INVESTIGATING THE IMPACT OF COVID-19 CRISIS ON THE ECONOMY OF USA AND INDIA: A COMPARATIVE STUDY USING MACHINE LEARNING

Mahima Gupta\(^1\), Dhaval Kumar Srivastava\(^2\), Sucheta Shukla\(^3\), Smitha Mony Sreedharan\(^4\), Ankur Saxena\(^5\)

\(^1\)mgupta161998@gmail.com, \(^2\)dhavalksrivastava@gmail.com, \(^3\)suchetashukla10@gmail.com, \(^4\)smsreedharan@amity.edu, \(^5\)asaxena1@amity.edu

Amity Institute of Biotechnology, Amity University, AUUP, Sector- 125, Noida, Uttar Pradesh, India, 201313.

ABSTRACT

Coronavirus disease is an infectious respiratory tract disease caused by SARS-CoV-2 virus. The spread of this pandemic had unprecedented effect on human life and world economy. With every increasing number of infected cases, latest technologies like Machine learning (ML) and Artificial intelligence (AI) are being employed for interpreting and solving the COVID-19 crisis. In the present work, we have compared the impact of COVID-19 crisis on the economy of USA and India using Artificial Intelligence and Machine Learning. We have applied Logistic Regression on the collected dataset to answer various questions related to the impact of COVID on economy of both the countries and have predicted future trends. The role of AI in setting benchmark for all future predictions and uses have also been outlined.

Keywords: COVID-19, Economy, India, USA, Machine Learning, Artificial Intelligence, Logistic Regression, Prediction

I. INTRODUCTION

The year 2020 will be remembered in the history as the year with COVID-19 pandemic gripping the entire world. The disease ripped off the human psyche and installed a sense of fear, shock and global economy decline everywhere. This virus was first reported in Wuhan, China on 31\(^{st}\) December 2019. Coronavirus (CoV) is a family of virus that causes illnesses and diseases some as deadly as MERS-CoV (Middle East Respiratory Syndrome and SARS-CoV (Severe Acute Respiratory Syndrome), and now the novel coronavirus COVID-19 too has been added to the list.

1.1 The COVID-19 Pandemic

Coronavirus (CoV) is a family of virus that causes illnesses and diseases some as deadly as MERS-CoV (Middle East Respiratory Syndrome and SARS-CoV (Severe Acute Respiratory Syndrome)[1][17]. This virus spreads by droplets in the air from the saliva or dose discharge of an infected person when they cough or sneeze. People with pre-existing medical problems like cardiovascular disease, diabetes, chronic respiratory disease, and cancer are more predisposed to develop this virus. The only way to prevent from being infected is to practice social distancing and remain quarantined when least of the symptoms start to develop. Prevention can be done by washing the hands regularly with a soap or by an ethanol-based hand sanitizer, and by avoiding touching the face with hands and wearing mask or covering the nose and mouth while going out in public. This virus infects different people differently. Some people need immediate attention and hospitalization, while some develop mild to moderate illness and recover without hospitalization. The most common symptoms include fever, dried cough, tiredness; and some fewer common symptoms are headache, sore throat, diarrhea, conjunctivitis, loss of taste or smell, aches and pains. Currently, there are no specific vaccines or treatments available for covid-19, but there are many ongoing clinical trials for a potential vaccine and treatment.

1.2 Comparison of U. S. A’s and Indian Economy
The US has the largest and the most powerful economy in the world, while India is the fastest growing economy. US ranks 1st on the GDP scale while India ranks 5th. US has the economic superpower because it constitutes almost a quarter of the global economy, has advanced infrastructure, technology and an abundance supply of natural resources. Whereas, India’s service sector contributes to more than 60% of its economy and making itself the fastest-growing sector in the world. If reports are to be believed, with its current speed and pace development, by 2030, India may overtake the USA, and become the second largest economy next to China.[2][18]

1.3 Effect of COVID-19 on the Economy

The impact of the pandemic has taken a toll on the economies of both the nations with effects being highly disruptive. The economies have gone into recession which means that there is economic shrinkage and the growth has stopped. In April 2020, the World Bank and the rating agencies downgraded India’s growth with the lowest figures it has seen in the past three decades since India’s economic liberalization in the 1990s. Similar recession conditions are seen in the US too, with main effect on it’s tourism, travel and hospitality industry. The Indian GDP estimated for Q4FY20 is at 3.1%; whereas for US at 2.5%. [2][19] Unemployment rates in both the countries have shot up. In this research paper we look forward to see the difference in the economic conditions of India and USA, which country had a higher scale of economic drop in the COVID-19 pandemic, how the global lockdown took over the financial conditions, and how the economies will recover after the pandemic. We might find answers to these questions using algorithms of Machine Learning.

