CLIMATE CHANGE PREDICTION USING TIME SERIES FORECASTING

Ishita Aggarwal
CSE, University at Buffalo
Person Number: 50431309
iaggarwa@buffalo.edu

Nikhil Manali
CSE, University at Buffalo
Person Number: 50419517
nmanali@buffalo.edu

Sagar Sonawane
CSE, University at Buffalo
Person Number: 50431730
ssonawan@buffalo.edu

Mahesh Desai
CSE, University at Buffalo
Person Number: 50419649
mdesai3@buffalo.edu

Abstract

The climate change is one of the most eminent concern of humankind in the current times. The change in surface temperature, increase in carbon dioxide levels, all pose a threat to us. The best way to study the impact of climate change is by analyzing the trend in data. Deep Learning has some effective techniques like Time series forecasting to help with this problem. Time series forecasting is a type of supervised learning problem in Machine Learning. is a technique which predicts future events by performing an analysis on historical time stamped data. The assumption is that future trends will hold similar to past data. It has its applications in several fields such as Statistics, Earthquake Prediction, Weather Forecasting and so on. In the paper An Application of Time Series Analysis for Weather Forecasting proposed by Agrawal et. al in 2012, have used a Multi-layer perceptron Neural Network on time stamped Indian Meteorological Department data to forecast the weather. In this project we investigate different methods such as Long Short Term Memory (LSTM), Prophet, Gated Recurrent Unit (GRU) for applying time series forecasting principles to Surface Temperature Data, USA CO2 Data and Natural Disaster Data to predict future data to analyze the alarming rate at which Climate Change is impacting planet Earth.

Keywords: Time Series Forecasting, Climate Change

1. Introduction

1.1. Problem Statement:

Climate change is the subject of how weather patterns change over decades or longer. It is impacted by both nature and humans. Climate concerns have been a part of history since the 1800s. In the 1800s scientists studied the rise in CO2 levels leading to the greenhouse effect. Between the years 1940-1970, due to the post war aerosol pollutants in the air, the temperature dropped thereby leading to theories of another ice age. From the 1980s we have been experiencing global warming. The Australian wildfire, the East African drought, the South Asia floods are few examples of natural disasters that beg for climate change action.

How are we causing climate change? From leaving the refrigerator door open to running small loads of laundry, we are very prone to wasting energy. Wastage leads to more consumption, thereby leading to more production which causes a stark increase in the CO2 levels. A Boeing 747-400 travelling 3500 miles at 80% passenger capacity produces 101 gm per passenger mile. Figure 1 shows the carbon footprint of the foods that what we eat.

Climate is ever changing, then why is Climate Change a sudden concern? Wolff et al. in their paper Climate Change Evidence and Causes, mention that All climate changes have been disruptive. Some climate changes have led to extinction of species, while others have impacted the land surface and ocean population.

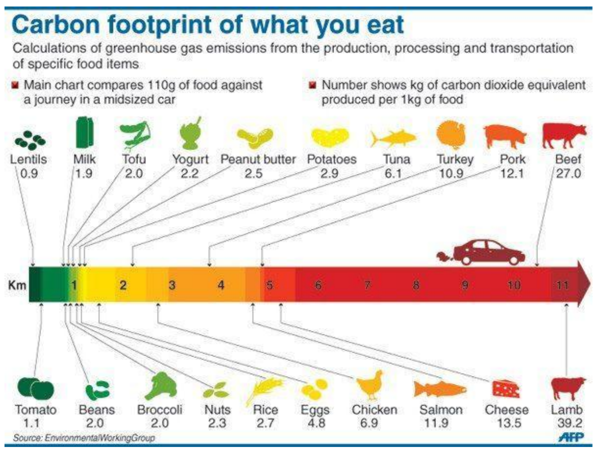


Figure 1: Carbon Footprint of what we eat

The current climate change is faster than the past events, making it difficult for humans to adapt.

One of the biggest issues that we face with Climate Change action is the inability to know the future. The solution to this problem is Deep Learning principle of Time Series Forecasting.

1.2. <u>Aim:</u>

Our goal is to create a climate change application that can be extended to any climate database with minimal changes to predict future data and extract inferences from the same.

In this project, we validated three climate databases using three different Deep Learning models which are well known for time series forecasting.

