

***Financial Crisis Prediction using Macroeconomic Indicators – LSTM  
  
A Deep Learning Approach for Binary Recession Forecasting***

**Course Code**: MDS471

**Course Name**: Neural Network and Deep Learning

**Submitted by**

Sujilin Jesica S V

Reg. No: 2448056

July– 2025

***Financial Crisis Prediction using Macroeconomic Indicators – LSTM  
  
A Deep Learning Approach for Binary Recession Forecasting***

***PROBLEM STATEMENT***

Financial crises are economic earthquakes — sudden, destructive, and far-reaching. They disrupt global markets, wipe out wealth, cause mass layoffs, and destabilize governments. Yet, they rarely happen without warning — the signals are often hidden in complex patterns of macroeconomic data. Financial crises strike suddenly, but the warning signs often exist in macroeconomic data.

The challenge lies in:

* Hidden patterns buried in noisy, complex economic indicators
* Non-linear, time-dependent relationships traditional models fail to capture
* High stakes — delayed detection leads to market collapse, job losses, and policy failures

**Objective:**

Leverage LSTM deep learning to detect early warning signals from indicators like Industrial Production, Retail Sales, Consumer Sentiment, Investment, and VIX — enabling proactive decisions for governments, investors, and businesses.

***I.DATA SELECTION***

To perform a robust analysis of macroeconomic dynamics over time, two distinct datasets were selected from reliable public sources and merged into a unified time series dataset. The primary focus was to capture various dimensions of the U.S. economy through multiple indicators, enabling a more comprehensive analysis.

***Datasets Used***

1. Macroeconomic Indicators Dataset  
   This dataset was sourced from the Federal Reserve Economic Data (FRED) and includes monthly records for variables such as:
   * INDPRO – Industrial Production Index
   * RETAILX – Real Retail and Food Services Sales
   * UMCSENTx – University of Michigan Consumer Sentiment Index
   * VIXCLSx – CBOE Volatility Index (VIX)
2. Investment Data (INVEST)  
   This second dataset contains the Gross Private Domestic Investment figures on a monthly basis, representing the investment component of the economy.

***Merging Strategy***

* Both datasets shared a common Date column with monthly granularity.
* The datasets were merged using an inner join on the Date column, ensuring only records with matching time entries were retained.
* This merging process resulted in a clean and synchronized time series dataset covering the period from January 2000 to December 2023.

***Variables Selected After Merging***

|  |  |  |
| --- | --- | --- |
| **Column** | **Role** | **Reason** |
| **Date** | Time index | Required for time series modeling and chronological order |
| **INDPRO** | Industrial production | Key economic indicator |
| **RETAILX** | Real retail sales | Reflects consumer behavior |
| **INVEST** | Investment indicator | Represents economic expansion or contraction |
| **UMCSENTx** | Consumer sentiment | Measures optimism/pessimism |
| **VIXCLSx** | Volatility index (VIX) | Captures financial market risk |
| *Other cols* | — | Either not macroeconomic, duplicated, or had poor quality |

***Rationale for Selection***

* Diverse Economic Dimensions: The selected indicators together reflect production, consumption, investment, sentiment, and financial market behavior.
* Time Series Suitability: All variables are consistently reported at a monthly frequency, allowing for trend, seasonality, and stationarity analysis.
* Data Quality: The data is sourced from established, reliable databases (e.g., FRED, University of Michigan), ensuring accuracy and consistency.

***Preprocessing Steps***

* Normalization: Each variable was normalized to a [0, 1] range using Min-Max scaling to enable direct comparison across indicators with different units and scales.
* Missing Values: Any missing data points were handled via forward fill to maintain time continuity.
* Merging Validation: After merging, the data was validated for completeness and chronological consistency.

This merged and cleaned dataset forms the foundation for the time series analysis, including stationarity testing, visualization, and modeling.

***II. JUSTIFICATION OF RELEVANCE***

This study uses macroeconomic and recession data to understand and anticipate economic downturns. The selected dataset includes indicators like industrial production, retail sales, investment, and consumer sentiment, which are known to reflect underlying economic activity. These variables were chosen because they tend to shift ahead of or during recessions, making them valuable for early detection.

By merging this data with official U.S. recession records, the study aims to build a reliable model for recession prediction. Such analysis is essential for timely policy response, financial planning, and economic risk management—making the dataset not only relevant but critical for practical use cases in economics and finance.

