ESJ Social Sciences

Macroeconomic Forecasting Examining the COVID-19 Pandemic Using Selected Countries: A Machine Learning LSTM (Long Term Short Term Memory) Approach

Rolando A. Santos
Lakeland Community College, USA
Brian W. Sloboda
University of Maryland, Global Campus, USA

Doi:10.19044/esj.2022.v18n12p1

Submitted: 21 December 2021
Accepted: 18 April 2022
Published: 30 April 2022

Cite As:
Santos R.A. & Sloboda B.W. (2022). *Macroeconomic Forecasting Examining the COVID-19 Pandemic Using Selected Countries: A Machine Learning LSTM (Long Term Short Term Memory) Approach.* European Scientific Journal, ESJ, 18 (12), 1. https://doi.org/10.19044/esj.2022.v18n12p1

Abstract

The disease COVID-19 caused by the virus SARS-CoV-2 has initially disrupted the Chinese economy after the first cases were reported in December 2019 in Wuhan city in Hubei province of China. The virus continued to spread throughout the rest of the world. This spread of the virus led to the official designation of the COVID-19 pandemic by the World Health Organization (WHO) in late February 2020, which resulted in the disruption of these economies due to the stringent lockdowns and restrictions in travel disease's evolution. The disruptive economic impact is highly uncertain, making it difficult for policymakers to craft an appropriate policy response to these macroeconomic disruptions. To better understand possible economic outcomes, this paper explores the use of the machine learning approach LSTM to assess the economic forecast in some selected countries. The empirical results from this paper demonstrate that there are temporary disruptions in macroeconomics in the short run and these economies rebound. The recovery of each selected country may be different as the forecast would imply.

Keywords: Pandemics, infectious diseases, macroeconomics, machine learning, LSTM

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Section I: Introduction

When the news of the spread of the disease COVID19 caused by the virus SARS-CoV-2[1] hit the United States after being triggered in December 2019 in Wuhan city in Hubei province of China, there were series of closures of business establishments in the middle of March 2020. It started in San Francisco, California, then other states followed, such as Ohio, New York, etc. With these closures, the unemployment rate increased quickly, and so goes the rest of the economy. This paper aims to examine the effects of COVID19 on GDP forecasting and determine its implication to some selected countries using LSTM methodology. The adverse impact of such a pandemic can affect factors like GDP, unemployment, industrial production, and interest rates. This paper would look at the macroeconomic forecasting for selected countries and assess their recovery prospects.

In March 2020, the number of people filed for unemployment had gone up to 6.6 million workers (WSJ, 2020) as the coronavirus hit the United States' economy, marking an abrupt end to the nation's historic, decade-long run of job growth. The number of Americans filing for claims was nearly five times the previous record. Millions of US businesses have announced layoffs or furloughs as their cash flows dry up. Several state and local authorities have ordered nonessential businesses to close in response to the novel coronavirus pandemic, bringing the great American job machine to a sudden halt.

Retail sales, a measure of purchases at stores, gasoline stations, restaurants, bars, and online, fell at a seasonally adjusted 8.7 percent in March 2020 (WSJ, 2020) from a month earlier, the most significant month-to-month decline since the record began in 1992. Clothing store sales have declined by more than half as spending on vehicles, furniture, sporting goods, and electronics has fallen by double digits. The Federal Reserve has also said that US Industrial Production fell by 5 percent in March, the most significant drop since World War II. The initial impact in the housing market was a drastic drop of 30 percent, while the US stock indexes have dropped by approximately 2 percent (WSJ, 2020).

The paper first summarizes the existing literature on the economic impacts of past pandemics. Section 3 outlines the data sources and variables used in the analysis and the machine learning LSTM (Long Term Short Term Memory) Approach. Section 4 describes the results from the use of this model. Section 5 concludes the paper by summarizing the main findings and discussing policy implications.

[1] The World Health Organization (WHO) defines the disease as COVID-19 but the actual virus causing COVID-19 is called the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Section 2: Past Pandemics and The Economic Impacts

There are only a few studies of the economic costs of large-scale outbreaks of infectious diseases. Schoenbaum (1987) is an example of an early analysis of the economic impact of influenza. Meltzer et al. (1999) examined the likely economic effects of the US influenza pandemic and evaluated several vaccine-based interventions. At a gross attack rate (i.e., the number of people contracting the virus out of the total population) of 15-35 percent, the number of influenza deaths is 89 – 207 thousand, and an estimated mean real economic impact for the US economy is $73.1- $166.5 billion.