2. Data

2.1. <u>Earth Surface Temperature Data:</u>

The dataset was put together by Berkley Earth, which is affiliated with Lawrence Berkley National Laboratory. The Berkley Earth Temperature study takes into consideration 1.6 billion temperature reports from 16 pre-existing archives.

The dataset mentioned above has datapoints recorded at a regular time interval of 1 month, thereby making it efficient for time series forecasting. The data is from 1750 to 2015.

2.2. <u>US Carbon Dioxide Emissions Data:</u>

A study done by overshootday.org on the planet's biocapacity suggests that the ecological footprint of United States is 8.1 gha per person and earth's biocapacity is 1.6 gha per person. Therefore, we would need 5.1 earths to survive if everyone lived like USA residents.

Keeping the above point in mind, we considered the US Carbon Dioxide Emissions data as our second database to analyze the carbon footprint. The dataset consists of yearly data from 1800 to 2020. It is a challenge for machine learning but with advanced technologies like Prophet from Meta, we were able to predict the future data.

2.3. <u>Global Natural Disasters Data:</u>

Natural Disasters are a key to understanding the impact of climate change. Due to the increase in temperatures, we are witnessing heat waves in South Asia like never before. Across the world, droughts are becoming longer, and more extreme, tropical storms are becoming more severe, glaciers are melting faster, and we are witnessing unnatural events.

Texas is one of the warmest states in the United States. The winter storms in Texas with temperatures dropping to -13℃ and people becoming vulnerable to extreme cold conditions due to power outage is one of the extreme impacts of climate change. The dataset consists of yearly data from 1900 to 2018 with the number of reported natural disasters.

3. Models

Time series forecasting is used to predict future values or classification at a particular point of time in future. This prediction problem requires the data to have a time component, thereby making it a data driven prediction model. It has a lot of applications namely forecasting the spread of COVID, prices of bitcoin and so on.

As per an article by Gavita Regunath, Time series forecasting has 4 major components:
Trend: Increase/decrease in series of data over long time.
Seasonality: Fluctuations in pattern over day/month/week/year.

Cyclical Variations: Data fluctuates at irregular intervals.
Random Variations: Random data fluctuations.

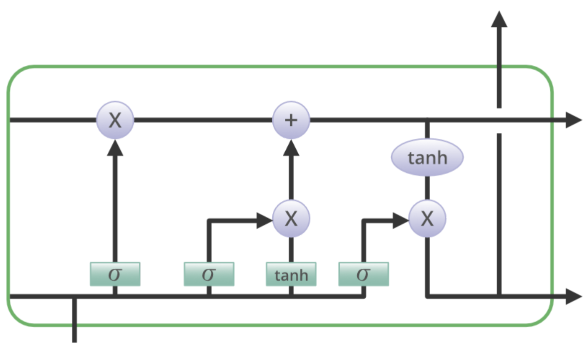
While there are many time series algorithms, we decided to select three to compare the results and choose the model best suited for our data.

We also added another code to view the change in surface temperature per country over the years. This animation is aimed at increasing awareness through visual mediums based on data.

3.1. Long Short-Term Memory (LSTM):

What is LSTM? It is special kind of recurrent neural network that is capable of learning long term dependencies in data. This is achieved because the recurring module of the model has a combination of four layers interacting with each other.

How does it work? LSTM has a chain structure that contains four neural networks and different memory blocks called cells.



Information is retained by the cells and the memory manipulations are done by the gates.

There are three gates :

1. **Forget Gate:** The information that is no longer useful in the cell state is removed with the forget gate. Two inputs x_t (input at the particular time) and h_t-1 (previous cell output) are fed to the gate and multiplied with weight matrices followed by the addition of bias. The resultant is passed through an activation function which gives a binary output. If for a particular cell state the output is 0, the piece of information is forgotten and for output 1, the information is retained for future use.

