Journal of Data Science 0 (0), 1–20
July 2026

DOI: 10.6339/26-JDS1223
Data Science in Action

Predicting Business Cycles Using Deep Learning Models

Binod Rimal[1,*], Sebastian Tardieu[1], Ramchandra Rimal[2], and
Hum Nath Bhandari[3]

[1] 401 W Kennedy Blvd, Tampa, FL, 33606, University of Tampa, USA
[2] 1301 E Main St, Murfreesboro, TN, Middle Tennessee State University, USA
[3] 1 Old Ferry Road, Bristol, RI, 02809, Roger Williams University, USA

Abstract

Forecasting business cycles and macroeconomic trends is inherently challenging due to their complex and non-linear relationships with volatile and noisy economic factors. However, the growing availability of large-scale economic data, coupled with advances in computational power, creates new opportunities to extract meaningful information and develop robust predictive models. Deep learning methods are well suited for handling noisy and complex data; nevertheless, their application to business cycle prediction remains at an early stage. To address this gap, this study develops an end-to-end computational framework that implements state-of-art deep learning architectures for identifying business cycle phases. The key contributions of this work include systematic input feature selection across a broad range of economic sectors, advanced preprocessing techniques for noisy data, the development of a customizable and reproducible computational framework, a data-driven approach to hyperparameter tuning, and the use of robust model selection strategies. Twenty two models based on two architectures—Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU)—were implemented and demonstrated strong predictive performance, highlighting the effectiveness of the proposed approach for forecasting macroeconomic trends. Among these, the GRU model with 512 neurons achieved the best results, with an accuracy of 86.42%, precision of 92.03%, recall of 93.32%, and an F1-score of 92.34%. Overall, the findings provide valuable insights that can support informed decision-making by policymakers and other stakeholders.

Keywords *forecasting; GRU; LSTM; recession; time series*

1 Introduction

Economic activity naturally alternates between periods of growth, known as expansions, and periods of decline, known as recessions. This cyclical pattern, called the business cycle or economic cycle, describes the economy's movement through recurring phases of expansion and contraction (Weinstock, 2023; Emsbo-Mattingly et al., 2014). In the United States, the National Bureau of Economic Research (NBER) is widely recognized for identifying and dating business cycles (Weinstock, 2023). The NBER determines cycle dates based on the peaks and troughs of economic activity and defines a recession as a significant and persistent decline in activity that is widespread across the economy. To make this determination, the agency considers a range of indicators, including real gross domestic product (GDP), total employment,

---

[*] Corresponding author. Email: brimal@ut.edu.

Received December 31, 2025; Accepted February 22, 2026

actual sales, and industrial production (Weinstock, 2023; Emsbo-Mattingly et al., 2014). A key measure of overall economic performance is inflation-adjusted real GDP, which reflects the total value of goods and services produced within a country. Quarterly fluctuations in GDP provide the signal for expansions or contractions of economy, shaping policy decisions and market expectations (Ma and Zhang, 2016; Bessec and Bouabdallah, 2015). Within each business cycle, real GDP typically passes through four stages: expansion, peak, contraction, and trough (Weinstock, 2023). Economists have systematically documented the recurrence of these patterns over the past century, even as the specific features of individual cycles varied in magnitude and durations (Hamilton, 2005). For example, in the United States, the COVID-19 recession technically lasted only two months while the previous recession started in 2007 lasted for 18 months. Sometimes, economic recession can drag down into economic depression, as experienced during the 1930s, when the Great Depression lasted 44 months. On the other hand, sometimes the economy may experience a slowdown but avoid recession, referred to as soft-landing by many economists as we have observed in 2024.

Multiple factors influence business cycles, including labor market conditions—measured by the unemployment rate, the duration of unemployment, and average weekly hours in manufacturing-which indicate the economy's overall strength (Buser et al., 2024; Ehrenberg et al., 2021). Other indicators such as industrial production, housing starts, and consumer sentiment also play vital roles. Interest rates and treasury bond yields affect borrowing costs and liquidity, while the yield curve often serves as a predictor of economic outlooks (Mitchell, 2023; Chodorow-Reich and Wieland, 2020). Price stability, monitored through the consumer price index, reflects inflationary pressures that can influence purchasing power and monetary policy (Cacciatore and Ghironi, 2021; Bianchi et al., 2023). Together, these measures offer a comprehensive view of macroeconomic trends, guiding both policymaking and market strategies.

Understanding and modeling business cycles has long been a central challenge in macroeconomics, as researchers seek to explain the persistence, co-movement, and volatility of economic fluctuations across time and regions. Early contributions, such as the Real Business Cycle (RBC) framework, emphasized real shocks, particularly technology shocks as fundamental drivers of cycles, highlighting mechanisms like the time-to-build process for capital formation and the efficient response of rational agents to productivity changes (Kydland and Prescott, 1982; Long Jr and Plosser, 1983). While these models successfully replicated key patterns in employment, productivity, and output dynamics, they often struggled to account for nominal factors such as wage and interest rate behavior without additional assumptions (King and Rebelo, 1999). Cross-country analyses further revealed that business cycles exhibit both synchronization and divergence, shaped by regional specialization, mobility, and the extent of global integration (Woitek, 1993; Krugman, 1993). Further empirical studies have also underscored the limitations of purely real or sticky-price models in fully explaining observed macroeconomic regularities, pointing to the need for frameworks that integrate both real and nominal dynamics (King and Watson, 1996). More recent research has expanded the scope to international and supranational dimensions, showing that globalization fosters overlapping, polycentric cycles influenced by common shocks that propagate unevenly across economies (Graff, 2011; Sella et al., 2016).

An alternative approach to the study of business cycles builds on a hierarchy of disequilibrium macroeconomic models that integrate inventory dynamics, distributive-share and wage–price cycles, and monetary demand factors. This framework, known as the Keynes-Metzler-Goodwin (KMG) model (Chiarella and Flaschel, 2000), has been extensively developed and applied in the literature. Asada et al. (2003) examined the implications of manipulating monetary aggregates, which influence inflation indirectly, while contrasting this with interest rate adjustments that have a more direct effect on inflation dynamics. Chiarella et al. (2005) demonstrated that business cycles may emerge within the KMG framework when the model's linear core loses stability through a Hopf bifurcation, and that nonlinearities in macroeconomic factors such as investment can generate persistent and bounded limit cycles even in the absence of external shocks. Their analysis further showed that stronger demand and more aggressive price wage adjustments can drive the system into cyclical regimes, with the model reproducing key historical features of U.S. business cycles. Subsequent contributions have expanded the framework. Yoshida and Asada (2007) investigated the role of policy lags and dynamic properties in disequilibrium settings, offering insights into the experimentation of monetary policy rules within the KMG context; Asada et al. (2011) incorporated richer stock flow interactions, portfolio behavior, and policy mechanisms; and Sołtysiak (2023) proposed an improved variant of the KMG system, reformulated in intensive form with local stability proofs for the steady state. Most recently, Di Guilmi et al. (2023) extended the KMG tradition by embedding endogenous political choices and behavioral factors, further broadening its applicability to contemporary macroeconomic dynamics.

These developments highlight the complexity of modern business cycles and the necessity for the continuous innovations in developing new state-of-art predictive models to uncover the structural, institutional, and global interactions. While early studies focused on theoretical development and empirical validations using econometric model for business cycle. Recent studies utilized those models along with new tools in predictive modeling. For instance, Korobilis (2018) explores econometric methods to handle high-dimensional macroeconomic models, highlighting inference challenges with more parameters than observations. However, this work focuses mainly on univariate linear regression and provides a foundational understanding of the economic dynamics that predictive models aim to forecast. The study of Thorsrud (2020) constructs a business cycle index by decomposing newspaper content into a time series of news topics and applying a time-varying dynamic factor model with a latent threshold mechanism. This approach classifies business cycle phases and provides high-frequency information on economic fluctuations. Fernández-Villaverde and Guerrón-Quintana (2020) developed a business cycle model incorporating financial frictions and uncertainty shocks, showing that such shocks can sometimes lead to expansionary effects. This research examines the variability in macroeconomic time series and its impact on economic fluctuations, highlighting the critical role of uncertainty in macroeconomic dynamics and providing a foundation for future studies. Cacciatore and Ravenna (2021) examines how wage rigidity and uncertainty shocks worsen economic downturns by increasing firms' hiring conservatism and profit-risk premiums. It further concludes that uncertainty rises in recessions due to higher unpredictability in economic outcomes, even without additional shocks.