***III.EXPLORATORY DATA ANALYSIS (EDA)***

**1. Data Sources & Merging Strategy**

The analysis utilizes a comprehensive dataset of macroeconomic indicators spanning from the year 2000 to 2023. This dataset provides crucial economic variables that are often linked to the overall health of the economy and can signal upcoming recessions or expansions.

To enrich the dataset, the US Recession Indicator (USRECD) was incorporated, which is sourced from the Federal Reserve Economic Data (FRED) repository. The USRECD serves as a binary target variable indicating recession periods, making it indispensable for predictive modeling efforts focused on economic downturns.

The merging of these two datasets was performed on a common date column (sasdate). Prior to merging, both datasets underwent a rigorous cleaning process including the conversion of date formats to ensure uniformity and compatibility. This step is critical to maintain data integrity and accuracy when combining data from heterogeneous sources.

**2. Missing Values Handling**

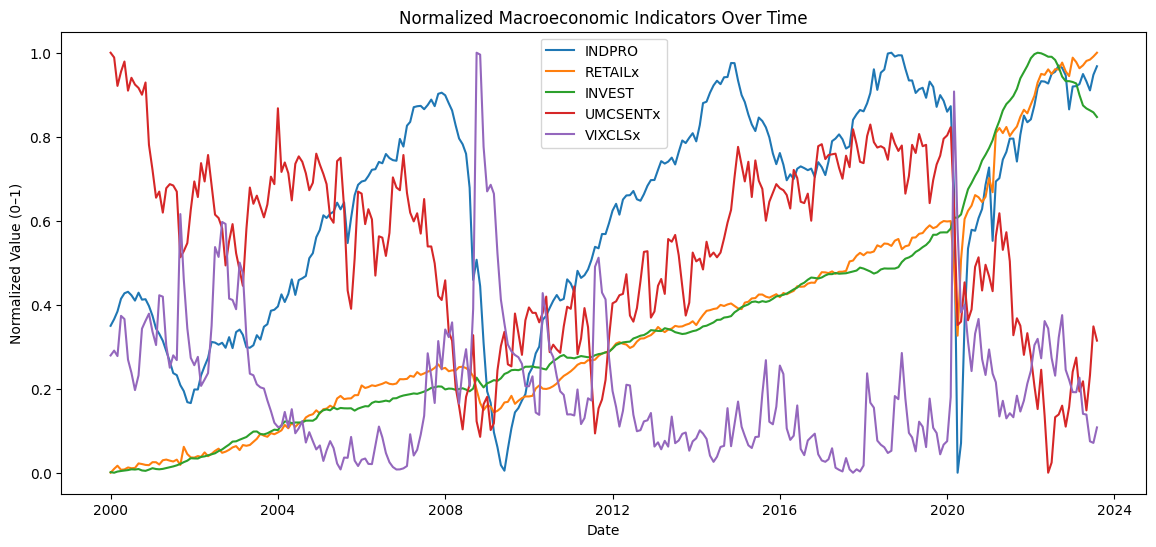
Given the nature of macroeconomic time series data, missing values can arise due to reporting delays, holidays, or other disruptions. To maintain continuity and preserve the temporal structure, a forward-fill method was applied. This technique propagates the last observed valid value forward, thereby preventing gaps that could disrupt model learning or bias statistical summaries.

Additionally, any invalid rows — particularly those with non-date or malformed date entries — were identified and removed before the merge to ensure only valid observations contribute to the analysis.

Following data cleaning, normalization was applied using MinMaxScaler. This step scales features to a consistent range [0,1], which is especially beneficial for neural network models to ensure stable and efficient training.