2. **Input gate**: The addition of useful information to the cell state is done by the input gate. First, the information is regulated using the sigmoid function and filter the values to be remembered similar to the forget gate using inputs h_t-1 and x_t. Then, a vector is created using tanh function that gives an output from -1 to +1, which contains all the possible values from h_t-1 and x_t. At last, the values of the vector and the regulated values are multiplied to obtain the useful information

3. **Output gate:** The task of extracting useful information from the current cell state to be presented as output is done by the output gate. First, a vector is generated by applying tanh function on the cell. Then, the information is regulated using the sigmoid function and filter by the values to be remembered using inputs h_t-1 and x_t. At last, the values of the vector and the regulated values are multiplied to be sent as an output and input to the next cell.

What are the advantages of using LSTM?

It tackled the problem of long-term dependencies of RNN in which the RNN cannot predict the word stored in the long-term memory but can give more accurate predictions from the recent information.

As the gap length increases RNN does not give an efficient performance. LSTM can by default retain the information for a long period of time.

It is used for processing, predicting, and classifying on the basis of time-series data.

3.2. Prophet (Meta Open Source):

What is Facebook Prophet? It is an open-source algorithm for generating time-series model using timed data. It provides a package with intuitive parameters that are easy to tune. The goal was to ensure that someone who lacks expertise in forecasting models would also be able to make meaningful predictions.

How does it work? Letham et al in their paper Prophet: forecasting at scale explain how prophet works. It is an additive regression model with 4 components:
Equation: $y(t) = g(t) + s(t) + h(t) + \varepsilon_t$
1. Prophet detects changepoints from data and pieces together a liner or logistic growth curve trend (y(t))
2. Yearly seasonal component using Fourier series. s(t)

3. Weekly seasonal component using dummy variables.
4. User provided list of important holidays. h(t)
ε_t is the error term for any unusual changes.

The prophet model is fit using Stan (a state-of-the-art model for statistical modelling and high-performance statistical computation). Stan performs the max posteriori for parameters in less than 1 second, thus enabling prophet to estimate parameters using Hamilton Monte Carlo algorithm.

Using time as a regressor, prophet is trying to fit several linear and nonlinear functions of time as components. The forecasting problem is framed as a curve-fitting exercise instead of looking explicitly at time-based dependence of each observation within a time series.

What are the advantages of using Prophet?
1. The flexibility to adjust parameters leads to improvement in forecasts.
2. It can handle missing observations and outlier problems very easily.
3. Because of using Stan, it is faster and more accurate.

3.3. Gated Recurrent Unit (GRU):

What is GRU? A gated recurrent unit (GRU) is a type of recurrent neural network that uses connections between nodes to accomplish machine learning tasks such as time-series forecasting. GRU aid in the adjustment of neural network input weights in order to avoid the vanishing gradient problem that affects recurrent neural networks.

How does it work? GRU uses, the so-called, update gate and reset gate. Basically, these are two vectors which decide what information should be passed to the output. It is a 2 component 2 step process:

1. Update Gate: The update gate helps the model to determine how much of the past information (from previous time steps) needs to be passed along to the future. This is calculated by the formula -

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1})$$

2. Reset gate: Reset gate is used from the model to decide how much of the past information to reset(forget). This is calculated by the same formula as the update gate but the difference comes in the weights and the gate's usage.

Step 1. Current memory content: This step introduces a new memory content which uses the reset gate to store the relevant information from the past. It is calculated by the formula -

$$h_t^{'} = \tanh(Wx_t + r_t \odot Uh_{t-1})$$

i.e. input x_t is multiplied by a weight W and h_(t-1) is multiplied by weight U which is then used to get a Hadamard product with reset gate r_t. This determines what is to be removed from the past time steps.

Step 2. Final memory at current step: At the end, the model needs to calculate h_t (a vector which has the relevant information for the current unit). This is done with the use of an update gate and by the formula:

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h_t^{'}$$

What are the advantages of using GRU?

1. GRU can be implemented on smaller training data with better accuracy than LSTM.

2. GRU is more computationally efficient.

3. GRU is simpler and more easy to modify

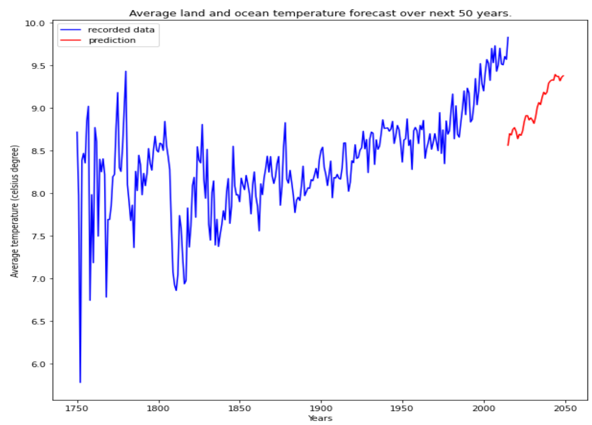
4. Results

We viewed each of the results for all three databases to understand the impact data points have on each of the chosen models.