Studies of the macroeconomic effects of the SARS epidemic in 2003 found a significant impact on economies through large reductions in consumption of various goods and services, an increase in business operating costs, and re-evaluation of country risks reflected in increased risk premiums in thirty countries. Shocks to other economies were transmitted according to the degree of the countries' exposure, or susceptibility, to the disease. Despite a relatively small number of cases and deaths, the global costs were significant and not limited to the directly affected countries (Lee and McKibbin, 2003).

Bloom et al. (2005) used the Oxford economic forecasting model to estimate the potential economic impact of a pandemic resulting from the mutation of the avian influenza strain. They assume a mild pandemic with a 20 percent attack rate, a 0.5 percent case-fatality rate, and a consumption shock of 3 percent. Scenarios include two-quarters of demand contraction only in Asia (combined effect 2.6 percent Asian GDP or US$113.2 billion); a longer-term shock with a more extended outbreak and more considerable shock to consumption and export yield a loss of 6.5 percent of GDP (US$282.7 billion). Global GDP is reduced by 0.6 percent, global trade of goods and services contracts by $2.5 trillion (14 percent). Open economies were typically more vulnerable to international shocks.

Garret (2007) speculated about the possibilities of a future pandemic. The US Centers for Disease Control and Prevention's forecast of fatalities can recover 200,000 and would cost the economy over $160 million or roughly 1.5% of GDP. Because there is almost a complete absence of economic data from the Spanish Influenza (1918-1922), Garrett looked for evidence in newspaper articles printed during the pandemic, particularly at the local levels. Between that and the evidence in earlier economic studies, he found a geographic variation in the disease's effects that is unlikely in our far more interconnected nation a century later. Cities, unsurprisingly, had "higher mortality rates than rural areas of the states." Cities like Little Rock, Arkansas saw general merchant business declines of 40 percent, and even the retail grocery business reduced by one-third. A specific department store reported a more than 50 percent cut in daily income, but at least it was still operating.

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Though there was a flu-related "increase in demand for beds, mattresses, and springs," the city's businesses were "losing $10,000 a day on average ($133,500 in 2006 dollars). This is an actual loss, not a decrease in business that may be covered by an increase in sales when the quarantine order is over."

The Memphis Street Railway reported that 124 of its 400 employees were too sick to work on one day. A depopulated telephone company begged the public to make fewer unnecessary calls. Coal mine operators reported a 50 percent cut in production, with some mining camps forced to shut down from raging infections. Garrett explained the possibility of a post-pandemic increase in wage and income growth on "a greater increase in capital per worker, and thus output per worker"—which might not work out the same way from a starting point of 2020 rather than 1920.

Yet most of the 1918 pandemic's effects "were short-term," Garrett concluded. Most businesses suffered a significant revenue loss, especially those in the service sector. However, companies that specialized in healthcare-related products experienced an increase in revenue. It also caused a shortage of labor that resulted in higher wages due to people getting sick and dying.

Keogh-Brown, M. et al. (2008) presented a selection of model results to outline the potential impact of pandemic influenza. Their results suggested that a pandemic of the type experienced in 1957 or 1968/69 would harm GDP of approximately 0.5 percent and would produce losses to household consumption of up to 1 percent, a slight increase in government expenditure, and some minor impacts on exchange rates. Sectoral results from their model are tiny, so the overall economic impact of the pandemic itself would seem to be of relatively minor concern. However, the introduction of a school closure policy, even if restricted to the pandemic's peak only, caused a significant increase in the working population shock and dramatically increased the economic impact of the pandemic. Under a peak pandemic school closure policy, the GDP losses of between 5 percent and 8 percent. Also, household consumption could fall by almost 13 percent during the pandemic, and government expenditure could rise by up to 6 percent in some countries. These results highlighted the power of pandemic mitigation policies, however beneficial from the health perspective, magnified the economic impact. The effect of school closure that they have modeled may prove a worst-case scenario because parents would make alternative arrangements for the care of their children. Conversely, if school closures would last longer than the four weeks assumed in their study, it would reduce some parents' ability to locate child care and remain home longer. Consequently, this would harm the economy. While some mitigation policies would have a detrimental effect on the economy, their results showed that antivirals and vaccines proved very beneficial in dampening the negative

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

economic impacts resulting from the school closures. The economic impact of school closure, together with antivirals and vaccines, was approximately twice as significant as the impact of the disease itself but is much smaller than the economic impact of the scenario that considers school closure. While there is much uncertainty surrounding the nature of future pandemics, their study highlighted the need for further investigation into the potential economic impact of pandemic influenza. Further research into this subject would provide valuable insights for policymakers and form an essential blueprint in the preparedness plan for future pandemics.