From the perspective of investment portfolio construction, understanding the business cycle is essential for both institutional and individual investor. Because the broad stock market often moves in tandem with the underlying business cycle, informed investors can position their portfolios to align with different stages of the cycle (Emsbo-Mattingly et al., 2014). For example, during the early expansion or recovery phase, sectors sensitive to interest rate changes and government stimulus—such as consumer discretionary and financial often outperform defensive sectors like utilities and consumer staples. Likewise, high-growth and economically sensitive sectors, such as information technology and industrials, tend to thrive during sustained expansions. Investors with a long-term horizon and higher risk tolerance may increase exposure to these sectors to capture the upside of a bull market. However, as the economy approaches its peak and growth begins to slow, risks can outweigh rewards in these volatile sectors. In such cases, gradually shifting toward more defensive, less interest-rate-sensitive sectors—such as utilities and healthcare—can help preserve capital and reduce exposure to downturns. This shows that it is essential to construct a disciplined investment framework that incorporates knowledge obtained at the moment by using multi-faceted macroeconomic factors. It gives investors a better understanding of the dynamics of the given phase of the business cycle in terms of potential length and impact. Thus, predicting business cycles is a significant research problem. However, it is inherently difficult to forecast due to its intricate relationship with numerous other factors such as non-linear behavior and high volatility. Achieving accurate forecasts demands a robust approach that addresses these challenges by integrating pertinent economic data while filtering out irrelevant information. However, many predictive models encounter difficulties identifying the optimal set of input features or striking the right balance between model complexity and predictive power. Furthermore, most statistical models, especially those designed for time series forecasting, rely on certain assumptions that may not always hold. Additionally, the selection of models and their complexity by forecasters is often influenced by their economic intuition and judgment, which can introduce biases and lead to flawed assumptions (Valaitis and Villa, 2024).

There are mainly three schools of thought in predicting business cycles: statistical, machine learning, and deep learning. Statistical models have long been used to predict business cycles, with techniques such as autoregressive integrated moving average, logistic regression, and dynamic factor models being prominent. These models focus on time-series analysis and identifying patterns within economic data. However, despite their utility, they often struggle with timeliness and efficiency in prediction, especially during periods of economic volatility. The studies of Vrontos et al. (2021); Siami-Namini and Namin (2018); Tehranian (2023); Beutel et al. (2019) illustrate the limitations and comparative performance of traditional statistical models versus emerging ML approaches. Recent advancements in ML have revolutionized business cycle prediction by offering sophisticated algorithms capable of capturing complex nonlinear relationships in economic data. Studies by Goulet Coulombe et al. (2022), Hall (2018), Paruchuri (2021), Liu et al. (2022), Tehranian (2023), Zheng et al. (2023), and Nosratabadi et al. (2020) demonstrate the superiority of ML techniques, particularly in handling large, diverse datasets and improving prediction accuracy. These approaches often outperform traditional models, especially during high macroeconomic uncertainty, financial stress, and economic downturns (Soybilgen, 2020; Bluwstein et al., 2023; Fouliard et al., 2021). Deep learning, particular machine learning approaches, offer a distinct advantage by automating many decisions typically made by forecasters, thereby reducing the impact of subjective discretion. Central to deep learning is optimizing models capturing non-linear interactions between the multitude of factors and the long-term temporal dependencies to predict the future without relying on assumptions. They can handle large datasets more effectively than traditional econometric models and automatically learn the best features from raw data through layers of abstraction. This ability to learn high-level representations makes them powerful in capturing intricate relationships in the data, making them effective to predict financial crisis (Sezer et al., 2020; Wasserbacher and Spindler, 2022; Chatzis et al., 2018; Hopp, 2022; Venkateswarlu et al., 2022).

There are several gaps in the literature for forecasting business cycles using deep learning. Firstly, business cycles are influenced by a wide variety of data sources, making a significant obstacle in the data integration. Additionally, deep learning models are often susceptible to overfitting, mainly when the dataset is limited or contains noise. These models must demonstrate robustness and the ability to generalize to new, unseen data. Moreover, there is a lack of standardized benchmarks and datasets to compare different deep learning models in business cycle forecasting to ensure consistent and fair evaluations. To address these challenges and bridge the exist-

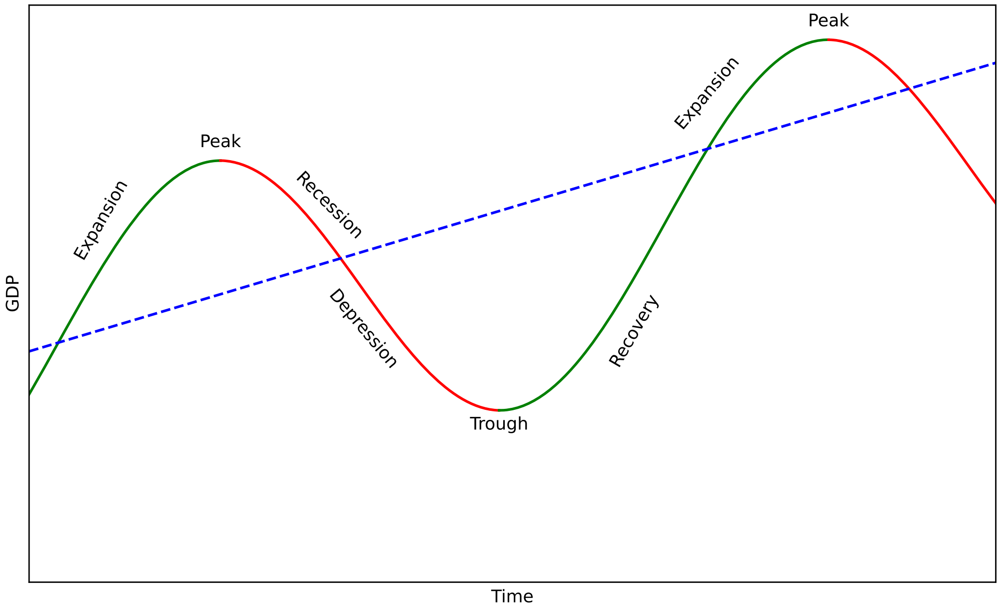

Figure 1: Illustration of the phases of the business cycle relative to the GDP curve, with non-recession periods shown in green and recession periods shown in red.

ing research gap, our study undertakes the following efforts: (a) identifying and selecting relevant features from a comprehensive set of input variables across diverse economic sectors, (b) maintaining data quality and transparency by extracting information from publicly accessible reliable sources, (c) implementing advanced data preprocessing techniques including denoising, (d) developing an integrated and customizable computational framework using deep learning techniques, (e) experimenting with several deep learning model configurations with various degrees of complexities, (f) designing a robust data-driven approach for hyperparameter tuning and model selection, (g) ensuring the reproducibility of results, serving as a baseline case study in this domain.

The rest of the paper is organized as follows. The data collection and feature selection procedure are explained in Section 2. Section 3 offers brief descriptions of the deep learning models used in this study along with optimization techniques and performance evaluation metrics for binary classification. The experimental design, model construction, and obtained results are discussed within Section 4. Section 5 presents the conclusion and future work, summarizing the findings and indicating potential directions for further research. Ethical considerations and implications are discussed in Section 6. The paper concludes with acknowledgment and a list of references.