4.1. Long Short-Term Memory (LSTM):

4.1.1. *Earth Surface Temperature Data:*

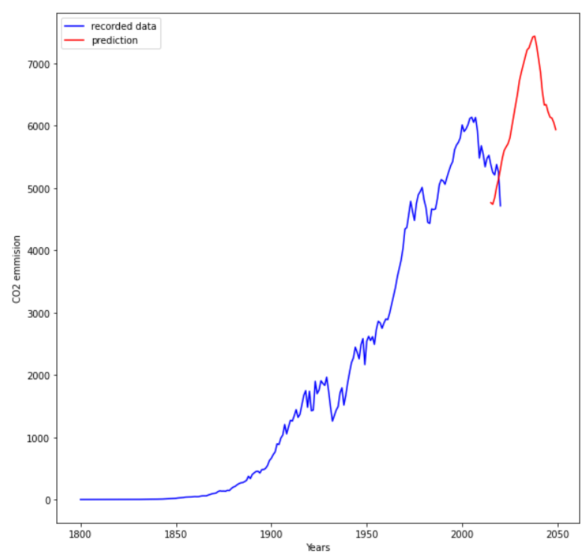
As you can see the prediction of the average surface temperature will more or less will be equal to current avg. temp as predicted by LSTM. It results are possible due to current efforts on containing global warmings



The accuracy for this dataset by LSTM model is: 90.92%.

4.1.2. ***US Carbon Dioxide Emissions Data:***

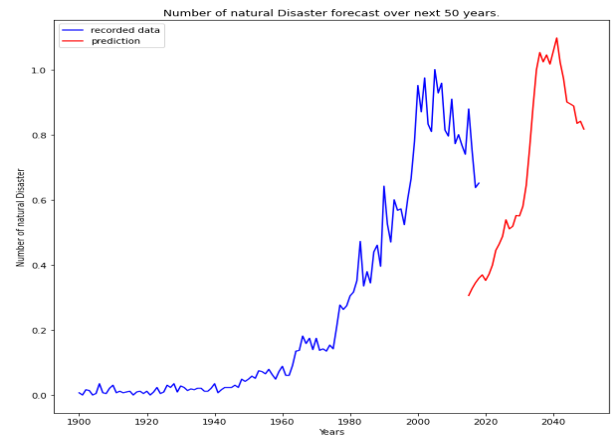
The CO2 levels in USA has increased from 4600 million metric tonne in 2018 to a predicted 5800 million metric tonne in the year 2050 which is followed by a peak of 7400 million metric tonne in 2039 years. The below graph shows the CO2 emission from US each year from 1800 to 2050.



The accuracy for this dataset by LSTM model is: 95.91%.

4.1.3. ***Global Natural Disasters Data:***

As per the trends and prediction done by LSTM the number of natural disaster will continue to follow the current trans but start falling sharply after some years.



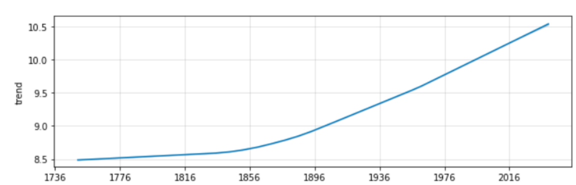
The accuracy for this dataset by LSTM model is: 90.74%

