Forecasting Unemployment Using Long Short-Term Memory Networks

Tristin Thilmany

CONTENTS

***Abstract*—Abstract**

I. INTRODUCTION

The health of the global economy is closely tied to the dynamics of labor markets, where the unemployment rate serves as a key indicator. The unemployment rate represents the percentage of the labor force that is actively seeking employment but unable to find it [15]. In the context of evolving economic landscapes, accurately forecasting unemployment rates is essential for decision making. Governments, businesses, and individuals rely on these forecasts to anticipate labor market conditions, form policy, and make choices. However, traditional methods for unemployment forecasting often fail to capture non-linear, economic relationships. For years, various statistical techniques have been used to predict unemployment rates. One primary method is the autoregressive integrated moving average (ARIMA) model. ARIMA models have been among the most popular models used in univariate, or single variable, forecasting. Although these methods can be useful, they often struggle to capture these complex relationships. [7] Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), better predicts these non-linear relationships that influence unemployment rates. By using LSTM's ability to learn from historical data and identify patterns, we can increase the accuracy and reliability of forecasting unemployment. Earlier, accurate predictions of unemployment changes can contribute to economic stability and enhanced job security.

II. BACKGROUND

A. What is Unemployment?

Unemployment is the number of individuals in the labor force who are actively seeking employment but are unable to find jobs. The unemployment rate is a key indicator of the health of an economy, as it reflects the degree to which people are working. High unemployment rates can lead to social and economic issues. As a consistently measured metric, changes in the unemployment rate serve as a key indicator of general economic trends. [16] The unemployment rate is calculated as the percentage of the labor force that is unemployed:

$$Unemployment\ Rate = \frac{Unemployed}{Labor\ Force} \cdot 100 \qquad (1)$$

where $Unemployed$ is the number of unemployed people, representing individuals who are actively seeking employment but are currently without a job. $LaborForce$ is the labor force includes all employed and unemployed people. [15].

There are three main types of unemployment: cyclical, structural, and frictional. Cyclical unemployment, tied to economic changes, can be predicted using LSTM neural networks making them a valuable tool for policymakers to anticipate and address unemployment cycles. [10]

B. Time Series Forecasting

A time series is a sequence of data points collected over time, typically at regular intervals. [11] Time series can be univariate, with one data point per time interval, or multivariate, with multiple data points per time interval. Time series data often exhibit distinct characteristics that are important to consider when forecasting. These characteristics include:

- **Trend:** The overall direction of the series, such as an upward or downward trend.
- **Seasonality:** Patterns that repeat at regular intervals, such as daily, weekly, or yearly cycles.
- **Cyclicality:** Patterns that occur over longer, irregular periods, without a fixed frequency.

- **Irregular Component:** The remaining data outside the trend that represent random fluctuations or noise. [11]

III. Technical Analysis of LSTMs

A. Neural Networks

Neural networks are a class of machine learning algorithms inspired by the structure and function of the human brain. They consist of interconnected nodes organized in layers. Each connection between neurons has a weight associated with it, which determines the strength of the connection. A typical neural network has an input layer, one or more hidden layers, and an output layer. The input layer receives the initial data, and the output layer produces a result. The hidden layers perform computations on the data and learn to recognize patterns. Neural networks learn by adjusting the weights of the connections between neurons. This process, called training, involves feeding the network with a large dataset and iteratively adjusting the weights until the network can accurately predict the desired output. The trained network can then be used to make predictions on new, unseen data. [2]

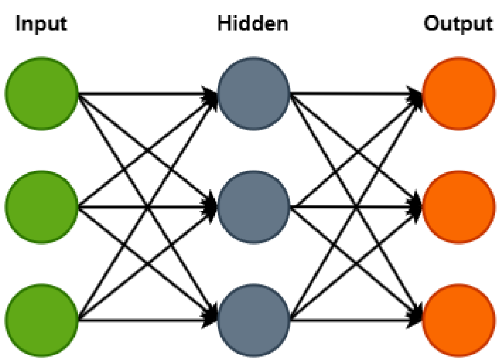

Fig. 1. A basic structure of an Artificial Neural Network. Connections between neurons have associated weights, which determine the strength of the connection. During the learning process, the network adjusts these weights to optimize its performance. Neural Networks learn by adjusting the weights and biases of the connections between neurons to minimize the difference between its predictions and the actual target values. [1]

Neural Networks are typically constructed of 3 layers:

- **Input Layer:** This layer receives the initial data or signals for the network to process. Each node in this layer represents a feature or input variable.
- **Hidden Layers:** These intermediate layers perform computations on the data received from the input layer. Each node in a hidden layer applies an activation function to a weighted sum of its inputs, introducing non-linearity into the model.
- **Output Layer:** This layer produces the final results or predictions based on the processed information from the hidden layers. The number of nodes in this layer depends on the specific task, such as classification or regression.

B. Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are designed to process sequential data by maintaining a memory of past inputs. However, traditional RNNs face a significant challenge when dealing with long-term dependencies: the vanishing and exploding gradient problem. This is when the model either overfits the training data and fails to predict real-world data, or under-fits and cannot accurately predict any data.