Both Artificial Intelligence (AI) and Machine Learning (ML) have represented themselves as a powerful tool for fighting against Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus Strain type-2 (SARS-CoV-2), which has become an unprecedented Pandemic, since its Resurfacing in 2019. The collaboration of the White House with different tech companies and research institutes on March 16, 2020, published a call to artificial intelligence (AI) researchers World-wide for producing a novel text and data-mining techniques to contribute towards COVID-19-related research [3][20].

Artificial intelligence (AI) is a field and study of computer systems that performs operations that requires human intelligence for its completion or execution such as speech recognition, face detection, decision making etc. introduction of Artificial Intelligence (AI) in different fields have led to industrial revolution 4.0 that affect all the different aspects of human activities. Artificial intelligence includes various formal algorithms that can be used to perform various kinds of subtask in a study such as decision tree, logistic regression, nearest neighbour etc[3, 4].

Machine learning is a technique for data analysis that carries out analytical model designing. It is a further part of the branches of artificial intelligence that is based on the vision that system can learn from a dataset provided, identify different patterns in data and make decisions on its own with minimal human efforts or intervention. Machine Learning can be two types: Supervised Learning and Unsupervised Learning, with an only pertaining difference i.e. existence of labels in the datasets[5, 6].

Table 1: Summary of ML Classifiers Used for Evaluation of Risk Assessment and Prediction of GDP/ Economy/ Financial Crisis

<table>
<thead>
<tr>
<th>SNO</th>
<th>Objective of the paper</th>
<th>ML algorithms used</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Bankruptcy prediction</td>
<td>Support vector machines (SVMs), multiple discriminant analysis (MDA), logistic regression analysis (Logit), and three-layer fully connected back-propagation neural networks (BPNs)</td>
<td>[7]</td>
</tr>
<tr>
<td>2.</td>
<td>Integrated approaches based on GA and ML are used for slope stability prediction</td>
<td>Logistic regression, decision tree, random forest, gradient boosting machine, support vector machine, and multilayer perceptron neural network, were used for the relationship modelling and firefly algorithm (FA) was used for the hyper-parameters tuning.</td>
<td>[8]</td>
</tr>
</tbody>
</table>
In this research study, we have focused on the economic impact of the pandemic on India and United States of America. Economy in India is expected to lose ₹32,000 crore everyday in the first 21 days of lockdown and about 53% of businesses were significantly affected[1]. The supply chains were put under stress because only the manufacture of essential services was allowed. Small businesses, start-ups, hotels, airlines cut the salaries of their employees and the daily wage workers were the worst affected as they had nowhere to go. While in USA, there was a similar scenario but with more repercussions in spite of it being a developed nation. In USA, the conditions are as bad as the Great Depression of 1929 with the economy fallen really low. Unemployment remains to be the biggest issue due to stress in the global supply chain during the lockdown conditions. The economies of both the countries will suffer as long as the lockdown continues, and the after-work life of lockdown will also play a major role. The economy recovery after the lockdown ends will vary from industry to industry. Depending upon