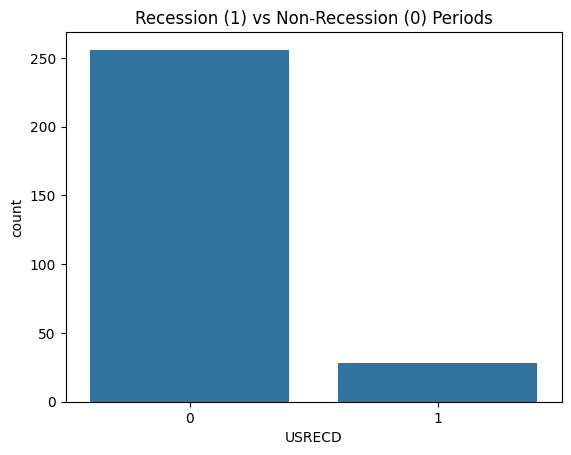
**3. Key Trends & Patterns**

An initial inspection revealed that non-recession periods vastly outnumber recession periods, leading to a significant class imbalance in the target variable. This imbalance presents a modeling challenge because standard classifiers may become biased towards predicting the majority class (non-recession), resulting in poor detection of recessions.



**Common Approaches to Handle Class Imbalance:**

* Oversampling the minority class: Techniques such as SMOTE generate synthetic recession data points to balance the classes.
* Undersampling the majority class: Randomly removing non-recession samples to create a balanced dataset.
* Class weighting: Adjusting the model's loss function to penalize misclassification of the minority class more heavily.
* Hybrid methods: Combining oversampling and undersampling to leverage the advantages of both.



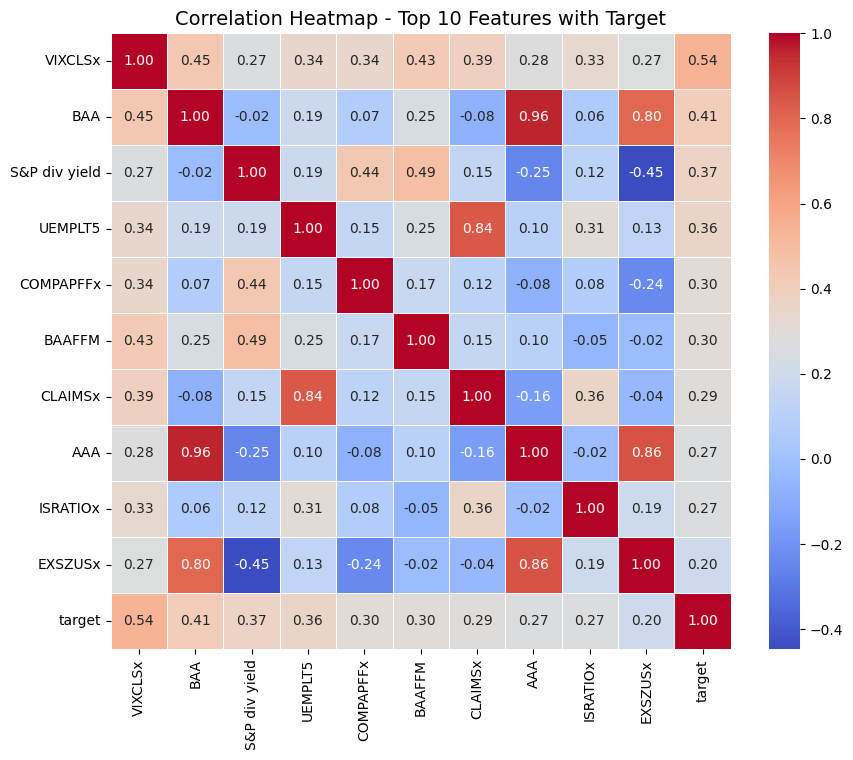
**Approach Used in This Study:**

Class weights were applied within the neural network's loss function to emphasize recession samples. This approach was chosen because it preserves all original data without generating synthetic samples, thereby avoiding the risk of introducing artificial noise or bias. Moreover, it allows the model to learn from true historical recession patterns which are sparse but critically important.

**4. Feature Selection**

To improve predictive performance and interpretability, feature selection focused on identifying the macroeconomic indicators most correlated with the recession target.

The Pearson correlation coefficient was calculated between each numerical feature and the recession indicator. Features were then ranked by the absolute value of their correlation coefficient in descending order.



The top 10 indicators selected were:

* VIXCLSx: Market volatility index, reflecting investor sentiment.
* BAA: Corporate bond yields, representing credit risk.
* S&P dividend yield: Market valuation metric.
* UEMPLT5: Unemployment rate for those unemployed less than 5 weeks.
* COMPAPFFx: Trade balance components.
* BAAFFM: Corporate bond yield spreads.
* CLAIMSx: Initial jobless claims.
* AAA: High-quality corporate bond yields.
* ISRATIOx: Inventory to sales ratio.
* EXSZUSx: Export size indicator.