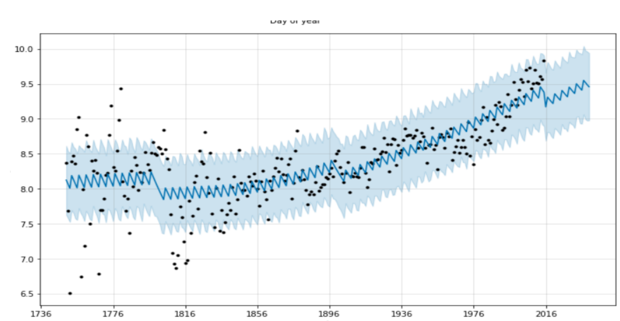
4.2. <u>Prophet (Meta Open Source):</u>

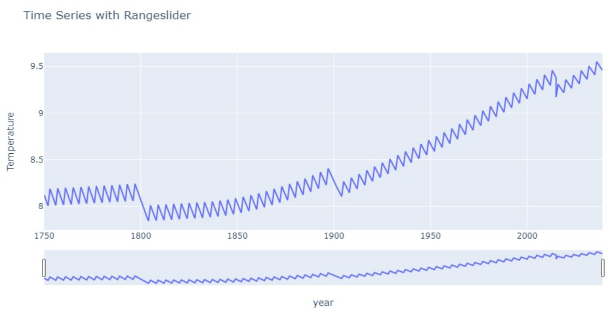
4.2.1. ***Earth Surface Temperature Data:***

The rate at which the earth surface temperature has been increasing is very alarming. From the predicted values obtained using Prophet database, we can see that the earth's surface temperature increased by 1.33°C from 1750 to 2039.

The graph for trend vs date is as below:







The x axis depicts all years and y axis depicts the surface temperature.

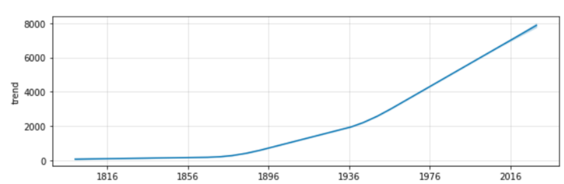
Both graph show us an increase in surface temperature data over the years.

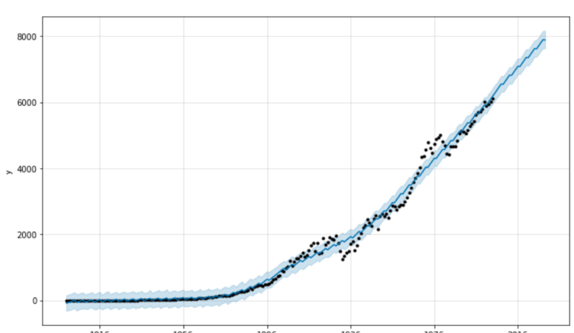
The Accuracy of Prophet for this model is: 90.15%, which is proof that we can rely on the predicted data.

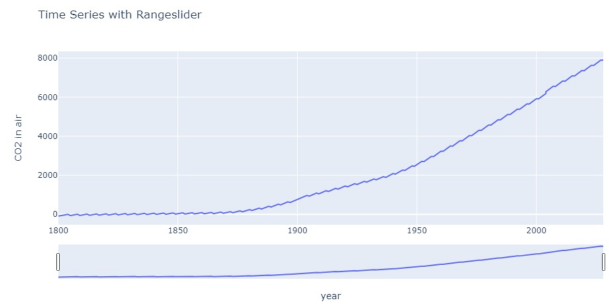
4.2.2. *US Carbon Dioxide Emissions Data:*

The CO2 levels in USA has increased from 0.253 million metric tonne to a predicted 7868.761 million metric tonne in the year 2028 which is about a 3.1million% increase over a span of just 228 years!

The graph for trend vs date is as below:







The x axis depicts all years and y axis depicts the CO2 emissions in USA.

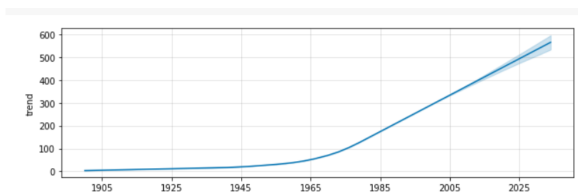
Both graph show us an increase in CO2 emissions over the years.