Jorda, Singh, and Taylor (2020), pandemics have a long-lasting effect, especially on the real interest rate. The impact of interest on assets could last for decades (20 years on average). In some instances, it would take the natural rate of interest to go back to its original state after 40 years. This trend was consistent in most European countries. However, when it comes to real wages, they tend to increase after a pandemic. The upward trend in real wages was attributed to labor shortages resulting from the deaths.

In brief, the spread of infectious diseases often leads to a substantial decline in consumer demand, especially for travel and retail sales service. Also, if the virus is quite contagious, people may avoid social interactions, as witnessed during the COVID-19 pandemic. The economic impact or the adverse demand shock becomes substantive in countries with more extensive service-related activities and a high density of population, e.g., Hong Kong or Beijing, China. More importantly, the psychological shock ripples throughout the world, not just to the countries of local transmission of the virus because the world is closely connected via international travel.

Section 3: Data and Methodology

Data Sources

The data included the following variables: Real GDP, industrial production, unemployment rate, retail sales, and federal funds rate from January 1995 through February 2021 monthly. The data were obtained from Trading Economics (https://tradingeconomics.com/).[2] We shall use the LSTM (Long Term Short Term Memory) for forecasting purposes and evaluate its performance as a forecasting tool. We will be using countries like the United States, Germany, China, and Australia. We chose these countries to represent each continent , see how well the recovery from the pandemic was, and evaluate its economic impacts using forecasting methods.

[2] The authors used API to extract the data for this paper from Trading Economics.

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Methodology

The method of analysis that we will use in this paper is Machine Learning, emphasizing LSTM (Long Term Short Term Memory). The motivation for using the LSTM model because our data is time series. We would like to see the impact of past values as it is incorporated in the current values of the variables involved. LSTM networks are a type of RNN (Recurrent Neural Network). The LSTM modules are typically called cells rather than neurons and contain a series of gates. A diagram of an LSTM cell can be seen in Figure 1. Each LSTM cell (A) has a form of longer-term memory in the form of a cell state that is updated through time. A forget gate (i.e. h_{t-1}) at the new input and the hidden state decides which information in the cell state can be safely ignored. The input gate (x_t) then decides what information from the new input should be added to the cell state to be remembered. The sigmoid function (σ) decides which information is important to keep from the tanh output. Finally, the output gate (h_{t+1}) takes information from the cell state, input, and hidden state and generates the output for the current time step. In this way, LSTM networks can remember information through many timesteps, making them ideal for finding longer-term trends in data. At the same time, the LSTM cell still uses the hidden state and therefore has short-term memory as well. Overall, LSTM networks can be a powerful tool in time series forecasting (Olah, 2015).

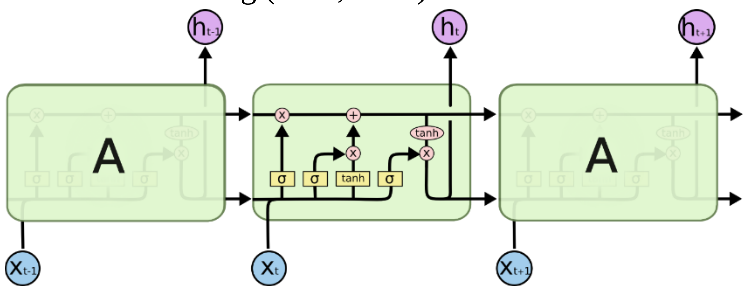


Figure 1: LSTM Structure

Before the data were used to train the LSTM network, they were split into a training set and a test set. The data were split by assigning the first 80 percent to the training set and the last 20 percent to the testing set. Next, the DateTime columns were removed from the training and test sets, leaving four-column data input in each set. Those columns were then normalized to between zero and one using feature scaling. Finally, the training and testing sets were split into input and target arrays. Each row of the input arrays contained a vector of length 249 and represented the input for a single training example. Each row of the target arrays contained a vector of length 4 and represented thefour4 target values for a single training example. Each of the four target values was the next 10 GDP production values directly following

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

the corresponding input vector of 249 values. The time window of inputs and outputs is then shifted by four values so that the first value in a given input vector is the same as the 11th value in the previous input vector. The outputs and targets do not overlap but rather are continuous in time.