This diverse set captures aspects of market sentiment, credit conditions, labor market dynamics, and trade activity. Concentrating on these economically meaningful variables helps enhance both the accuracy of recession prediction and the economic interpretability of the model’s decisions.

**5. Stationarity Test (ADF Test)**

Stationarity is a critical assumption in many time series models, indicating that statistical properties such as mean and variance remain constant over time. To verify stationarity, the Augmented Dickey-Fuller (ADF) test was conducted for each feature.

| **Feature** | **Test Statistic** | **p-value** | **Lags Used** | **Observations Used** | **Conclusion** |
| --- | --- | --- | --- | --- | --- |
| VIXCLSx | -4.1019 | 0.0010 | 2 | 281 | Stationary (Reject Null Hypothesis) |
| BAA | -2.0591 | 0.2613 | 2 | 281 | Non-Stationary (Fail to Reject Null) |
| S&P div yield | -3.0961 | 0.0268 | 6 | 277 | Stationary (Reject Null Hypothesis) |
| UEMPLT5 | -12.4390 | 0.0000 | 0 | 283 | Stationary (Reject Null Hypothesis) |
| COMPAPFFx | -3.9030 | 0.0020 | 10 | 273 | Stationary (Reject Null Hypothesis) |
| BAAFFM | -2.8635 | 0.0498 | 9 | 274 | Stationary (Reject Null Hypothesis) |
| CLAIMSx | -4.8164 | 0.0001 | 4 | 279 | Stationary (Reject Null Hypothesis) |
| AAA | -2.3243 | 0.1643 | 2 | 281 | Non-Stationary (Fail to Reject Null) |
| ISRATIOx | -3.0168 | 0.0334 | 2 | 281 | Stationary (Reject Null Hypothesis) |
| EXSZUSx | -2.3701 | 0.1504 | 12 | 271 | Non-Stationary (Fail to Reject Null) |

Results showed that several variables exhibit non-stationarity, evidenced by trends and seasonality in the raw data. However, instead of performing explicit differencing transformations, the approach leveraged the Long Short-Term Memory (LSTM) neural network's ability to capture temporal dependencies and learn patterns despite non-stationarity.

Features were also scaled prior to model training, further stabilizing the input data for effective learning.