2 Data Preparation and Variables Selection

Figure 1 represents an illustrative example of a typical business cycle, showcasing the fluctuations in economic activity over time as measured by GDP. An expansion phase is marked by a green line indicating increasing economic activity. This upward trend culminates at a peak where economic growth reaches its highest point. Following the peak, the economy enters a recession phase, depicted by a red line, where economic activity declines. This downturn continues into a depression, the cycle's lowest point, labeled as a trough. After the trough, the economy experiences a recovery, shown again by a green line, as economic activity improves. This recovery

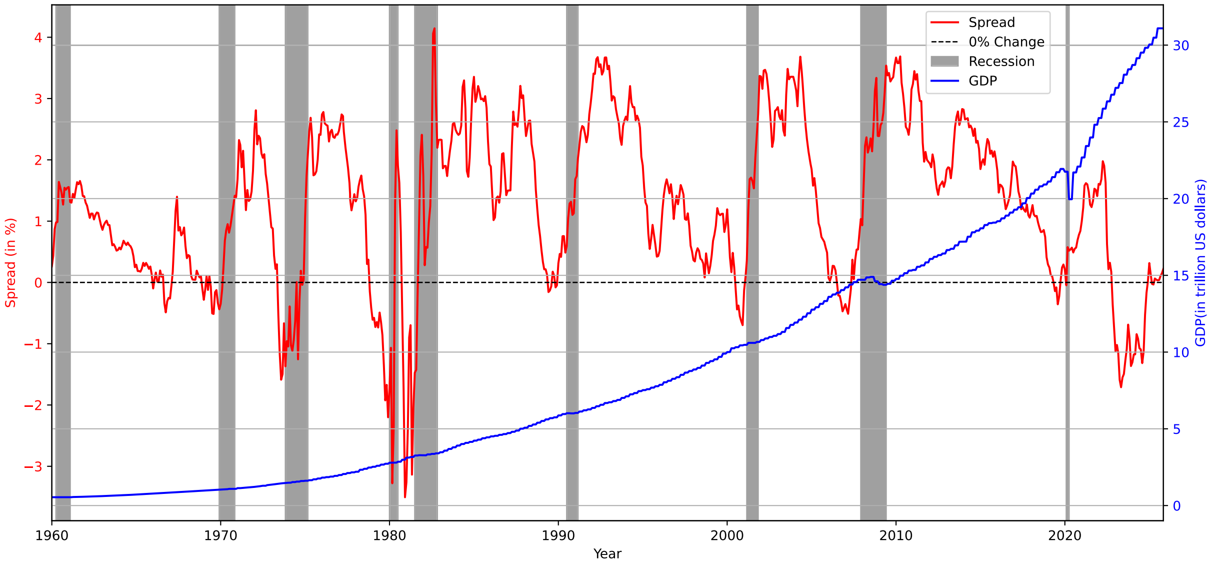

Figure 2: US GDP and spread (10-year minus 3-year treasury yields) along with highlights of official recession periods.

eventually leads back into another expansion, continuing the cyclical pattern. In this study, the trough to peak period is defined as a non-recession (1) period represented in green, and the peak to trough is recession (0) phase in red. The graph also includes a blue dashed line representing a steady growth trend, highlighting the long-term upward trajectory of economic progress despite short-term fluctuations.

Figure 2 illustrates the relationship between GDP growth, spread, and the business cycle, highlighting periods of recessions with a gray bar. The red curve illustrates how the term spread between the 10-year and 3-year Treasury yields evolved from 1960 to 2025, and its relationship with business cycles. Periods when the spread dips below the zero line represent yield-curve inversions—an event historically recognized as one of the strongest predictors of U.S. recessions. Nearly every recession is preceded by a negative term spread, often with a lead time of about 6 to 18 months. While real GDP (shown by the blue curve) follows a long-run upward trajectory, reflecting sustained economic growth, each recession creates a noticeable disruption in that trend. Overall, the graph highlights a consistent and powerful relationship between yield-curve inversions and subsequent economic downturns, underscoring the spread's value as an early-warning indicator for recession risk.

Table 1 lists periods of economic expansion and contraction from 1960 to 2025, as defined in business cycle dating information by the NBER (NBER, 2024). Over these 791 months, there were nine recessions, totaling 95 months of economic downturn. This study aims to classify business cycle phases—recessions (0) and non-recession (1) using sequential modeling in deep learning.

This study incorporates 13 input features from seven sectors of the economy: income/output, labor market, housing, consumption, interest rates, prices, and the stock market, as listed in Table 2. These variables represent diverse macroeconomic sectors (McCracken and Ng, 2016). These macroeconomic factors sufficiently represent their impacts on economic changes within the United States (Hall, 2018). All variables are extracted from the Federal Reserve Economic Data (FRED) database, maintained by the Federal Reserve Banks of St. Louis and New York,

Table 1: Business cycles, their durations, and respective binary class as dated by the NBER (NBER, 2024).

Peak to Trough	Trough to Peak	Duration (month)	Business Cycle
Apr 1960 - Feb 1961	-	10	0
-	Feb 1961 - Dec 1969	106	1
Dec 1969 - Nov 1970	-	11	0
-	Nov 1970 - Nov 1973	36	1
Nov 1973 - Mar 1975	-	16	0
-	Mar 1975 - Jan 1980	58	1
Jan 1980 - Jul 1980	-	6	0
-	Jul 1980 - Jul 1981	12	1
Jul 1981 - Nov 1982	-	16	0
-	Nov 1982 - Jul 1990	92	1
Jul 1990 - Mar 1991	-	8	0
-	Mar 1991 - Mar 2001	120	1
Mar 2001 - Nov 2001	-	8	0
-	Nov 2001 - Dec 2007	73	1
Dec 2007 - Jun 2009	-	18	0
-	Jun 2009 - Feb 2020	128	1
Feb 2020 - Apr 2020	-	2	0
-	Apr 2020 - Nov 2025	67	1

using the abbreviated symbols provided in the table, except for the S&P 500 Index, which is obtained from Yahoo Finance.

Income/Output Variable: GDP is an important factor in the output and income sectors. It serves as a primary quantitative indicator of the business cycle by capturing fluctuations in economic output, income, and spending. Policymakers, businesses, and investors rely on GDP trends to assess the current phase of the cycle and to design appropriate fiscal, monetary, and strategic responses. The quarterly data of GDP is accessed from the Federal Bank of St. Louis and then converted into monthly data by filling forward method in order to maintain the consistent frequency with other input variables.

Labor Market Indicators: The civilian unemployment rate, average unemployment duration, average weekly hours in manufacturing, and industrial production together provide a broad picture of labor market conditions by capturing employment levels, job availability, labor utilization, and production activity. The unemployment rate and duration reflect the availability of jobs and the persistence of joblessness, while average weekly hours indicate changes in labor demand that often occur before hiring or layoffs. Industrial production links labor market conditions to overall economic activity, as changes in output directly influence employment needs. Combined, these indicators effectively represent labor market dynamics across different phases of the business cycle. The data of all of these variables are maintained by the Federal Bank of St. Louis and publicly available in monthly basis.

Housing Market Indicator and Consumer Sentiment: The index of total new privately owned housing units started and the consumer sentiment index are important business cycle predictors because they reflect forward-looking behavior in housing and consumption. Housing starts respond quickly to changes in interest rates, credit conditions, and economic

Table 2: List of input features, sampling frequencies, and their abbreviations.

Sector/Feature	Frequency	Abbreviation
Income/Output		
Gross Domestic Product	Quarterly	GDP
Labor Market		
Civilian Unemployment Rate	Monthly	UNRATE
Average Unemployment Duration	Monthly	UEMPMEAN
Average Weekly Hours (Manufacturing)	Monthly	AWHMAN
Industrial Production: Total Index	Monthly	INDPRO
Housing		
Total New Privately Owned Housing Units Started	Monthly	PERMIT
Consumption, Orders and Inventories		
Consumer Sentiment Index	Monthly	UMCSENT
Interest Rates		
Effective Federal Funds Rate	Monthly	FEDFUNDS
3 Month Treasury Bill Minus Federal Funds Rate	Monthly	TB3SMFFM
10 Year Treasury Constant Maturity Minus Federal Funds Rate	Monthly	T10YFFM
Spread	Monthly	-
Prices		
Consumer Price Index (All Items)	Monthly	CPIAUCSL
Stock Market		
S&P500 Index	Daily	ˆGSPC

expectations, making them a leading signal of expansion or slowdown. Consumer sentiment captures households' confidence about their financial prospects, which strongly influences spending decisions, orders, and inventory levels. Together, these indicators provide early insight into shifts in economic activity and potential turning points in the business cycle. Both variables are extracted from the Federal Bank of St. Louis and they are available in monthly basis.