The Keras Python library was used to build, train and test the LSTM network. The LSTM model was built upon Keras's sequential class. A single hidden layer of 50 neurons and an output layer of 6 neurons were added. The "Adam" optimizer was chosen for training. Once built, the LSTM was trained and tested using the training and test sets, respectively. The model underwent 25 epochs of training with a batch size of 10, a dropout rate of 0.2, and the "Adam" optimizer function. In addition to the test results, a 10-fold cross-validation process was used to evaluate the model. This entire building, training, and testing process was conducted twice with the time-interpolated data sets and once with the linearly interpolated data sets. 3. Results The single-layer LSTM network received 192 timesteps of the US GDP and its input vectors and forecasted 2 timesteps into the future. Each timestep was one month, so the model received just over 10 months of data as input and forecasted up to one month into the future. Both the single-layer models and the 5-layer model to which they were compared had 50 neurons in each hidden layer. The RMSE is the standard deviation of the residuals and measures how well a regression fits a set of data.

Section 4: Empirical Results of the Model

The actual versus forecasted plots of each country (Germany, China, Australia, and the US) showed similar results (see Figure 2,4,6 and 8). It mimics the actual values, especially in the case of Germany and Australia. China, on the other hand, the forecasted values reflect the trend but are more volatile and not as smooth as compared to other countries. In all the selected countries, the dip due to the pandemic was captured differently from each other. In some cases, the dip in the forecasted values is not as deep compared to other countries. A good example of this would be the United States which it showed the dip but not as deep as the actual impact of the pandemic (see Figure 8).

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Figure 1: Germany Validation Loss:

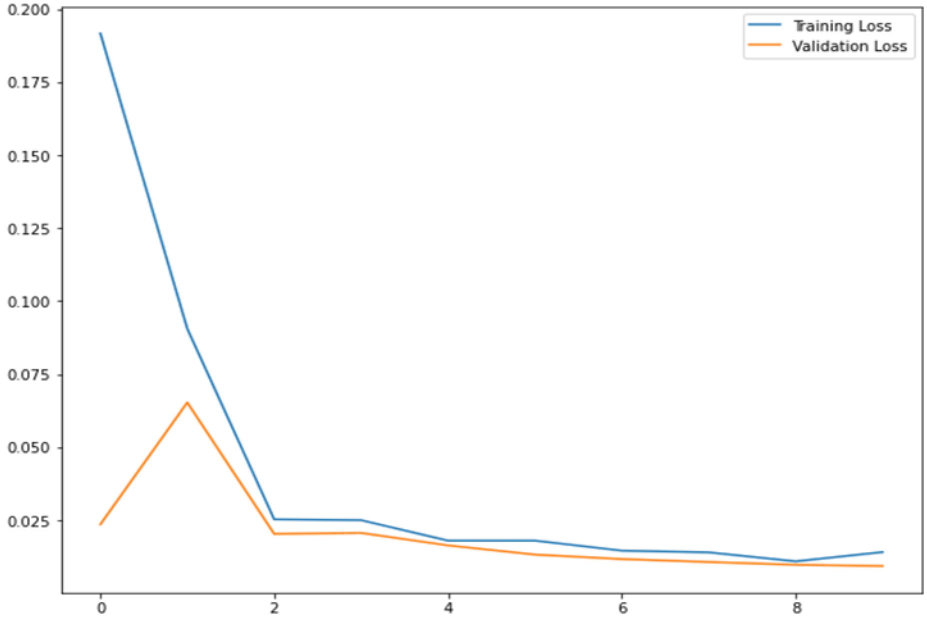


Figure 2: Germany Actual vs. Predicted:

MSE: 1.6361
RMSE: 1.2791

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Figure 3: China Validation Loss:

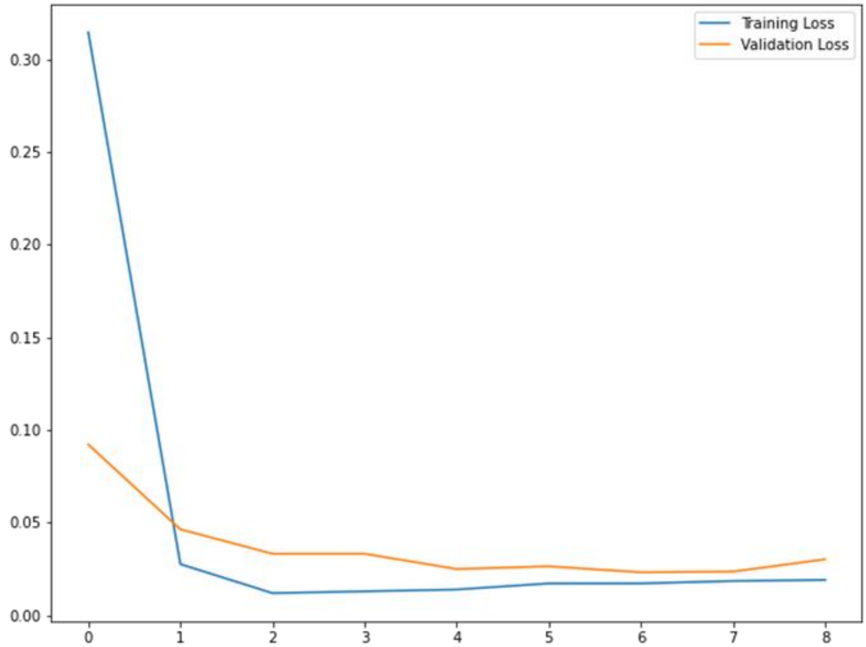
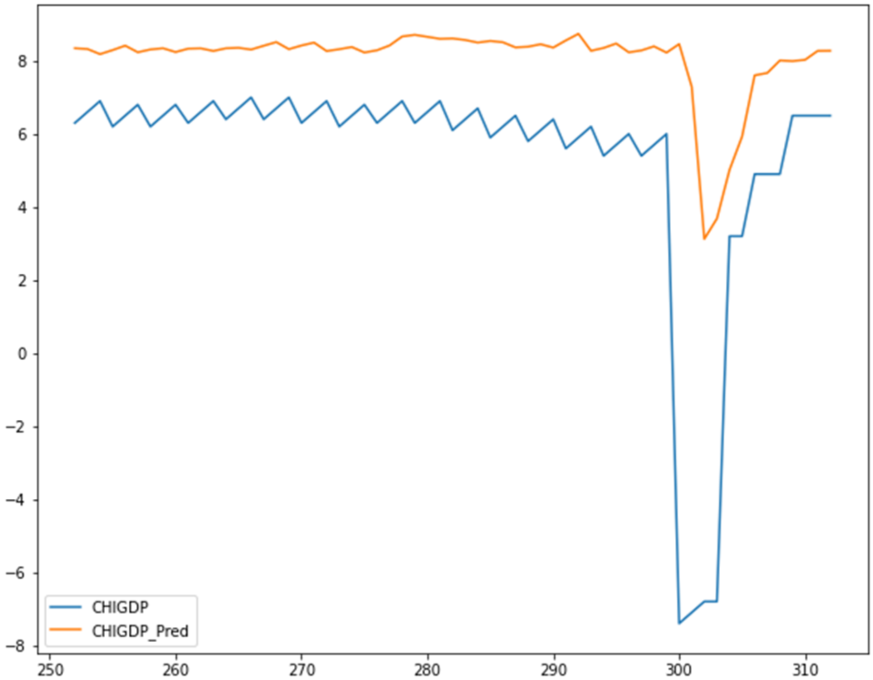


Figure 4: China Actual vs Predicted LSTM:



MSE: 15.0971
RMSE: 3.8855

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Figure 5: Australia: Training vs. Validation Loss