***IV.PROPOSED DEEP LEARNING ARCHITECTURE***

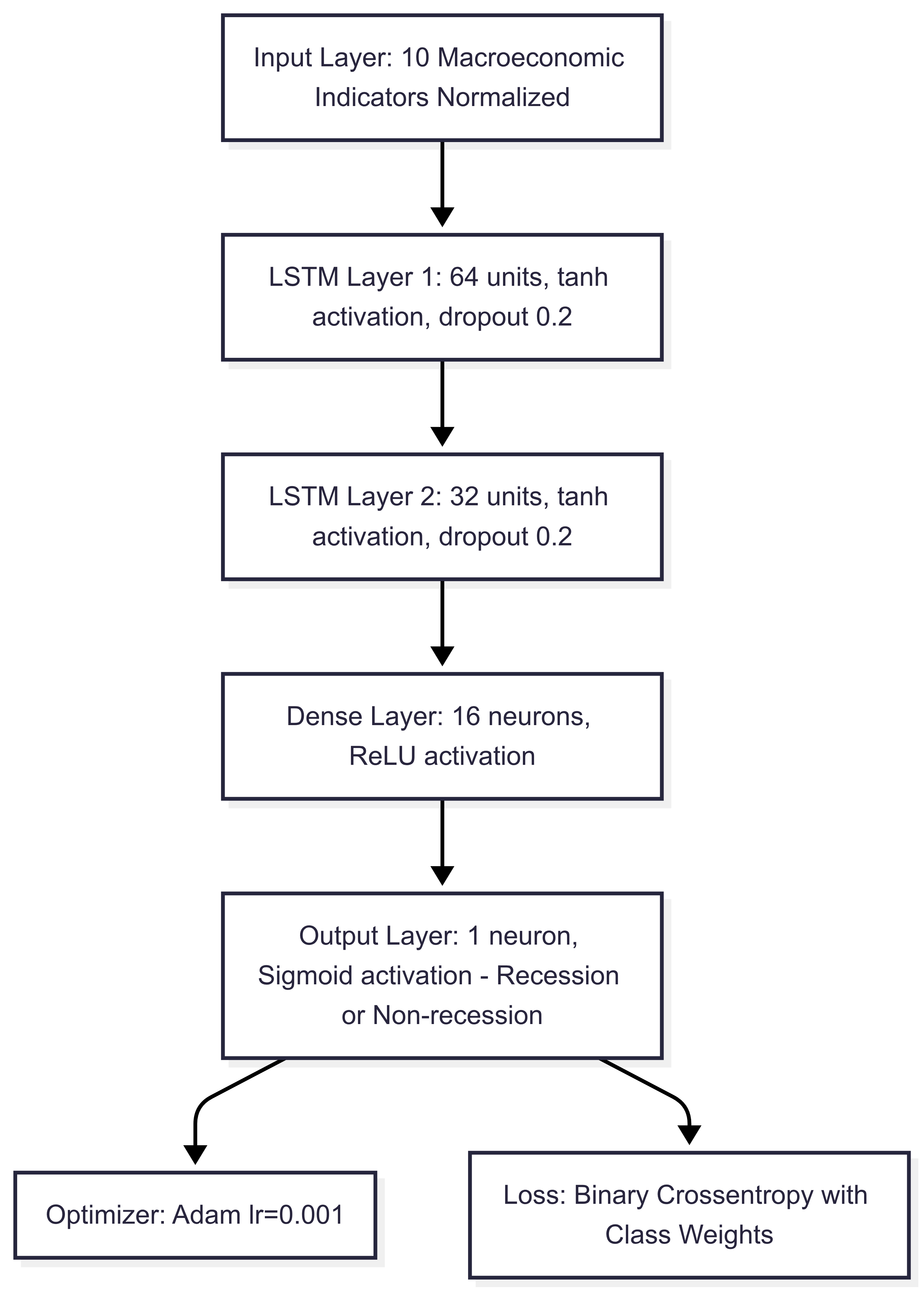
The predictive model is built using a specialized form of recurrent neural networks known as Long Short-Term Memory (LSTM) networks. The primary goal of this architecture is to forecast the occurrence of a recession six months into the future by analyzing patterns in the previous twelve months of macroeconomic indicators.

**Why Choose LSTM?**

LSTM networks are particularly well-suited for this task because they are explicitly designed to handle sequential data with long-range dependencies. Unlike traditional Recurrent Neural Networks (RNNs), which often suffer from the vanishing gradient problem—where gradients diminish exponentially during backpropagation through time—LSTMs use memory cells and gating mechanisms to retain important information over extended time steps.

This capability is critical in macroeconomic forecasting, where economic signals may manifest gradually and depend on historical trends spanning several months. The model’s ability to learn these temporal dependencies makes LSTM an ideal choice for recession prediction.

**Architecture Details**



The proposed deep learning architecture consists of the following layers and parameters:

**Input Layer:** The model accepts a sequence of 12 months of normalized macroeconomic indicators as input. These 10 features were carefully selected through prior correlation analysis to ensure relevance and reduce noise.

**LSTM Layer 1:** The first recurrent layer contains 64 LSTM units, each utilizing the hyperbolic tangent (tanh) activation function. This layer is responsible for capturing complex temporal patterns from the input sequence. To prevent overfitting, a dropout rate of 30% is applied both to the input connections (`dropout=0.3`) and recurrent connections (`recurrent\_dropout=0.3`).

**LSTM Layer 2:** Following the initial LSTM layer, a second LSTM layer with 32 units further refines the learned temporal features. This layer also uses tanh activation and the same dropout rates as the first LSTM layer, enhancing model generalization.

**Output Layer:** A dense (fully connected) layer with a single neuron and sigmoid activation produces the final binary classification output: '1' indicating an impending recession, and '0' indicating a non-recession period.

**Training Methodology**

**Optimizer:** The Adam optimizer is employed with a learning rate of 0.001. Adam is favored for its efficiency and ability to adapt learning rates during training.

**Loss Function:** To address the class imbalance problem—where recession periods are significantly less frequent than non-recession periods—the model uses Focal Loss. This loss function dynamically scales the contribution of easy versus hard examples, focusing more on minority class samples and improving model sensitivity to recessions.