Interest Rates: To obtain a comprehensive measure of interest rate and exchange rate conditions, we analyzed the effective federal funds rate along with key interest rate spreads, including the 3-month Treasury bill minus the federal funds rate, the 10-year Treasury constant maturity minus the federal funds rate, and the term spread. These variables capture both short-term monetary policy stance and the slope of the yield curve, which are closely linked to expectations about future economic activity. The interest rates are extracted from the Federal Bank of St. Louis database and the spread is extracted from the Federal Bank of New York.

Consumer Prices and Broad Stock Market Index: As an indicator of price dynamics, we used the consumer price index (CPI), which reflects inflationary pressures in the economy and plays an important role in the business cycle by influencing real purchasing power, interest rates, and monetary policy decisions. Rising CPI inflation often accompanies periods of economic expansion, when demand strengthens, while slowing or declining inflation is typically associated with economic downturns. To capture financial market performance and investor expectations,

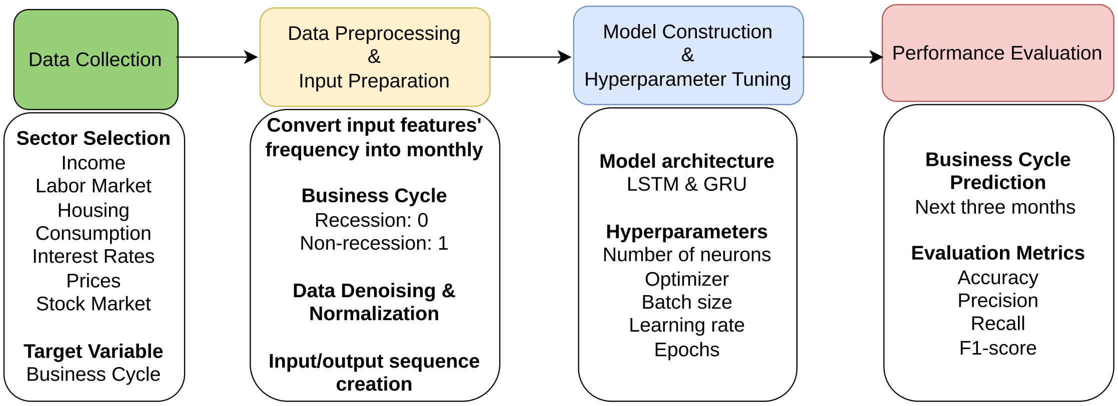

Figure 3: Experimental framework for business cycle prediction using LSTM and GRU model architectures.

we utilized the S&P 500 stock index, which is a forward-looking indicator that reflects market participants' expectations about future corporate earnings and economic conditions. Stock market movements often lead changes in real economic activity, making the S&P 500 a valuable predictor of turning points in the business cycle. The monthly data of CPI data is extracted from the Federal Bank of St. Louis while the daily S&P 500 index data is extracted from Yahoo Finance and then sampled to monthly data in order to maintain consistency in the combined data.

3 Methods

The modeling approach is based on deep learning which is a type of machine learning that uses artificial neural networks as its core model architecture. It has proven to be a valuable tool for solving various real-world problems, ranging from everyday household problems to security, banking, healthcare, finance, and several other scientific advancements. Deep neural networks are made up of several layers of interconnected computational nodes, also called neurons, where each node is responsible for learning a specific feature of the data, mirroring the functioning of a human brain. It is constructed by incorporating multiple hidden layers with several neurons, which makes it capable of learning complex patterns and nonlinear relationships of the underlying data. The effectiveness of deep neural network models has elevated performance to human levels, and in some cases even surpassed them, leading to their widespread utilization in various practical applications (Silver et al., 2017; Esteva et al., 2017; Zhou and Troyanskaya, 2015).

Figure 3 outlines our research methodology for predicting business cycles using LSTM and GRU models. The process begins with data collection from various sectors, including income, labor market, housing, consumption, interest rates, prices, and stock market data, with the business cycle as the target variable, classified as either recession(0) or non-recession (1). The data is then preprocessed by taking monthly feature values, denoising, normalizing the values, and creating input-output sequences for model training. In the model construction phase, LSTM and GRU architectures are employed, and hyperparameters such as the number of units, optimizer, batch size, learning rate, and epochs are tuned to optimize performance. Finally, the framework evaluates the models' ability to predict the business cycle over the next three months—using performance metrics such as accuracy, precision, recall, and F1-score. This structured approach ensures robust time-series analysis and reliable prediction of economic trends.

3.1 Long Short-Term Memory (LSTM)

LSTM stands as a prevalent deep learning method within Recurrent Neural Networks (RNNs) for predicting time series data. While conventional RNNs excel in retaining information in the short term, they often struggle with learning long-term dependencies due to the issue of vanishing gradients. This issue occurs when gradients in deep neural networks become extremely small during training, causing earlier layers to learn very slowly or not at all. LSTM addresses this challenge by employing memory cells, which effectively combat the problem of vanishing gradients. Its architecture comprises an input layer, a hidden layer, a cell state, and an output layer. The pivotal element of LSTM architecture is the cell state, which traverses the network chain with linear interactions, thus ensuring the preservation of information flow. Through a gate mechanism, LSTM selectively maintains or modifies information within the cell state, facilitated by components such as the sigmoid layer, hyperbolic tangent layer, and pointwise multiplication operation (Hochreiter and Schmidhuber, 1997; Atkinson and Shiffrin, 1971; Bhandari et al., 2022; Rimal, 2022).

Let (x_t, y_t) be the pair of input and output of the model where $x_t \in \mathbb{R}^{k\times 1}$ is the input feature and $y_t \in \mathbb{R}$ is the output at time t for $t = 1, 2, \ldots, n$. Here, k and n are the number of input features and observations. Furthermore, to incorporate the time step (m) in LSTM and GRU architectures, input sequence X_t is created by taking m consecutive sequence $x_t : x_{t+m-1}$, which is a matrix of shape $k \times m$ for $t \in \{1, 2, \ldots, n-m-1\}$. The hyperparameter m is chosen during training process via grid search over the given candidate set. The LSTM network works as a memory system with gate mechanism, where gate is a computational node also known as decision step. The forget gate f_t decides which old information is no longer important and should be erased. Then, the input gate i_t looks at the new information coming in and chooses what's worth remembering. The cell's memory c_t is then updated by combining what was decided to keep from the past with the important new details. Finally, the output gate o_t decides which parts of this updated memory to share as the network's output for the current step. This careful filtering at each stage helps LSTMs remember useful patterns for long periods while ignoring irrelevant noise.

At time t, the respective gates and layers compute the following functions:

$$\begin{aligned}
i_t &= \sigma(W_i x_t + W_{hi} h_{t-1} + b_i),\\
f_t &= \sigma(W_f x_t + W_{hf} h_{t-1} + b_f),\\
o_t &= \sigma(W_o x_t + W_{ho} h_{t-1} + b_o),\\
\tilde{c_t} &= \tanh(W_c x_t + W_{hc} h_{t-1} + b_c),\\
c_t &= f_t \otimes c_{t-1} + i_t \otimes \tilde{c_t},\\
h_t &= o_t \otimes \tanh(c_t),
\end{aligned}$$

where σ and $\tanh$ represent the sigmoid and hyperbolic tangent activation functions respectively, the operator $\otimes$ is the element-wise product, $W \in \mathbb{R}^{d\times k}$, $W_h \in \mathbb{R}^{d\times d}$ are weight matrices, and $b \in \mathbb{R}^{d\times 1}$ is the bias vector. Moreover, d is the number of neurons in the hidden layer. The output h_t of LSTM is a feature representation of the input sequence X_t at time t, which can be expressed as $(h_t, c_t) = LSTM(X_t, h_{t-1}, c_{t-1}, w)$.