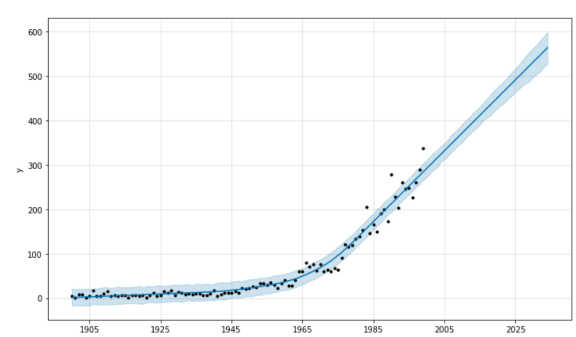
The Accuracy of Prophet for this model is: 75.47%, which is proof that we can rely on the predicted data.

4.2.3. *Global Natural Disasters Data:*

The number of natural disasters globally has increased from 5 a year to a predicted 558.5 in the year 2033 which is about 11070% increase over a span of just 133 years.

The graph for trend vs date is as below:





The x axis depicts all years and y axis depicts the global number of diasters.

Both graph show us an increase in number of natural disasters over the years.

The Accuracy of Prophet for this model is: 80.65%, which is proof that we can rely on the predicted data.

4.3. <u>Gated Recurrent Unit (GRU):</u>

4.3.1. *Earth Surface Temperature Data:*

From the predicted values obtained using GRU, we can see that the earth's surface temperature increased by 1.082°C from 1750 to 2050.

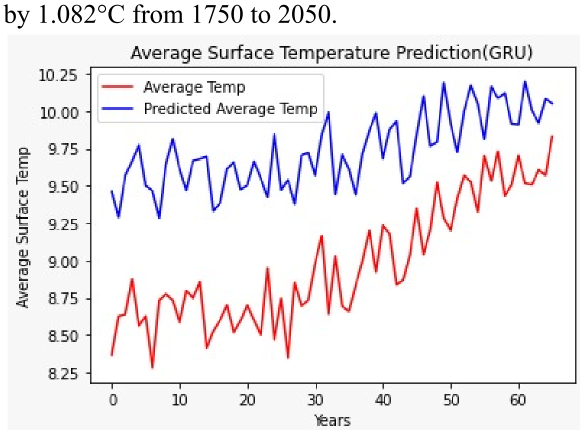


Fig. GRU Test Result

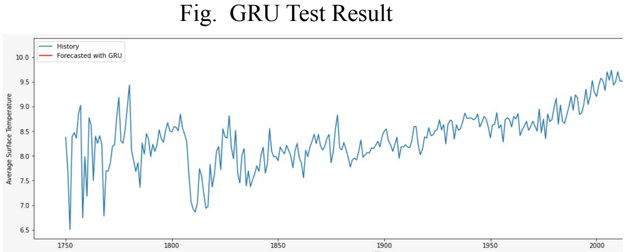


Fig. GRU Surface Temperature Prediction

The Mean Absolute Percentage (MAP) Error for this dataset was calculated to be 2.856%
Accuracy of GRU for this dataset = 97.14%

The x-axis defines years and the y-axis defines the surface temperature.

4.3.2. *US Carbon Dioxide Emissions Data:*

The GRU has predicted a rise in CO_2 levels in USA from 0.253 million metric tonnes in 1800 to 5189.46 million metric tonnes in the year 2050. which is about a 2 million% increase over a span of just 250 years.

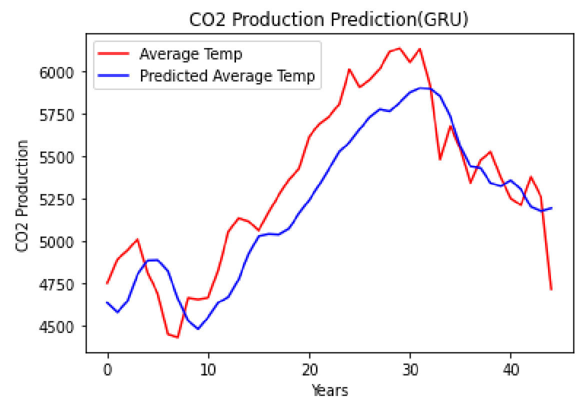


Fig. GRU Test Result

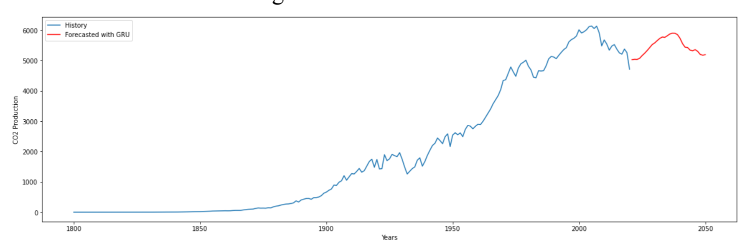


Fig. GRU CO2 Production Prediction Result

The Mean Absolute Percentage (MAP) Error for this dataset was calculated to be 4.043%
Accuracy of GRU for this dataset = 95.956%

The x-axis defines years and the y-axis defines CO2 production in USA.