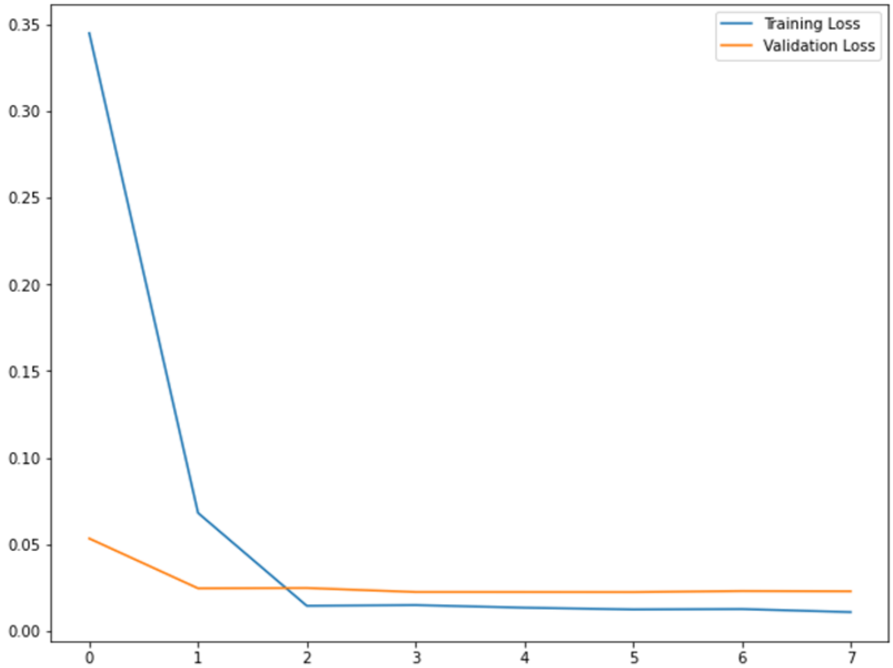
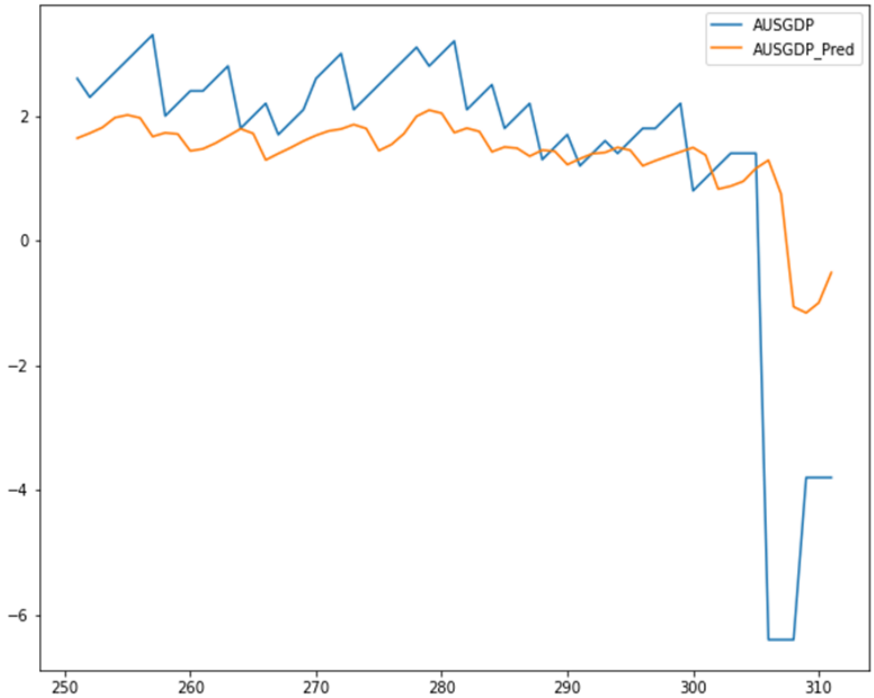


Figure 6: Australia: Actual vs. Predicted Value



MSE: 3.2043
RMSE: 1.7900

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Figure 7: US Training vs. Validation Loss