3.2 Gated Recurrent Unit (GRU)

GRU is a simplified variant of the LSTM model architecture. Compared to LSTM, the GRU simplifies the gate mechanism by combining the forget and input gates into a single update gate and merging the cell state and hidden state, which reduces complexity and computation while often performing similarly (Cho, 2014; Pokhrel et al., 2022; Rimal, 2022).

At time t, the respective gates and layers compute the following functions:

$$\begin{aligned} u_t &= \sigma(W_u x_t + W_{uh} h_{t-1} + b_u), \\ r_t &= \sigma(W_r x_t + W_{rh} h_{t-1} + b_r), \\ c_t &= \tanh(W_c x_t + W_{ch}(r_t \otimes h_{t-1}) + b_c), \\ h_t &= (1 - u_t) \otimes h_{t-1} + c_t \otimes u_t, \end{aligned}$$

where $W_u, W_r, W_c \in \mathbb{R}^{d \times k}$, $W_{uh}, W_{rh}, W_{ch} \in \mathbb{R}^{d \times d}$ are weight matrices, and $b_u, b_r, b_c \in \mathbb{R}^{d \times 1}$ are bias vectors. The output h_t of GRU is a feature representation of the input sequence X_t at time t and is calculated as $h_t = GRU(X_t, h_{t-1}, w)$.

3.3 Optimizers and Performance Metrics

In deep learning, optimizers play a crucial role in efficiently updating model parameters to minimize the loss function. The models were trained using the weighted binary cross-entropy loss function, which quantifies the difference between the predicted probabilities and the true binary labels (Mao et al., 2023). This study have used three optimizers during the model training. Adam combines the benefits of momentum and adaptive learning rates by maintaining estimates of both first and second moments of the gradients, making it one of the most widely used optimizers in practice (Kinga and Ba Adam, 2014). Nadam extends Adam by incorporating Nesterov Accelerated Gradient into its momentum term, which helps in achieving faster convergence and better generalization in certain tasks (Dozat, 2016). On the other hand, Adagrad adapts the learning rate for each parameter based on the history of gradients (Duchi et al., 2011). Together, these optimizers provide flexible options for training deep neural networks depending on the data structure and optimization challenges. In the LSTM and GRU layers implemented in Keras, the network parameters were initialized to promote stable training and mitigate vanishing or exploding gradients. The input weights were initialized using the Glorot uniform method and all biases were set to zero (Hochreiter and Schmidhuber, 1997; Cho, 2014).

Performance metrics are quantitative measures that evaluate how effectively a classification model distinguishes between different classes. Accuracy, precision, recall, and F1-score are used to evaluate the performance of the implemented models (Kelleher et al., 2020). In the context of business cycle classification, where 0 represents recession and 1 represents non-recession, performance metrics capture different dimensions of model effectiveness. Accuracy reflects the overall proportion of correctly classified periods; however, it can be misleading because non-recession periods typically dominate the data. As a result, a model may achieve high accuracy by predominantly predicting the non-recession class. Precision assesses the reliability of non-recession (1) predictions, which is important when false positives—predicting non-recession during an actual recession—could delay timely policy interventions. Recall, in contrast, evaluates the model's ability to correctly identify recession (0) periods, a critical consideration since failing to detect a recession (false negatives) may lead to severe economic misjudgments. The F1-score provides a balanced measure by jointly considering precision and recall, ensuring that the model does not favor overall accuracy at the expense of correctly identifying recessions. Collectively, these metrics offer a more nuanced evaluation than accuracy alone, enabling a clearer assessment of trade-offs between false positives and false negatives in business cycle classification.

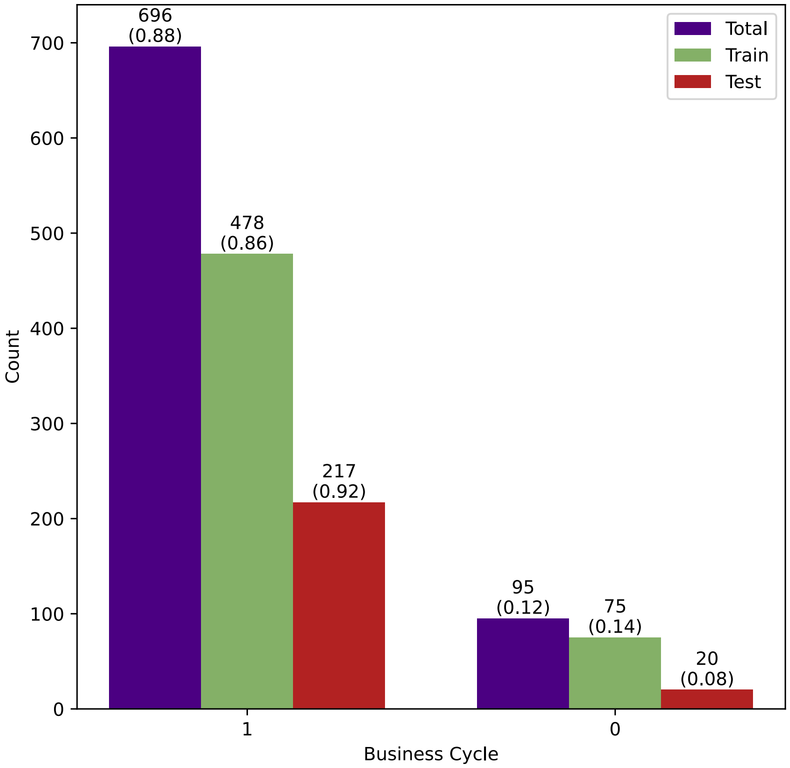

Figure 4: Distribution of recession (0) and non-recession (1) observations in the full dataset, training set, and test set.

4 Experimental Design and Results

We conducted our experiments in several stages. First, the dataset was split into train and test sets using different test sizes: 10%, 15%, 20%, 25%, and 30%. Each split was evaluated to check how well the class distributions (recession (0) and non-recession (1)) were preserved in both sets. The best balance was obtained with a 30% test size. In this split, the train set contains the first 553 months of data (January 1960 to February 2006), while the test set includes the last 238 months (March 2006 to November 2025). The class distributions of 1's and 0's are 88% and 12% in the full dataset, 86% and 14% in the train set, and 92% and 8% in the test set, as shown in Figure 4.

4.1 Input Preparation

The input features included in this study are highly volatile, we implemented the singular spectrum analysis to denoise the input features (Sella et al., 2016). Since the input features had values on very different scales, we applied min–max normalization defined as $x_{scaled} = \dfrac{x - x_{min}}{x_{max} - x_{min}}$ to bring them into a consistent range [0, 1], where x_{min} and x_{max} are minimum and maximum value in the data. After normalization, the data were represented as a 2D array with dimensions corresponding to the number of observations and features. However, the LSTM and GRU model require inputs in 3D shape, the data were reshaped into a 3D array in the form of number of observations, time steps, and number of features.

Table 3: List of optimal hyperparameters for the respective model configurations.

No. of neurons	LSTM		GRU	
	Optimizer	Batch size	Optimizer	Batch size
4	Adam	16	Nadam	64
8	Adam	8	Nadam	64
16	Adam	8	Nadam	8
32	Adam	8	Nadam	16
64	Adam	16	Nadam	64
128	Adam	8	Adagrad	64
256	Adam	32	Adam	8
512	Adagrad	8	Adam	16
750	Adagrad	64	Nadam	8
1024	Adagrad	64	Adagrad	64
1250	Adagrad	64	Adagrad	8

4.2 Model Construction and Hyperparameter Tuning

We designed single-layer LSTM and GRU models with 11 different configurations, using neuron counts of 4, 8, 16, 32, 64, 128, 256, 512, 750, 1024, and 1250 in the first layer to capture a wide range of model complexities. Although we also experimented with multi-layer architectures, their performance was inferior, consistent with findings from earlier research (Bhandari et al., 2022; Pokhrel et al., 2022; Bhandari et al., 2024; Rimal et al., 2024). Each model was trained on input sequences of varying time steps (m), namely 3, 6, 12, 24, and 36 months, where the time step specifies the length of past observations used to forecast the next value. Empirical evaluation indicated that a time step of six months yielded the most effective performance.