---

<table>
<thead>
<tr>
<th></th>
<th>Chaotic genetic ant colony algorithm (FSCGACA) based classification model.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.</td>
<td>A prediction horse race using a one-step ahead model validation design to evaluate how each ML algorithm, data set, and variable selection method weighs on predictive accuracy.</td>
</tr>
<tr>
<td></td>
<td>Linear Methods- 4-Quarter Moving Average (4QMA), Forward Stepwise Regression, LASSO Regression, Ridge Regression.</td>
</tr>
<tr>
<td></td>
<td>Nonparametric Methods- Regression Trees, Random Forests (RF), Gradient Boosting, Multi-Adaptive Regression Splines (MARS), Support Vector Regression (SVR) with Radial Basis Function (RBF).</td>
</tr>
<tr>
<td>5.</td>
<td>Financial time series prediction.</td>
</tr>
<tr>
<td></td>
<td>ℓ2,1-norm and Random Fourier Mapping based Extreme Learning Machine (ℓ2,1 RF-ELM).</td>
</tr>
<tr>
<td>6.</td>
<td>To predict bankruptcy one year prior to the event, and compare their performance with results from discriminant analysis, logistic regression, and neural networks.</td>
</tr>
<tr>
<td></td>
<td>Support vector machines, bagging, boosting, and random forest.</td>
</tr>
<tr>
<td>7.</td>
<td>To predict the direction of US stock prices by integrating time-varying effective transfer entropy (ETE) and various machine learning algorithms.</td>
</tr>
<tr>
<td></td>
<td>Logistic regression, multilayer perceptron, random forest, XGBoost, and long short-term memory network.</td>
</tr>
<tr>
<td>8.</td>
<td>To study transmission mechanisms across stock markets along with effects from bond and currency markets and to examine crash event propagation in international stock markets.</td>
</tr>
<tr>
<td></td>
<td>Classification Trees, Support Vector Machines, Random Forests, Neural Networks, Extreme Gradient Boosting, and Deep Neural Networks.</td>
</tr>
<tr>
<td></td>
<td>Random Forests, Support Vector Machines, Neural Networks, Conditional inference trees.</td>
</tr>
<tr>
<td>10.</td>
<td>To predict the market performance of Karachi Stock Exchange (KSE) on day closing using different machine learning techniques.</td>
</tr>
<tr>
<td></td>
<td>Single Layer Perceptron (SLP), Multi-Layer Perceptron (MLP), Radial Basis Function (RBF) and Support Vector Machine (SVM).</td>
</tr>
</tbody>
</table>

---

[1] Turkish Journal of Physiotherapy and Rehabilitation; 32(2) ISSN 2651-4451 | e-ISSN 2651-4451 | www.turkjphysiotherrehabil.org
how the organizations reform their practices and how the consumer need is, the curve can be a U-shaped (slow recovery rate) or a v-shaped (quick recovery rate). [17][21]

Keeping in mind the main purpose of this study we have employed Machine Learning for analysis. Machine learning (ML) is a branch of Artificial Intelligence (AI) that relates the problem of learning from data samples to the general concept of inference. This works in two steps, first, the estimation of unknown dependencies from the dataset; second, the use of these unknown dependencies to predict new outcome for the system. Using Regression and Support Vector Machine or SVM techniques of ML we have predicted the trend of economic impact of COVID-19 on these two countries in question and have weighed them out in order to make the analysis easy for future studies.

**Regression Model for Prediction**

Logistic Regression is a Machine Learning algorithm which is used for the classification problems, it is a predictive analysis algorithm and based on the concept of probability. Logistic Regression uses a more complex ‘Sigmoid function’ or also known as the ‘logistic function’ instead of a linear function. Logistic Regression is used when the dependent variable(target) is categorical. Logistic regression makes ample use of the logistic function as it outputs values between 0 and 1 which we can use to model and predict responses. Logistic Regression is a significant machine learning algorithm because it provides probabilities and classify new data using continuous and discrete datasets.

The process of regression offers the forecast of the condition of a resulting variable at a specific point of time with the assistance of other corresponded autonomous factors. The regression technique, dissimilar to the grouping techniques, yields constant incentive inside a given range. But this is not all; maintaining a constant performance standard is of utmost importance, as with dynamic datasets, constant altering of the algorithms for the desired output is required and to do that a constant performance standard has to be maintained. This can be achieved by applying an evaluation metric to evaluate the result.

**II. MATERIALS AND METHODS**

3.1 Set-up of Environment

The dataset for this paper was collected from https://www.nasdaq.com/market-activity/stocks/usa/historical and https://in.finance.yahoo.com/quote/%5EBSESN/history/ for the months of January’2020 - July’ 2020 for India and USA. Since our objective was to find the difference in the economies of two of the world’s top five nations due to COVID-19 pandemic, we made a month and date-wise tabular form of the data to see the stock market trend of these two nations.

The data was collected in the .xlsx format which was later converted into the .csv format for further formatting, training, validation and then testing. The CSV data file was then uploaded on Jupyter notebook and then analysis was done with python 3.7 (64- bit Graphical Installer).

3.2 Pre-processing and Importation of Libraries

The dataset which was collected was an unstructured dataset having many null values or unstructured values which cannot undergo pre-processing. This unstructured data was now supposed to be structured by removing all sort of null values or data that is not appropriate for the study or that may cause issues and errors while execution of the codes/ Algorithms. The structuring of the data also helps in organizing the information according to the choice of the user and all sort of unwanted variables are removed from the same, which cannot be understood by the system.