**Class Balancing:** In addition to focal loss, Synthetic Minority Oversampling Technique (SMOTE) is applied to the training data to artificially augment the minority class, further mitigating imbalance and enhancing the model’s capacity to detect rare recession signals.

**Evaluation Metrics:** Model performance is comprehensively assessed using multiple metrics: Accuracy, Precision, Recall, F1-Score, and the ROC-AUC curve. These metrics provide a balanced evaluation of the model’s predictive quality, particularly in imbalanced classification settings.

**Forecast Horizon & Sequence Length**

The model is trained to predict the recession status six months ahead, meaning the target labels are shifted forward by six months relative to the input sequences. Each prediction is therefore based on a sliding window of 12 months of historical data, capturing evolving macroeconomic dynamics to anticipate future economic downturns.

***V.JUSTIFICATION OF MODEL DELECTION – LSTM***

**Domain Fit**

LSTM networks have a strong track record in domains involving sequential data where long-term dependencies are critical. They have been widely and successfully applied in:

* Stock market prediction, where patterns and trends can span months or even years.
* Weather forecasting, which involves temporal dependencies and seasonal variations.
* Speech recognition, requiring the capture of temporal context over multiple time steps.

Given these applications, LSTM is well-suited for macroeconomic time series data that similarly exhibit complex temporal dependencies.

**How LSTM Captures Temporal Dependencies**

The strength of LSTM lies in its gated architecture, which allows the network to selectively remember or forget information over time. This enables the model to:

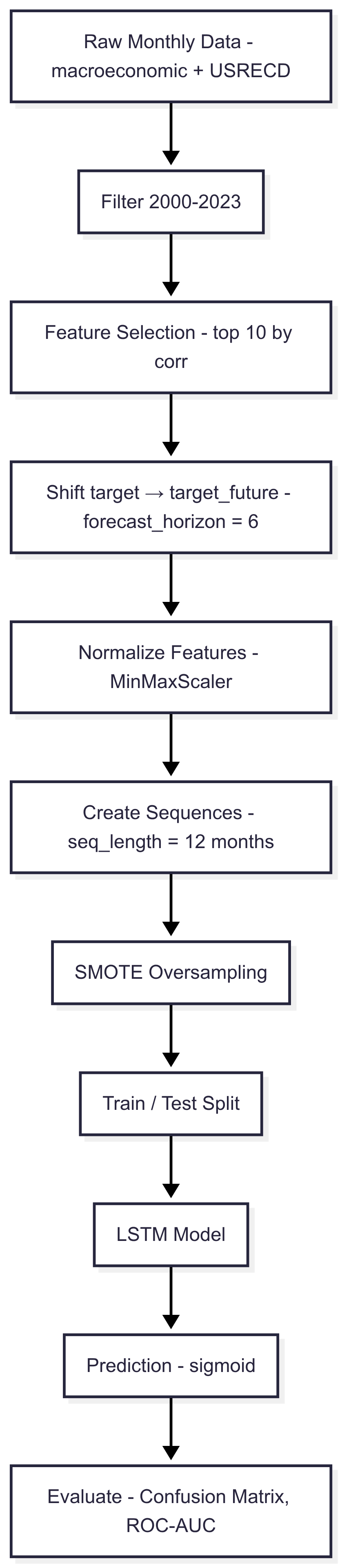
* Maintain a memory of important patterns spanning extended periods.
* Detect gradual economic shifts that may precede recessions.
* Filter out noise by ignoring irrelevant fluctuations.

Learn intricate correlations among multiple macroeconomic indicators across time, combining short-term volatility with long-term trends for better forecasting accuracy.