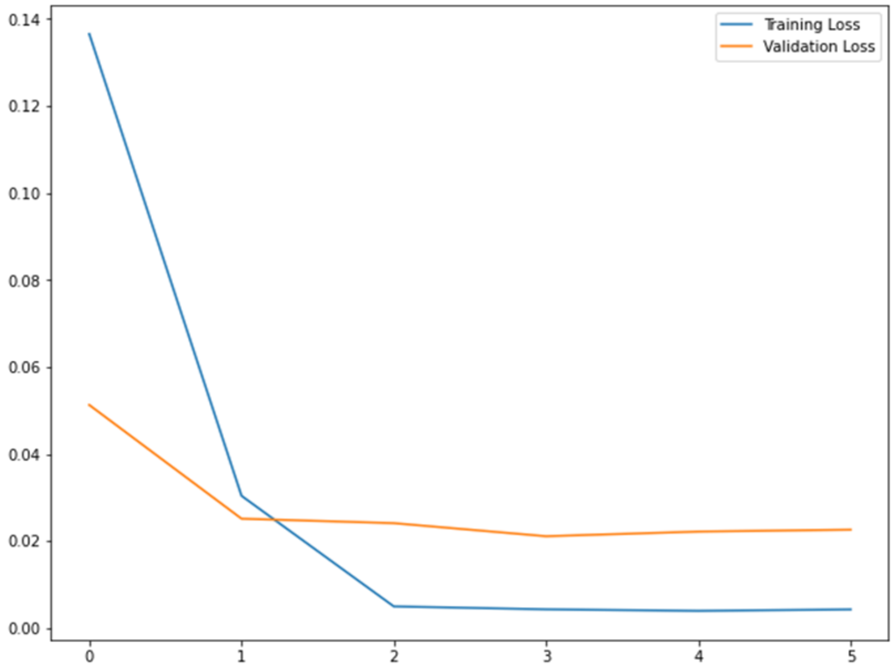
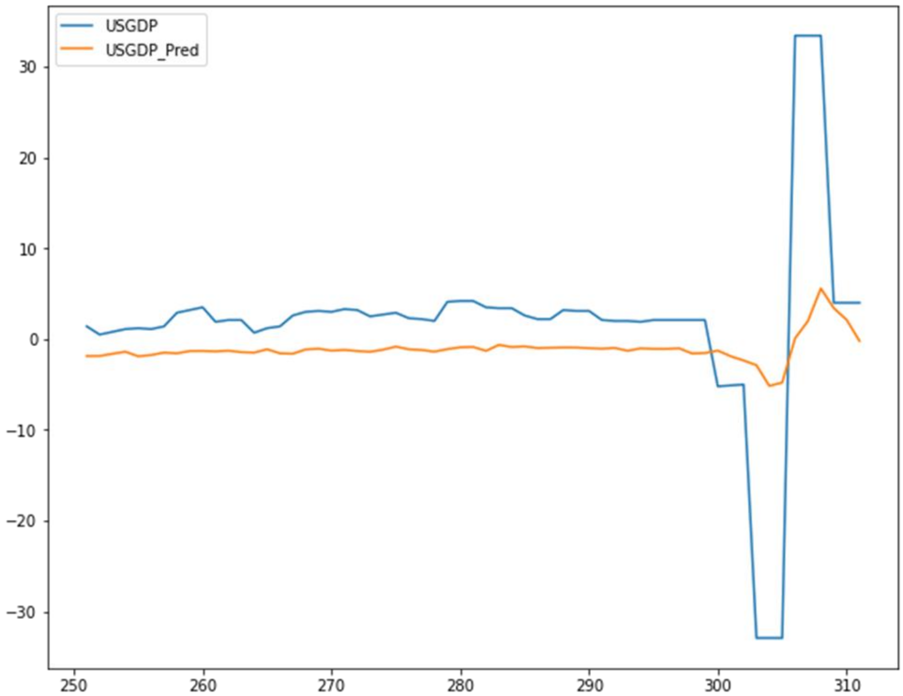


Figure 8: Actual vs Predicted US (LSTM)



MSE = 99.3383
RMSE = 9.968

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

Table 1 shows results from testing the two single-layer LSTM models with the test sets and compares their performance of the pre-existing 5-layer LSTM network. The RMSE is the standard deviation of the residuals and measures how well a regression fits a set of data. The assumption is that the lower the MSE (Mean Square Error) or RMSE (Root Mean Square Error), the more robust the model is. There is no specific ideal number for MSE, but the lower, the better. In addition to the test data, 10-fold validation was used to evaluate the model. The training and validation loss results for each country are shown in Figures 1, 3, 5, and 7. The results for each are that both the training and validation loss decrease over various epoch repetitions. This indicates that the model is fit, and the likelihood of forecast error is minimal. Although the actual forecast of the actual values shows a similar trend, the shock from the pandemic was not captured significantly. The RMSE calculation for the US is 9.9 percent which is significant, indicating that the prediction error is not that big.

Table 1: MSE and RMSE for Selected Countries:

	MSE	**RMSE**
Germany	1.6361	1.2791
China	15.0971	3.885
Australia	3.2043	1.79
USA	99.338	9.968

As to the forecasted values beyond the observed data, the US and China exhibited a remarkable forecast by 2022 and 2023. China's growth on average initially would be around 8-9 percent which is consistent with other forecasts, while the US has an average growth rate of 3-4 percent in 2021-2022. However, by 2023, growth rates are bound to increase on an average of 20-30 percent, which is expected to happen due to pent-up demand after a pandemic. Economists surveyed by *The Wall Street Journal* project US gross domestic product will grow by 6.4 percent this year (WSJ, 2020). Germany and Australia showed much less aggressive growth than their US and China counterparts. The growth is positive but not as robust as compared to the latter. In the case of Australia, we have seen a future decline in their growth rates but this is in a further future forecast. Caution must be taken for the long-term forecast as it may not necessarily be feasible even using this long-term, short-term memory method.