But before Structuring of the data, various Python Libraries were imported into the Jupyter Notebook as these provide predetermined and pretested set of helper functions/objects/modules which your application code calls for specific functionality. A typical library has a very limited scope depending upon the purpose of its usage; therefore, their API’s also tend to be smaller and require fewer dependencies. The libraries imported were: NumPy, Pandas, Matplotlib, Scikit-learn, SciPy & Seaborn.

3.3 Applying Logistic Regression

Regression is the process of foreseeing continuous values by Learning from different independent features. Wherein, we first performed Logarithmic transformation of the variables followed by some initial setup of the model and foresaw the yield depending on the given info. The anticipated value was then contrasted with the objective and the evaluation of the model execution was done. At that point the different parameters of the model were balanced iteratively so as to arrive at the ideal estimation of the exhibited metric.

3.4 Evaluation of the Regression Model
After Logistic Regression was applied, the evaluation of the model was done by utilizing evaluation metrics such as the logistic probability score function, which the user to obtain a predicted probability score of a given event using a logistic regression model. The logistic probability score works by specifying the dependent variable (binary target) and independent variables as input. The function runs the logistic model and fits the model and calculates the fitted probability of the event from the model. The accuracy of our regression model (in the form of LR score) used is 99% for both the cases of India and USA as shown in the figures below:

![Figure 4: accuracy for India’s regression model](image1)

![Figure 5: accuracy for USA’s regression model](image2)

Another parameter which we needed in our dataset was Relative Strength index (RSI) indicator. It was calculated using the closing rates on the stock market for each day. RSI reflects the weakness and the strength of the economy, and the daily trade price movement, so it was required for the fulfilment of our aim.

### III. RESULTS

After running our code using Logistic Regression algorithm in machine learning, and considering all the parameters used, the results for the study were found.

The RSI indicator graph showed the economic drop difference during the entire pandemic situation. Figures 6 and 7 show the trend of RSI or the relative strength index of India and USA respectively.

![Figure 6: RSI of India for the months of February ’2020-July ’2020](image3)

![Figure 7: RSI of USA for the months of February ’2020-July ’2020](image4)

In the case of USA, the economic drop occurred earlier in the global lockdown, during the late days of the month of February itself, while in India this drop occurred in the mid of March. Both the nations then saw a varied range till the month of June. India again experienced a low in mid-May ’2020 due to the extended lockdown; whereas in USA since the lockdown began to unfold, the economic conditions seemed to be pretty stable. In June ’2020, as the lockdown unfolded globally, the economy began to rise. In USA, there was a gradual increase but in India there was a sudden shoot up seen.

Another parameter used was the Volume of the stocks to see the number of trades or stocks exchanged for the required time duration. Figures 8 and 9 represent the Volume of the stock trade for India and USA respectively.
The volume-time graph depicted the underlying cause for the economic drop or rise. Since the starting of the year, the trade volume has remained at a constant low for India with its first hike in the month of early April’2020. While in USA, the trends show that the only highest peak in the stock exchange till July’2020 was observed in the early month of March’2020. The exchange volume of USA never reached a hike again but juggled up and down constantly. On the other hand, after April’s shot up in the volume, India remained constantly low on the trades because of the extended lockdown conditions. Now as the unlock 1 set in, another peak hike was observed for India in June’2020.

Now if we correlate the RSI and volume, we find that since the volume remained generally constant and did not shoot up or down suddenly for USA, there was a gradual increase in the RSI. But in India, since there were sudden hikes in the volume in 4th and 6th months, the RSI also was seen on the extreme levels during the same period.

IV. DISCUSSION

This machine learning model used to contrast the difference between the world’s two most influential economies, can also lead to predict the future trends of economy in these countries. As we already saw that the global pandemic of Coronavirus disease or COVID-19 has caused a drop in the world economy leading to recessions in most parts. Now the point arises that how fast and steadily can the countries move past the recession stage. This can also be answered through our regression model. Since we saw that India observed random change and drop in economy, we can say that in the future the increase won’t be gradual but will be erratic. Whereas for USA the economy was balanced out with variations, we can say that it will have a more subtle increase. Both the economies moved past the recession phase in the very initial unlock days.

REFERENCES