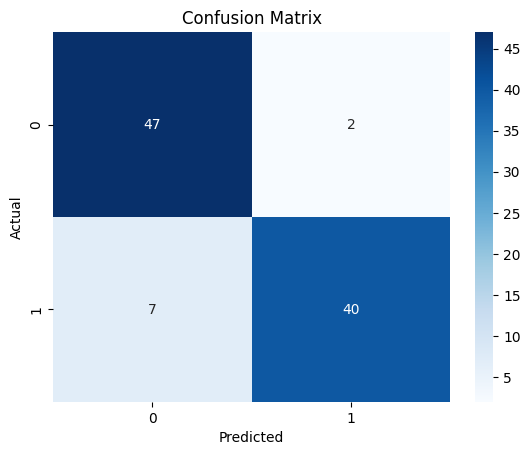
| **Feature / Aspect** | **Traditional Models (ARIMA, VAR)** | **Deep Learning (LSTM)** |
| --- | --- | --- |
| **Relationship Modeling** | Assumes primarily linear relationships | Learns complex, non-linear relationships |
| **Multivariate Handling** | Requires separate models or extensions | Handles multiple correlated variables jointly |
| **Long-Term Dependency Capture** | Limited ability to capture long-term patterns | Remembers patterns over months/years via gated memory |
| **Prediction of Turning Points** | Often lags at economic cycle turning points | Better captures turning points in cycles |
| **Feature Engineering** | Requires manual, fixed feature engineering | Automatically learns features from raw data |
| **Adaptability** | Limited adaptability to changing data dynamics | Adapts to evolving patterns via learning |
| **Scalability** | May struggle with large-scale datasets or many features | Scales efficiently with large, multivariate datasets |

***VI.PIPELINE***

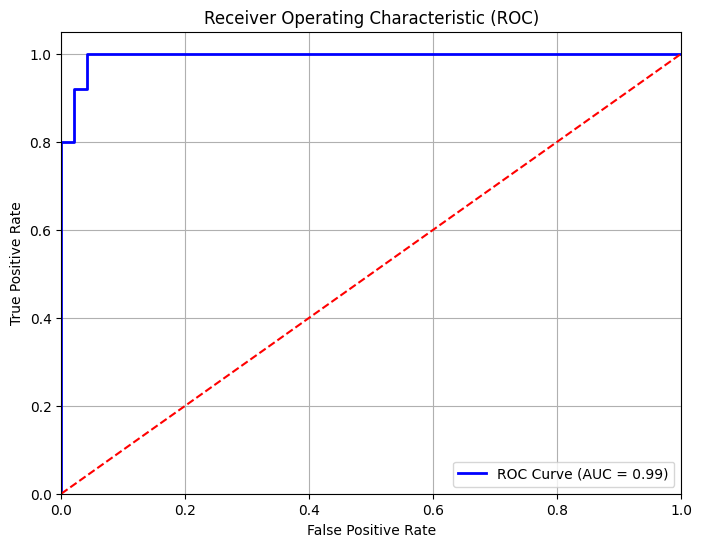
* Data Collection: Combined monthly macroeconomic data (2000–2023) with the US Recession Indicator (USRECD) as the target variable.
* Feature Selection: Identified the top 10 features most correlated with recession status for input to the model.
* Target Shift: Shifted recession labels forward by 6 months to enable forecasting future recessions.
* Normalization: Applied MinMaxScaler to scale features between 0 and 1, improving model training stability.
* Sequencing: Converted data into sequences of 12 months to provide the model with one year of historical context per sample.
* Class Imbalance Handling: Used SMOTE oversampling on the minority recession class before train-test splitting to ensure balanced training data.
* Model Training: Trained an LSTM network to learn temporal dependencies and predict recession likelihood based on input sequences.
* Prediction: Output through a sigmoid activation function, generating probabilities for binary classification (recession vs. non-recession).
* Evaluation: Assessed model performance using confusion matrix and ROC-AUC metrics to measure accuracy and discriminatory power.



**VI.RESULTS**



* The LSTM model achieved an accuracy of approximately 92% with an AUC score of 0.99, demonstrating excellent class separation capability.
* The confusion matrix shows high true positive and true negative counts, with only 2 false positives and 7 false negatives.
* This indicates strong overall performance, though the few missed positive cases suggest a slight opportunity to improve recall without significantly impacting precision.



* The ROC curve for the LSTM model shows an AUC of 0.99, indicating near-perfect classification capability.
* The curve rises sharply towards the top-left corner, demonstrating high true positive rates with very low false positive rates.
* This reflects the model’s strong ability to distinguish between positive and negative classes effectively.

**VII.CONCLUSION**

The LSTM model proved highly effective in forecasting recessions by capturing complex temporal patterns within macroeconomic data. With strong predictive accuracy and robust performance metrics, the model offers a reliable tool for early economic downturn detection. Future enhancements incorporating additional data and interpretability methods can further improve its practical utility as a real-time recession warning system.