4.3.3. *Global Natural Disasters Data:*

The number of natural disasters globally has increased from 5 a year to a predicted 301 in the year 2033 which is about 5920% increase over a span of just 133 years.

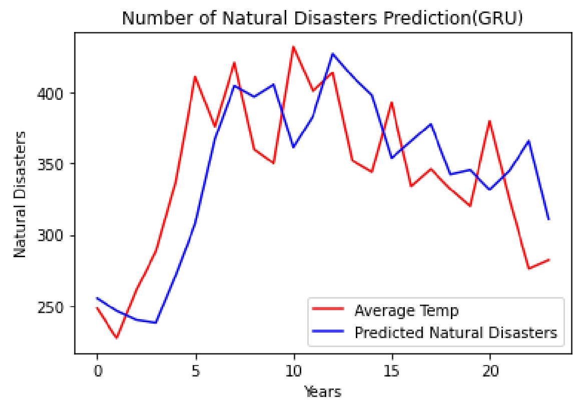


Fig GRU Test Result

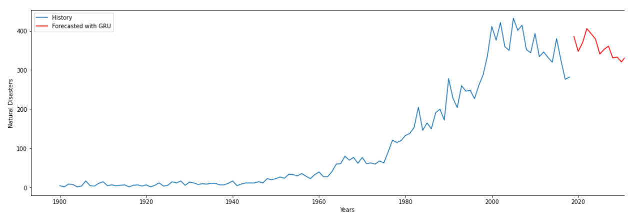


Fig GRU Natural Disaster Prediction Result

The Mean Absolute Percentage (MAP) Error for this dataset was calculated to be 11.324%

Accuracy of GRU for this dataset = 88.67%

The x-axis defines years and the y-axis defines the number of natural disasters.

4.4. Key Takeaways:

Below is the takeaways for each of the databases-

Hansen et al in their paper Global temperature change mention that an increase in surface temperature of more than 1°C is regarded as "dangerous" climate change with an increased likelihood of effects of sea level and extinction of species. Our data suggests we have already crossed that limit and current trends indicate that the temperatures will continue to rise.

Sperry et al in their paper 'The impact of rising CO2 and acclimation on the response of US forests to global warming' mention that while an increase in CO2 may be beneficial for trees, the rise in temperature levels would need the trees to rapidly adjust to the global climate change and even with the rapid increase in the same the probability of a favorable outcome is 71% but without any certainty. While the population increased 6107% - the CO2 emission levels increased more than 2 million% in the US alone. An analysis on world data would be even more terrifying!

Botzen et al in their paper 'The Economic Impacts of Natural Disasters: A Review of Models and Empirical Studies' conclude that an increase in a number of natural disasters would have dire economic implications on low-income countries. While it may not impact globally on a large scale but the impact on smaller countries would be impactful.

5. Conclusion

From the above experiments we observed that for each dataset, different models outperform each other. GRU, LSTM and Prophet gave an accuracy of 97.94%, 90.92% and 90.15% for the temperature dataset. 95.96%,95.91% and 75.47% for CO2 dataset. 88.67%, 90.74% and 80.65% for Disaster dataset. From this, we conclude that in most of the cases GRU outperformed other models as it performs better on low to medium size datasets.

CO2 molecules in the atmosphere absorb heat (infrared radiation) coming from the surface and re-radiate some of the heat back to the earth's surface in essence trapping the heat and creating a warming effect. The rise in CO2 emissions has been at an alarming rate which has led to a high increase in the surface temperature. Possible outcomes of increased land temperature are the increase in the severity of natural disasters such as droughts and storms. Increased land temperature can result in the melting of ice caps thus increasing sea levels which could result in beach erosion and higher chances of flooding.