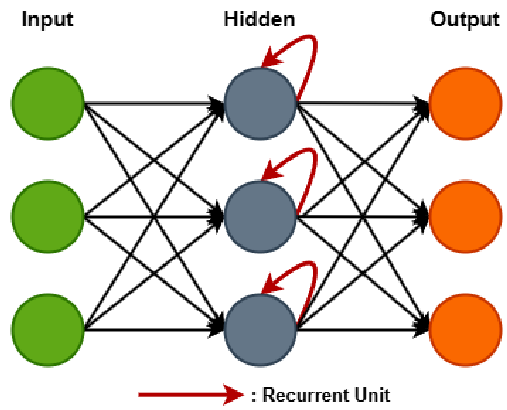

Fig. 2. Structure of a Recurrent Neural Network, showing the recurrent connection of the hidden state. These recurrent calls to hidden states differ from traditional neural networks.

C. Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) networks are a specialized type of recurrent neural network designed to overcome the vanishing gradient problem that hinders traditional RNNs. While traditional RNNs have a simple recurrent unit that feeds the hidden state back into the network at the next time step, LSTMs introduce a more complex 'memory cell' structure. This memory cell allows the network to better regulate the flow of information through time, enabling it to learn long-term dependencies more effectively.

The 3 main gates of an LSTM are what control the flow of information within the unit. [13]

- **Forget Gate (f_t):** Controls the flow of data out of the memory cell.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

- **Input Gate (i_t):** Controls the flow of data into the memory cell.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

- **Output Gate (o_t):** Controls the output of the memory cell to the rest of the network.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4}$$

where:

- W_i, W_f, and W_o are the weight matrices for the input, forget, and output gates, respectively.
- b_i, b_f, and b_o are the bias vectors for the respective gates.
- h_{t-1} is the hidden state from the previous time step.
- x_t is the input at the current time step.
- σ is the sigmoid activation function.

The sigmoid function (σ) is a common activation function used in neural networks. It takes any input value and maps it to an output value between 0 and 1. This is useful for representing probabilities or for controlling the flow of information in a network. The formula for the sigmoid function is:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{5}$$

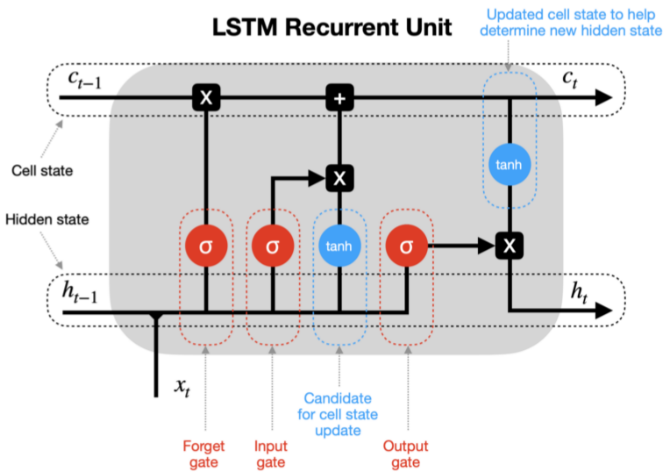

Fig. 3. Structure of an LSTM unit, with its forget, input, and output gates regulating the cell state. Internal computations not only change the information passed on, but also information kept in the LSTM unit's memory cell. [3]

By carefully controlling the flow of information through these gates, LSTMs can effectively mitigate the vanishing gradient problem and predict long-term data points. This makes them particularly well-suited for time series forecasting, where understanding historical trends and patterns is crucial for accurate predictions.

IV. Benefits and Problems of LSTMs for Unemployment Forecasting

A. Benefits

LSTMs can model complex, non-linear relationships between unemployment rates and various influencing factors. LSTMs are found to perform well with long-term data and during periods of *heightened macroeconomic uncertainty* [8]. LSTMs are designed to handle long-term dependencies in time series data, allowing them to capture the impact of past events on future unemployment rates more effectively than traditional RNN models. LSTMs can easily incorporate multiple input features, such as multiple economic indicators and other economic data. LSTMs can adapt to changing patterns in the data, making them potentially more robust to economic shifts and structural changes in the labor market. When compared to alternative models, LSTMs predicted more accurate, long-term results. [6].

B. Problems

LSTMs typically require large amounts of data to train effectively. If the historical unemployment data is limited or noisy, the model's performance will likely suffer. Training LSTMs can be computationally expensive, requiring significant processing power and time, especially for complex models with many layers and units. Furthermore, LSTM performance not only depends on how many data points and variables are included, but it also depends on the data's statistical characteristics. [4]. LSTMs have several hyperparameters that need to be carefully tuned to achieve optimal performance. This tuning process can be time-consuming and requires expertise. LSTMs are often considered "black box" models, meaning it can be difficult to understand exactly how they arrive at their predictions. This lack of interpretability can be a drawback in situations where understanding the underlying drivers of unemployment is important. They are vulnerable to unexpected external shocks or events (e.g., pandemics, sudden policy changes) that are not reflected in the historical data. These events can significantly impact unemployment rates and are difficult for any model to predict accurately.