Several hyperparameters, which govern the learning process but are not directly learned from the data, require systematic tuning. Specifically, we considered three optimizers: Adam, Adagrad, and Nadam which update network weights during training, and four batch sizes: 8, 16, 32, and 64 representing the number of samples processed before each weight update. In total, this resulted in $11 \times 3 \times 4 = 132$ unique configurations for each model architecture. To further enhance training stability, we employed an exponential decay learning rate schedule, initialized at 0.01 with a decay factor of 0.95, rather than a fixed learning rate. Each configuration was trained for up to 200 epochs, where an epoch corresponds to one complete pass through the training dataset. Early stopping with a patience of five epochs was applied to prevent overfitting, and to mitigate stochastic variability, each experiment was replicated ten times. The optimal hyperparameters were selected based on the highest average validation accuracy across these replications. Table 3 summarizes the optimal hyperparameters identified for each model configuration. For every model size, defined by the number of neurons, the table reports the optimizer and batch size that produced the best predictive performance for both LSTM and GRU model architectures.

4.3 Model Performance Evaluation

After selecting the best hyperparameters for each model, we trained each configuration with 30 replications over 200 epochs to account for the inherent randomness in model initialization. In particular, the randomness in the LSTM/GRU model arises from the random initialization

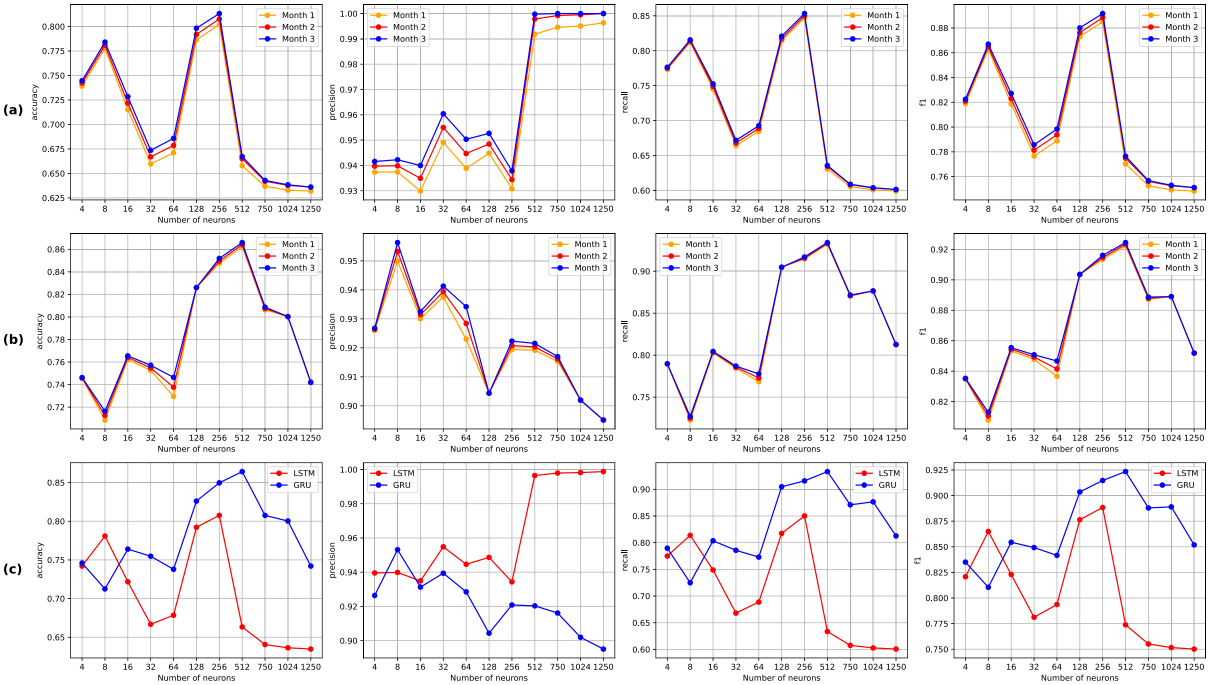

Figure 5: Multi-step overall average performance scores (accuracy, precision, recall, and F1) obtained from 30 replicates on the test data for (a) LSTM, (b) GRU, and (c) best-performing model from LSTM and GRU models.

of the weight matrices, which can lead to variations in performance across training runs. Each weight matrix in the model is initialized using the Glorot uniform initializer, a variant of the uniform distribution designed to maintain a consistent variance of activations across layers. By repeating the training process multiple times, we obtained more stable and reliable estimates of the model's performance.

The performance of individual predictions for the next three months from each model architecture on the test set is shown in Figure 5 (a)-(b). These figures indicate that both models with fewer neurons performed slightly better for the 3rd-month prediction, whereas larger neuron configurations exhibited similar performance levels overall.

We further calculated each performance metric's average and standard deviation on the test data. Table 4 summarizes the average percentage ± standard deviation from 30 replicates for each metric for the given models. The graphical representation of the average values in the above table is presented in Figure 5 (c). The Table 4 and Figure 5 (c) show a clear capacity–performance trade-off for both LSTM and GRU architectures as the number of neurons increases. For LSTM, performance improves as model capacity increases from very small configurations (4–16 neurons) to moderate sizes (128–256 neurons), where the highest accuracy (80.76%) and F1-score (88.84%) are achieved with relatively low variability. Beyond this range, further increases in neurons lead to a systematic degradation in accuracy, recall, and F1-score, despite precision approaching near-perfect values. This pattern indicates over-parameterization, where the model becomes overly conservative—predicting positives with extremely high precision but missing a substantial portion of true positives, as reflected by the sharp decline in recall. For GRU, performance exhibits a more stable and monotonic improvement with increasing neurons up to 512 neurons,

Table 4: Overall average ± standard deviation of performance scores (accuracy, precision, recall, and F1) obtained from 30 replicates for each model architecture on the test data.

Models	Neurons	Accuracy	Precision	Recall	F1-score
LSTM	4	74.20 ± 19.29	93.96 ± 3.62	77.49 ± 24.04	82.08 ± 18.39
	8	78.09 ± 12.22	93.98 ± 3.21	81.38 ± 14.38	86.49 ± 8.49
	16	72.17 ± 12.06	93.49 ± 2.01	74.89 ± 14.81	82.29 ± 9.22
	32	66.68 ± 8.20	95.48 ± 1.89	66.80 ± 10.14	78.12 ± 6.63
	64	67.85 ± 6.77	94.46 ± 1.05	68.87 ± 8.23	79.37 ± 5.39
	128	79.22 ± 5.67	94.87 ± 1.98	81.75 ± 6.83	87.64 ± 3.65
	256	80.76 ± 5.24	93.43 ± 1.90	85.01 ± 6.89	88.84 ± 3.45
	512	66.34 ± 3.68	99.65 ± 0.32	63.36 ± 4.10	77.39 ± 3.02
	750	64.06 ± 1.74	99.79 ± 0.23	60.76 ± 1.94	75.52 ± 1.44
	1024	63.64 ± 1.19	99.82 ± 0.20	60.29 ± 1.40	75.16 ± 1.02
	1250	63.47 ± 0.52	99.88 ± 0.13	60.07 ± 0.61	75.02 ± 0.46
GRU	4	74.59 ± 16.25	92.64 ± 4.20	78.97 ± 19.94	83.51 ± 13.45
	8	71.25 ± 13.65	95.32 ± 3.52	72.49 ± 17.00	81.05 ± 10.92
	16	76.40 ± 12.21	93.13 ± 3.61	80.36 ± 15.08	85.44 ± 8.24
	32	75.48 ± 8.93	93.94 ± 3.03	78.57 ± 12.17	84.93 ± 6.32
	64	73.79 ± 5.69	92.85 ± 1.25	77.30 ± 7.11	84.16 ± 4.28
	128	82.61 ± 6.23	90.43 ± 0.63	90.48 ± 6.82	90.35 ± 3.68
	256	84.97 ± 7.60	92.08 ± 2.08	91.58 ± 10.03	91.47 ± 5.06
	512	**86.42 ± 8.04**	92.03 ± 1.75	**93.32 ± 10.20**	**92.34 ± 5.09**
	750	80.75 ± 9.20	91.61 ± 2.52	87.10 ± 11.74	88.79 ± 6.29
	1024	80.03 ± 1.47	90.20 ± 0.16	87.65 ± 1.61	88.90 ± 0.91
	1250	74.20 ± 1.11	89.51 ± 0.14	81.27 ± 1.21	85.19 ± 0.73

which yields the best overall performance (accuracy: 86.42%, recall: 93.32%, F1-score: 92.34%). Compared to LSTM, GRU benefits more consistently from increased capacity, achieving a better balance between precision and recall. However, similar to LSTM, performance declines when the network becomes excessively large ($\geqslant 750$ neurons), again suggesting diminishing returns and possible overfitting. Overall, the results indicate that moderate-to-large architectures are optimal, with GRU outperforming LSTM across all metrics at their respective optimal sizes. Excessively large models do not improve generalization and instead reduce recall and overall predictive balance.