The oxford dictionary defines crisis as "a time of great danger, difficulty or doubt when problems must be solved or important decisions must be made". A lot of climate change movements are going on to increase awareness about the fact that climate change is not another climate event, climate change needs to be treated for what it is - a crisis! Based on the predicted data, we can see that the climate crisis is in fact a reality and concrete environmental decisions need to be put in place.

As per the worldwildlife.org, humans are heading towards the 6th mass extinction. While the first 5 extinctions were caused by natural disasters, the sixth mass extinction is driven by human activity, primarily (though not limited to) the unsustainable use of land, water and energy use, and climate change.

6. Future Works

While there have been substantial leaps and bounds in the field of Machine Learning and Deep Learning, the biggest challenge is that the fields rely predominantly on data. The more the data, the better the prediction but that would not necessarily be helpful in cases with limited data. Machine Learning technologies that can handle the problem of limited data will be helpful for future works. Predicting doomsday will be a future work.

7. References

York University Ecological Footprint Initiative. National Footprint and Biocapacity Accounts, 2022 edition. Produced for the Footprint Data Foundation and distributed by Global Footprint Network. Available online at: https://data.footprintnetwork.org.

H Nguyen, Kim Phuc Tran, S Thomassey, M Hamad. Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in Supply Chain Management. International Journal of Information Management, Elsevier, 2020. ffhal-03083642.

Qiang, L. The application of neural network to the analysis of earthquake precursor chaotic time series. *Acta Seimol. Sin.* **13,** 434–439 (2000). https://doi.org/10.1007/s11589-000-0025-8.

Abhishek Agrawal, Vikas Kumar, Ashish Pandey, Imran Khan. An Application of Time Series Analysis for Weather Forecasting. International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 2, Mar-Apr 2012, pp.974-980.

Paul Dix. Why Time Series Matters for Metrics, Real-Time Analytics and Sensor Data. Influxdata, Revision 5, July 2021.

Eric Wolff FRS, Inez Fung and others. Climate Change Evidence and Causes. The Royal Society, Update 2020.

https://www.infoworld.com/article/3622246/an-introduction-to-time-series-forecasting.html

https://www.advancinganalytics.co.uk/blog/2021/06/22/10-incredibly-useful-time-series-forecasting-algorithms

https://towardsdatascience.com/time-series-analysis-with-facebook-prophet-how-it-works-and-how-to-use-it-f15ecf2c0e3a

Ben Letham, Sean J Taylor. Prophet: forecasting at scale, Meta Research, Feb 23, 2017.

https://mc-stan.org/

https://www.analyticsvidhya.com/blog/2018/05/generate-accurate-forecasts-facebook-prophet-python-r/

https://www.kaggle.com/code/subbhashit/stock-time-series-using-gru

https://towardsdatascience.com/predictive-analytics-time-series-forecasting-with-gru-and-bilstm-in-tensorflow-87588c852915

https://www.kaggle.com/code/carabaestlein/forecasting-temperatures-using-fb-prophet

https://www.techopedia.com/definition/33283/gated-recurrent-unit-gru

https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be

James Hansen, Makiko Sato, Reto Ruedy, Ken Lo, David W. Lea, Martin Medina-Elizade. Global temperature change. Sept 26, 2006. https://doi.org/10.1073/pnas.0606291103

John S. Sperry, Martin D. Venturas, Henry N. Todd, Xiaonan Tai. The impact of rising CO2 and acclimation on the response of US forests to global warming. Nov 25, 2019. https://doi.org/10.1073/pnas.1913072116

W. J. Wouter Botzen, Olivier Deschenes, and Mark Sanders. The University of Chicago Press Journals. Volume 13, Number 2.

https://www.kaggle.com/datasets/danielrpdias/co2-and-greenhouse-gas-emissions

https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data?select=GlobalTemperatures.csv

https://ourworldindata.org/grapher/number-of-natural-disaster-events https://www.worldwildlife.org/stories/what-is-the-sixth-mass-extinction-and-what-can-we-do-about-it#:~:text=Unlike%20previous%20extinction%20events%20caused,energy%20use%2C%20and%20climate%20change.