V. Methodology

A. Data Collection and Preprocessing

LSTM-based forecasting lies in the quality and relevance of the data. Historical unemployment rate data serves as the primary input, often sourced from official government agencies. To increase the model's accuracy, additional economic indicators are incorporated. These may include other relevant macroeconomic variables or financial datasets from the same time periods. Demographic data, such as labor force participation rates and population statistics, can also be valuable. Data preprocessing is a critical step to ensure the data's suitability for LSTM training. This involves handling missing values, which can be addressed through imputation techniques or by excluding incomplete records. Time series data often requires transformation to ensure stationarity, which can be achieved through taking the logarithm of the dataset to stabilize variance. [14] Additional research has also shown that LSTM networks perform well with *wavelet denoising*. This is another process of taking random, or outlier, data points in order to better train the model. [9]

B. Model Training

The training process involves feeding the training data to the LSTM model, calculating the difference between the model's predictions and the actual unemployment rates, and adjusting the model's weights to minimize this difference. One common loss function is Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6}$$

where n is the number of samples, y_i is the actual value, and $\hat{y}_i$ is the predicted value. [12] Values closer to 0 indicate better performance.

C. Forecasting

Once the LSTM model is trained and evaluated, it can be used to generate forecasts of future unemployment rates. The model takes the most recent historical data as input and predicts the unemployment rate for the next time step. This process can be iterated to generate multi-step forecasts. The reliability of the forecasts depends on the model's accuracy and the stability of the underlying economic conditions. Uncertainty in the forecasts can be quantified using techniques such as prediction intervals or ensemble methods, providing a range of possible outcomes.

VI. Future Trends

TBD

VII. Conclusion

Long Short-Term Memory (LSTM) networks present a powerful approach to forecasting unemployment rates. Their ability to model complex, non-linear relationships and long-term dependencies in labor market data makes them particularly well-suited for this task. Unlike traditional methods, LSTMs can effectively predict economic indicators and adapt to changing of labor markets. The architecture of LSTMs, with its gating mechanisms and memory cells, addresses the limitations of standard Recurrent Neural Networks, allowing for improved handling of temporal data. The process of training and evaluating these models, including data preprocessing and performance metric analysis, is crucial for achieving accurate results. While challenges like data requirements, computational expenses, and interpretability exist [5], the potential of LSTMs to provide precise unemployment rate forecasts is substantial.

References

[1] Amazon Web Services, Inc. What is a recurrent neural network?, n.d.

[2] Erik Bollt. How neural networks work: Unraveling the mystery of randomized neural networks for functions and chaotic dynamical systems. *Chaos*, 34(12), December 2024.

[3] Saul Dobilas. Lstm recurrent neural networks – how to teach a network to remember the past. *Towards Data Science*, Feb 2022.

[4] Erdem Doğan. Lstm training set analysis and clustering model development for short-term traffic flow prediction. *Neural Computing and Applications*, 33(15):11175–11188, 2021.

[5] Daniel Hopp. Economic nowcasting with long short-term memory artificial neural networks (lstm), 3 2021.

[6] Ark O. Ifeanyi and Jamie B. Coble. A practical comparison of data-driven prognostics methods for energy systems. *Nuclear Science and Engineering*, 0(0):1–19, 2025.

[7] Joshua Noble. What are arima models? *IBM Think*, May 2024.

[8] Livia Paranhos. Predicting Inflation with Recurrent Neural Networks. Papers 2104.03757, arXiv.org, April 2021.

[9] Jiayu Qiu, Bin Wang, and Changjun Zhou. Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PLoS One*, 15(1):e0227222, 2020.

[10] Reserve Bank of Australia. Unemployment: Its measurement and types, 2023.

[11] Thomas Hanne Ryan Prater and Rolf Dornberger. Generalized performance of lstm in time-series forecasting. *Applied Artificial Intelligence*, 38(1):2377510, 2024.

[12] Tomasz Rymarczyk, Monika Kulisz, and Grzegorz Kłosowski. Influence of loss function on training the lstm network in wall moisture tomography. *International Journal of Applied Electromagnetics and Mechanics*, 73(4):353–367, 2023.

[13] Wei Song, Chao Gao, Yue Zhao, and Yandong Zhao. A time series data filling method based on lstm-taking the stem moisture as an example. *Sensors*, 20(18):5045, 2020.

[14] C Tamilselvi, Ranjit Kumar Paul, Md Yeasin, and A K Paul. Novel wavelet-LSTM approach for time series prediction. *Neural Comput. Appl.*, nov 2024.

[15] U.S. Bureau of Labor Statistics. Labor force statistics from the current population survey. https://www.bls.gov/cps/.

[16] U.S. Bureau of Labor Statistics. Why the unemployment rate still matters. January 2017.