In summary, the GRU model with 512 neurons performed the best overall based on accuracy, recall, and F1-score among all architectures and model configurations tested. The best GRU model is able to capture the upward economy perfectly. However, the model was less effective at identifying a downward trend. One of the reasons could be the imbalanced proportions of the class labels in the training data. To forecast recessions for the next three years, we retrained the best-performing model using all available data. Based on 30 replicates of this model, the results indicate a high probability (average ± standard deviation: $93.38\% \pm 5.41\%$) that no recession is expected during the next three years.

5 Conclusion

This study demonstrates that well-structured deep learning models can effectively capture complex patterns in business cycle dynamics. By combining a compact yet diverse set of macroeconomic indicators across key economic sectors, the proposed framework balances information richness with model stability, reducing noise while preserving meaningful temporal and cross-sectoral signals.

Experimental results indicate that GRU-based models with 512 neurons consistently achieve the strongest performance across multiple evaluation metrics when assessed over 30 independent training replications. Extensive hyperparameter exploration across 264 model instances enabled robust parameter selection and supports the reproducibility of the results. These findings reinforce the suitability of GRU architectures for modeling sequential dependencies in macroeconomic time-series data. The study places strong emphasis on experimental rigor through systematic evaluation of multiple model configurations, repeated training runs, and performance averaging, which together mitigate the stochastic effects inherent in deep learning models.

The framework explicitly addresses challenges common in macroeconomic datasets, including class imbalance and regime shifts during systemic crises. Training on data that include major historical events, such as the Global Financial Crisis and the COVID-19 pandemic, improves robustness. However, the results also indicate that purely data-driven models may have limited generalization when exposed to future crises with unseen structural characteristics. Overall, this work presents a reproducible and adaptable deep learning framework for business cycle forecasting. The approach contributes empirical evidence to the growing body of machine learning research on macroeconomic prediction and offers practical value for decision-making applications. In future, the study may focus on multi-scale temporal modeling utilizing transformer-based architectures, time-series data augmentation, and hybrid optimization strategies to further improve performance and generalization under rare and extreme economic conditions.

6 Ethics and Implications

The study follows established ethical standards by emphasizing transparency, fairness, and reproducibility. Publicly available datasets and open-source Python libraries are used, enabling independent evaluation and replication of the proposed methodology. Model performance is reported as the average of multiple out-of-sample replications, thereby improving the robustness and reliability of the findings. While the results offer useful insights for investment-related analysis, business cycle dynamics can be influenced by numerous external and uncertain factors beyond model outputs. Therefore, stakeholders are encouraged to complement the model's predictions with real-time data, broader economic indicators, and domain-specific judgment when making informed decisions.

Acknowledgments

The authors would like to thank Shane L. Fabbri for helping with data collection and preprocessing.

Funding

No funding was received to assist with the preparation of this manuscript.

Conflicts of Interest

We confirm that there are no known conflicts of interest with this publication.

Data and Code Availability

Data and codes will be made available upon request with corresponding author.

References

Asada T, Chiarella C, Flaschel P (2003). Keynes-Metzler-Goodwin model building: The closed economy. UTS School of Finance and Economics Working Paper, (124).

Asada T, Chiarella C, Flaschel P, Mouakil T, Proano C, Semmler W (2011). Stock-flow interactions, disequilibrium macroeconomics and the role of economic policy. *Journal of Economic Surveys*, 25(3): 569–599. https://doi.org/10.1111/j.1467-6419.2010.00661.x

Atkinson RC, Shiffrin RM (1971). The control of short-term memory. *Scientific American*, 225(2): 82–91. https://doi.org/10.1038/scientificamerican0871-82

Bessec M, Bouabdallah O (2015). Forecasting GDP over the business cycle in a multi-frequency and data-rich environment. *Oxford Bulletin of Economics and Statistics*, 77(3): 360–384. https://doi.org/10.1111/obes.12069

Beutel J, List S, von Schweinitz G (2019). Does machine learning help us predict banking crises? *Journal of Financial Stability*, 45: 100693. https://doi.org/10.1016/j.jfs.2019.100693

Bhandari HN, Pokhrel NR, Rimal R, Dahal KR, Rimal B (2024). Implementation of deep learning models in predicting ESG index volatility. *Financial Innovation*, 10(1): 75. https://doi.org/10.1186/s40854-023-00604-0

Bhandari HN, Rimal B, Pokhrel NR, Rimal R, Dahal KR (2022). LSTM-SDM: An integrated framework of LSTM implementation for sequential data modeling. *Software Impacts*, 14: 100396. https://doi.org/10.1016/j.simpa.2022.100396

Bianchi F, Nicolò G, Song D (2023). Inflation and real activity over the business cycle (No. w31075). National Bureau of Economic Research. url: http://www.nber.org/papers/w31075

Bluwstein K, Buckmann M, Joseph A, Kapadia S, Şimşek Ö (2023). Credit growth, the yield curve and financial crisis prediction: Evidence from a machine learning approach. *Journal of International Economics*, 145: 103773.

Buser T, Niederle M, Oosterbeek H (2024). Can competitiveness predict education and labor market outcomes? evidence from incentivized choice and survey measures. *Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_01439

Cacciatore M, Ghironi F (2021). Trade, unemployment, and monetary policy. *Journal of International Economics*, 132: 103488. https://doi.org/10.1016/j.jinteco.2021.103488

Cacciatore M, Ravenna F (2021). Uncertainty, wages and the business cycle. *The Economic Journal*, 131(639): 2797–2823. https://doi.org/10.1093/ej/ueab019

Chatzis SP, Siakoulis V, Petropoulos A, Stavroulakis E, Vlachogiannakis N (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications*, 112: 353–371. https://doi.org/10.1016/j.eswa.2018.06.032

Chiarella C, Flaschel P (2000). *The Dynamics of Keynesian Monetary Growth: Macro Foundations*. Cambridge University Press.

Chiarella C, Flaschel P, Franke R (2005). *Foundations for Disequilibrium Theory of the Business Cycle*. Cambridge University Press.

Cho K (2014). On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint: https://arxiv.org/abs/1409.1259.

Chodorow-Reich G, Wieland J (2020). Secular labor reallocation and business cycles. *Journal of Political Economy*, 128(6): 2245–2287. https://doi.org/10.1086/705717

Di Guilmi C, Galanis G, Proaño CR (2023). A baseline model of behavioral political cycles and macroeconomic fluctuations. *Journal of Economic Behavior & Organization*, 213: 50–67. https://doi.org/10.1016/j.jebo.2023.05.041

Dozat T (2016). Incorporating Nesterov momentum into adam. International Conference on Learning Representations (ICLR).

Duchi J, Hazan E, Singer Y (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(7): 2121–2159.

Ehrenberg R, Smith R, Hallock K (2021). *Modern Labor Economics: Theory and Public Policy*. Routledge.

Emsbo-Mattingly L, Dirk H, Lund-Wilde J (2014). The business cycle approach to equity sector investing, *Technical report, Fidelity Investments (AAR)*.

