main insights
- Case studies demonstrating model applications
- Policy brief format for government stakeholders

## 5. Conceptual Framework

Solution Pipeline Flowchart:

```
A[Historical Data Collection] --> B[Data Integration & Cleaning]
B --> C[Feature Engineering]
C --> D[Exploratory Data Analysis]
D --> E[Model Development]
E --> F[Model Evaluation]
F --> G[Business Impact Analysis]
G --> H[Deployment & Monitoring]

A1[Maddison Project] --> A
A2[World Bank Data] --> A
A3[Shock Identification] --> A

C1[Economic Indicators] --> C
C2[Resilience Metrics] --> C
C3[Temporal Features] --> C
```

E1[Random Forest] --> E
E2[XGBoost] --> E
E3[Neural Networks] --> E

F1[Cross-Validation] --> F
F2[Performance Metrics] --> F
F3[Feature Importance] --> F

G1[Policy Recommendations] --> G
G2[Risk Assessment] --> G
G3[Early Warning System] --> G

Project Dependencies:

1. Data Dependency: Maddison → World Bank → Feature Engineering
2. Technical Dependency: EDA → Feature Selection → Model Training
3. Validation Dependency: Historical Performance → Future Prediction Capability
4. Business Dependency: Model Accuracy → Policy Relevance → Stakeholder Adoption

---

# Timeline and Scope

## Project Phase Breakdown

Phase 1: Dataset Finalization and Problem Formulation

- Dataset Acquisition: Integrated Maddison Project and World Bank databases (2 hours)
- Business Problem Refinement: Defined economic resilience prediction framework
- Project Repository Setup: Organized modular project structure
- Stakeholder Analysis: Identified key user groups and success metrics
- Features research, design and implementation (8 hours approx.) Definition of categories of features: Economic Fundamentals, Financial Development, Trade Integration, Innovation Capacity, Temporal Features and Engineered Targets

Phase 2: Exploratory Data Analysis

- Comprehensive Data Profiling: Analyzed 1,292 observations across 78 features Statistical Relationship Analysis: Correlation matrices and dependency mapping

- Informative Visualizations: Time series plots, regional comparisons, shock impact analysis
- Insight Documentation: Preliminary findings on resilience / stability patterns

Phase 3: Data Preprocessing

Data Cleaning Implementation: Handling missing values and outliers

- Pipeline Development: Scikit-learn pipeline construction
- Data Splitting: Temporal and geographic stratification for train/validation/test. Split by size is not enough in this case due to changes by periods of time. This is crucial for economic data to prevent data leakage, we must not use future information to predict past events.

Phase 4: Model Development

- Baseline Model Implementation: Simple models for performance benchmarking
- Algorithm Comparison: Random Forest, XGBoost, SVM, Neural Networks
- Hyperparameter Tuning: Grid search and Bayesian optimization
- Cross-Validation: Time series and geographic validation strategies

Phase 5: Model Evaluation and Refinement

- Final Model Selection: Performance-based algorithm choice
- Test Data Evaluation: Unbiased performance assessment
- Business Metric Calculation: Policy-relevant impact quantification
- Results Interpretation: SHAP analysis and feature importance (optional). SHAP values explain how each feature contributes to a specific prediction for a single instance and are used in economic models.

Phase 6: Documentation and Reporting

- Code Documentation: Comprehensive commenting and cleanup
- Technical Report: Detailed methodology and findings documentation
- Executive Presentation: Stakeholder-focused presentation development
- Business Case: analysis and implementation roadmap

Phase 7: Final Review and Publication

- Quality Assurance: Code review and validation
- Final Submission: Complete project package delivery
- Peer Review Preparation and discussion material

## Project Structure

```
modern_economic_resilience/
├── data/                  # Raw and processed datasets
├── notebooks/             # Jupyter notebooks for analysis
    ├── 01_data_collection.ipynb
    ├── 02_exploratory_data_analysis.ipynb
    ├── 03_modeling.ipynb
    └── feature_engineering_advanced.ipynb
├── src/                   # Source code modules
├── config/                # Configuration files
├── models/                # Trained model artifacts
├── results/               # Model outputs and metrics
├── reports/               # Documentation and presentations
├── figures/               # Visualization outputs
```

## Potential Challenges and Research Areas

Technical Challenges:

1. Temporal Dependencies: Economic data has complex time-based relationships requiring sophisticated modeling
2. Data Sparsity: Historical data gaps may limit model training for certain countries/periods
3. Shock Heterogeneity: Different types of economic shocks may require specialized modeling approaches.
4. Model Interpretability: Balancing predictive accuracy with explainable insights for policy makers
5. 1,292 rows for 78 features is a very high ratio. Models like XGBoost can over-fit

Research Areas for Additional Learning:

1. Time Series Machine Learning: Advanced techniques for temporal economic data
2. Causal Inference: Methods to identify causal relationships in economic resilience
3. Ensemble Methods: Sophisticated combination techniques for improved prediction
4. Economic Theory Integration: Incorporating established economic principles into ML models

Risk Mitigation Strategies:

- Robust Cross-Validation: Multiple validation approaches to ensure model generalizability
- Feature Engineering Iteration: Continuous refinement based on domain expertise
- Incremental Development: Modular approach allowing for iterative improvements
- Regularization and permutation test to check leakage

### Success Criteria

This project will be considered successful when it delivers:

- A validated machine learning model achieving >85% accuracy in predicting economic stability
- Clear, actionable insights for policy makers and investors
- A robust, replicable methodology for ongoing economic resilience assessment
- Comprehensive documentation enabling future research and development

## Innovation and Impact

This project advances the dialogue between data science and post-growth economics by applying machine learning to reveal systemic vulnerabilities inherent to the growth paradigm.

- Combining Historical and Contemporary Data: Leveraging the Maddison Project's historical depth with World Bank's contemporary breadth
- Advanced Feature Engineering: Creating sophisticated resilience indicators beyond traditional economic metrics
- Multi-Shock Analysis: Examination of different crisis types and their patterns
- Actionable Intelligence: Translating complex ML outputs into concrete policy recommendations

The resulting framework will provide stakeholders with insight into economic vulnerability and resilience, enabling more informed decision-making in an uncertain global economy.