Section 5: Conclusion and Final Thoughts

The remarkable performance observed through deep learning-based approaches to the prediction problem is due to the "iterative" optimization algorithm used in these approaches to find the best results. By iterative, we mean to obtain the results several times and then select the most optimal one,

 ISSN: 1857-7881 (Print) e - ISSN 1857-7431

i.e., the iteration that minimizes the errors. As a result, the iterations help an under-fitted model be transformed into a model optimally fitted to the data. The actual versus forecasted values seem to show a robust fit, as evidenced by the RMSE. Each country has different forecasted values beyond the actual data that is given. US and China indicated a more robust recovery phase while Germany and Australia have a tone-dow down recovery growth. Although this paper was written while the pandemic was still ongoing, the results indicated some consistency with what other forecasters would have speculated.

References:

1. Brainerd, E. & Siegle, M. V. (2003). The Economic Effects of the 1918 Influenza Epidemic. Discussion Paper n. 3791. Centre for Economic Policy Research. London: CEPR.
2. Doherty, B. (2020). *Coronavirus Economics History Health*.
3. Bloom, E., V. d. Wit, et al., (2005). Potential economic impact of an Avian Flu pandemic on Asia. ERD Policy Brief Series No. 42. Asian Development Bank, Manila. http://www.adb.org/Documents/EDRC/Policy_Briefs/PB042.pdf.
4. Chou, J., Kuo, N. F., & Peng, S. L. (2004). Potential impacts of the SARS outbreak on Taiwan's economy. *Asian Economic Papers*, *3*(1), 84-99.
5. Cuddington, J. T., (1993a). Further results on the macroeconomic effects of AIDS: The dualistic, labour-surplus economy. *World Bank Economic Review 7*(3), 403-17.
6. Cuddington, J. T., (1993b). Modeling the macroeconomic effects of AIDS, with an application to Tanzania. *World Bank Economic Review 7*(2), 173-89.
7. Garrett, T. A. (2007). Economic effects of the 1918 influenza pandemic. *Federal Reserve Bank of St. Louis*, *26*.
8. Hyams, K. C., F. M. Murphy, et al., (2002). Responding to chemical, biological, or nuclear terrorism: The indirect and long-term health effects may present the greatest challenge. *Journal of Health Politics, Policy and Law 27*(2), 273-91.
9. Jordà, Ò., Singh, S. R., & Taylor, A. M. (2020). *Longer-run economic consequences of pandemics* (No. w26934). National Bureau of Economic Research.
10. Keogh-Brown, Marcus, McDonalds, Scott et al. *The Macroeconomic Costs of A Global Influenza Pandemic.* Available at: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=2755.

ISSN: 1857-7881 (Print) e - ISSN 1857-7431

11. Lee, J.W & McKibbin, W. (2004). Estimating the Global Economic Costs of SARS. in S. Knobler, A. Mahmoud, S. Lemon, A. Mack, L. Sivitz, and K. Oberholtzer (Editors), *Learning from SARS: Preparing for the Next Outbreak*, The National Academies Press, Washington DC (0-309-09154-3)
12. Litterman, R. B. (1979). Techniques of Forecasting Using VAR. Working paper no.115 Federal Reserve Bank of Minneapolis, Minneapolis, MN Nov 1979.
13. Meltzer, M. I., N. J. Cox, et al., (1999). The economic impact of pandemic influenza in the United States: priorities for intervention. *Emerging Infectious Diseases 5*(5), 659-71.
14. Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1394-1401). IEEE.
15. Olah, C. (2015). Understanding LSTM. Colahs Blog, https://colah.github.io/posts/2015-08-Understanding-LSTMs/.
16. Schoenbaum, S. C., (1987). Economic impact of influenza. The individual's perspective. *American Journal of Medicine 82*(6A), 26-30.
17. Sunstein, C. R., (1997). Bad Deaths. *Journal of Risk and Uncertainty 14*(3), 259-82.
18. *Wall Street Journal* (2020). Trump and the Stock Market.
19. Viscusi, W. K., J. K. Hakes, et al., (1997). Measures of mortality risks. *Journal of Risk and Uncertainty 14*(3), 213-33.
20. Yildirim, D., Toroslu, I., & Fiore, U. (2021). Forecasting Directional Movement In Forex data using LSTM with Technical and Macroeconomic Indicators. Financial Innovation.