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, ..., Thrun S (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639): 115–118. https://doi.org/10.1038/nature21056

Fernández-Villaverde J, Guerrón-Quintana PA (2020). Uncertainty shocks and business cycle research. *Review of Economic Dynamics*, 37: S118–S146. https://doi.org/10.1016/j.red.2020.06.005

Fouliard J, Howell M, Rey H, Stavrakeva V (2021). Answering the queen: Machine learning and financial crises (No. w28302). National Bureau of Economic Research. https://doi.org/10.3386/w28302

Goulet Coulombe P, Leroux M, Stevanovic D, Surprenant S (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5): 920–964. https://doi.org/10.1002/jae.2910

Graff M (2011). International business cycles: How do they relate to Switzerland? (No. 291) KOF Working Papers. url: https://hdl.handle.net/10419/54715 https://doi.org/10.3929/ethz-a-006742181

Hall AS (2018). Machine learning approaches to macroeconomic forecasting. *The Federal Reserve Bank of Kansas City Economic Review*, 103(63): 63–81.

Hamilton JD (2005). What's real about the business cycle? National Bureau of Economic Research. url: https://www.nber.org/papers/w11161 https://doi.org/10.3386/w11161

Hochreiter S Schmidhuber J (1997). Long short-term memory. *Neural Computation MIT-Press.* 9(8): 1735–1780.

Hopp D (2022). Economic nowcasting with long short-term memory artificial neural networks (LSTM). *Journal of Official Statistics*, 38(3): 847–873. https://doi.org/10.2478/jos-2022-0037

Kelleher JD, Mac Namee B, D'Arcy A (2020). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies.* MIT press.

King RG, Rebelo ST (1999). Resuscitating real business cycles. *Handbook of Macroeconomics*, 1: 927–1007. https://doi.org/10.1016/S1574-0048(99)10022-3

King RG, Watson MW (1996). Money, prices, interest rates and the business cycle. *Review of Economics and Statistics*, 78(1): 35–53. https://doi.org/10.2307/2109846

Kinga D, Ba Adam J (2015). A method for stochastic optimization. International Conference on Learning Representations (ICLR), 5(6).

Korobilis D (2018). Machine learning macroeconometrics: A primer. Working Paper. Essex Finance Centre Working Papers, Colchester. url: https://repository.essex.ac.uk/22666/

Krugman P (1993). Lessons of Massachusetts for EMU. (F Torres, F Giavazzi, eds.) 241–266. Cambridge University Press.

Kydland FE, Prescott EC (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 50(6) 1345–1370. https://doi.org/10.2307/1913386

Liu L, Chen C, Wang B (2022). Predicting financial crises with machine learning methods. *Journal of Forecasting*, 41(5): 871–910. https://doi.org/10.1002/for.2840

Long Jr JB , Plosser CI (1983). Real business cycles. *Journal of Political Economy*, 91(1): 39–69. https://doi.org/10.1086/261128

Ma Y, Zhang J (2016). Financial cycle, business cycle and monetary policy: Evidence from four major economies. *International Research Journal of Finance and Economics*, 21(4): 502–527. https://doi.org/10.1002/ijfe.1566

Mao A, Mohri M, Zhong Y (2023). Cross-entropy loss functions: Theoretical analysis and applications. In: *International Conference on Machine Learning*, 23803–23828. pmlr.

McCracken MW, Ng S (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4): 574–589. https://doi.org/10.1080/07350015.2015.1086655

Mitchell WC (2023). *Business Cycles and Their Causes*. Univ of California Press.

NBER (2024). U.S. business cycle expansions and contractions. Technical report. *National Bureau of Economic Research*. Accessed July 15, 2024.

Nosratabadi S, Mosavi A, Duan P, Ghamisi P, Filip F, . . . , Gandomi AH (2020). Data science in economics: Comprehensive review of advanced machine learning and deep learning methods. *Mathematics*, 8(10): 1799. https://doi.org/10.3390/math8101799

Paruchuri H (2021). Conceptualization of machine learning in economic forecasting. *Asian Business Review*, 11(1): 51–58. https://doi.org/10.18034/abr.v11i2.532

Pokhrel NR, Dahal KR, Rimal R, Bhandari HN, Khatri RK, . . . , Hahn WE (2022). Predicting nepse index price using deep learning models. *Machine Learning with Applications*, 9: 100385. https://doi.org/10.1016/j.mlwa.2022.100385

Rimal B (2022). Financial time-series analysis with deep neural networks (Doctoral dissertation, Florida Atlantic University). Available from ProQuest Dissertations & Theses Global. (2719490547). https://esearch.ut.edu/login?url=https://www.proquest.com/dissertations-theses/financial-time-series-analysis-with-deep-neural/docview/2719490547/se-2

Rimal R, Rimal B, Bhandari HN, Pokhrel NR, Dahal KR (2025). Real estate market prediction using deep learning models. *Annals of Data Science*, 12(4): 1113–1156.

Sella L, Vivaldo G, Groth A, Ghil M (2016). Economic cycles and their synchronization: A comparison of cyclic modes in three European countries. *Journal of Business Cycle Research*, 12(1): 25–48. https://doi.org/10.1007/s41549-016-0003-4

Sezer OB, Gudelek MU, Ozbayoglu AM (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90: 106181. https://doi.org/10.1016/j.asoc.2020.106181

Siami-Namini S, Namin AS (2018). Forecasting economics and financial time series: Arima vs.

LSTM. arXiv preprint: https://arxiv.org/abs/1803.06386.

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, ..., Chen Y (2017). Mastering the game of go without human knowledge. *Nature*, 550(7676): 354–359. https://doi.org/10.1038/nature24270

Sołtysiak D (2023). On the stability of a certain Keynes-Metzler-Goodwin monetary growth model. *Economics and Business Review*, 9(1): 26–64. https://doi.org/10.18559/ebr.2023.1.2

Soybilgen B (2020). Identifying us business cycle regimes using dynamic factors and neural network models. *Journal of Forecasting*, 39(5): 827–840. https://doi.org/10.1002/for.2658

Tehranian K (2023). Can machine learning catch economic recessions using economic and market sentiments? arXiv preprint: https://arxiv.org/abs/2308.16200.

Thorsrud LA (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, 38(2): 393–409. https://doi.org/10.1080/07350015.2018.1506344

Venkateswarlu Y, Baskar K, Wongchai A, Gauri Shankar V, Paolo Martel Carranza C, ..., Murali Dharan AR (2022). An efficient outlier detection with deep learning-based financial crisis prediction model in big data environment. *Computational Intelligence and Neuroscience*, 2022(1): 4948947. https://doi.org/10.1155/2022/4948947

Valaitis V, Villa AT (2024). Machine learning projection methods for macro-finance models. *Quantitative Economics*, 15(1): 145–173.

Vrontos SD, Galakis J, Vrontos ID (2021). Modeling and predicting us recessions using machine learning techniques. *International Journal of Forecasting*, 37(2): 647–671. https://doi.org/10.1016/j.ijforecast.2020.08.005

Wasserbacher H, Spindler M (2022). Machine learning for financial forecasting, planning and analysis: Recent developments and pitfalls. *Digital Finance*, 4(1): 63–88. https://doi.org/10.1007/s42521-021-00046-2

Weinstock LR (2024). Introduction to U.S. economy: The business cycle and growth, Congressional Research Service (CRS) in Focus. url: https://www.congress.gov/crs-product/IF10411

Woitek U (1993). *The G7-Countries: A Multivariate Description of the Business Cycle Stylized Facts Volkswirtschaftl. Fakultät D.* Ludwig-Maximilians-Univ., München.

Yoshida H, Asada T (2007). Dynamic analysis of policy lag in a Keynes–Goodwin model: Stability, instability, cycles and chaos. *Journal of Economic Behavior & Organization*, 62(3): 441–469. https://doi.org/10.1016/j.jebo.2004.10.014

Zheng Y, Xu Z, Xiao A (2023). Deep learning in economics: A systematic and critical review. *Artificial Intelligence Review*, 56(9): 9497–9539. https://doi.org/10.1007/s10462-022-10272-8

Zhou J, Troyanskaya OG (2015). Predicting effects of noncoding variants with deep learning–based sequence model. *Nature Methods*, 12(10): 931–934. https://doi.org/10.1038/nmeth.3547