Predicting Climate Induced Total Disasters.

By: Sean Klein, Kevin Russell, and Ataish Nehra

Advisor: Dr. Nafa Fatema

DS5500 Capstone Project Presentation

Table of content:

1. Student Objectives ....(Slide 3-4)
2. Problem Specifications ....(Slide 5)
3. Related Work ....(Slide 6)
4. Amendments ....(Slide 7)
5. Solution Design ....(Slide 8-19)
   - 5.1. Original Dataset ....(Slide 9)
   - 5.2. Data Pre-processing ....(Slide 10)
   - 5.3. Models Chosen ....(Slide 11)
   - 5.4. LSTM Model - Workflow ....(Slide 12)
   - 5.5. LSTM Architecture ....(Slide 13)
   - 5.6. RNN Model ....(Slide 14)
   - 5.7. Attention Model ....(Slide 15)
   - 5.8. Forward-Forward Model ....(Slide 16-17)
   - 5.9. Tool List ....(Slide 18)
   - 5.10. Time Schedule ....(Slide 19)
6. Demo ....(Slide 20)
7. Postmortem ....(Slide 21)
8. Results and Deliverables ....(Slide 22-24)
   - 8.1. Actual vs Prediction graphs ....(Slide 23)
   - 8.2. Comparison ....(Slide 24)
9. Contributions ....(Slide 25)
10. Future Enhancements ....(Slide 26)
11. Questions ....(Slide 27)

Student Objectives.

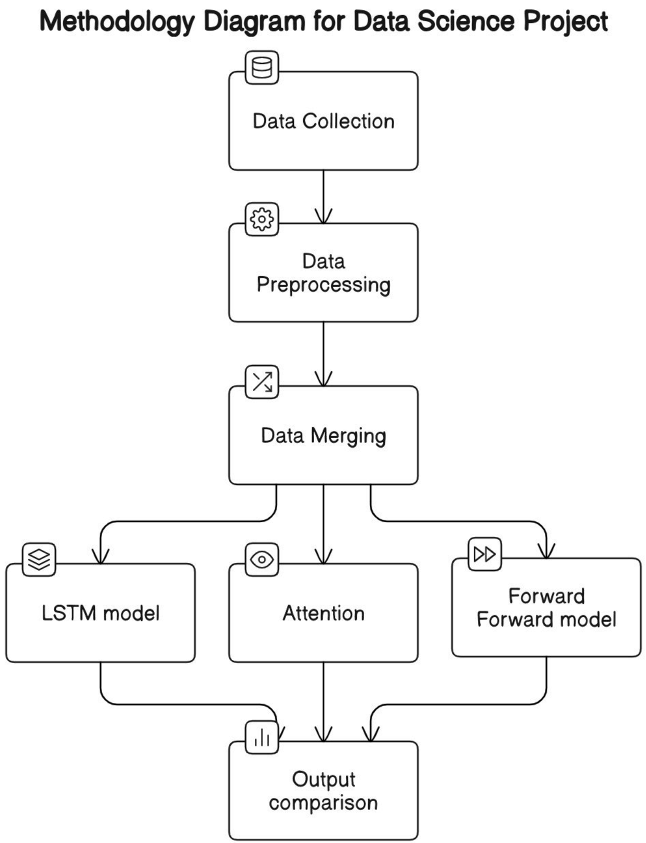
- **Purpose:**
  - Our project explores Climate Change to predict Total Natural Disasters using Machine Learning. This aligns with the course objective to expand our knowledge in Data Science
- **Learning Knowledge:**
  - We expect to learn how to program in TensorFlow, work with Jira project management software, build advanced deep learning architecture, understand the basic steps of research, and think critically of new ideas
- **Application of Knowledge:**
  - Sean and Ataish learned and used TensorFlow, Kevin learned and used Jira project management, Kevin learned how to perform basic research experimentation, and Kevin extended the Forward-Forward algorithm.

Student Objectives (Continued).

- **Critical Thinking and Problem Solving:**
    - Kevin extended the Forward-Forward algorithm to predict continuous values, Sean learned how to resolve issues with the attention model architecture, and Ataish learned how to better tune models.
- **Collaboration:**
    - All three of us learned how to collaboratively resolve issues with challenges as a team.
        - Issues with Dataset creation.
        - Tackling iterations together.
        - Brainstorm Ideas.
- **Presentation and Communication of Skills:**
    - We hope to convey through our presentation the importance of climate change.
    - Properly convey our methodology and findings in both written and verbal formats.

Problem Specification.

- **Problem Description:**
  - Climate change has had devastating impacts to the environment. Predicting Natural Disaster frequency can raise global awareness.
- **Scope:**
  - To predict Total Disasters with climate change data using advanced neural networks. This involves pre-processing, building of deep learning architectures, and research of an experimental method.
- **Impact:**
  - If there is no/minimal awareness to climate change, climate change can lead to an increase in natural disasters, which can have devastating impacts on the environments globally.
- **Challenges:**
  - Extending an experimental method may yield unsuccessful results.
  - Building deep learning architectures can be challenging.



Related Work.

Climate Change Dataset

- Tsui-Fang Hu et al utilized the IMF dataset to forecast inflation.
- We used the IMF dataset to predict frequency of Total Disasters as a time-series and continuous value problem.

LSTM Cell Model

- Liu et al. introduced an LSTM-based model tailored to predict greenhouse climate variables.
- The LSTM cell is known to perform sequence to sequence prediction.
- Ataish used the LSTM cell to predict Total Disasters as a time-series problem.

Multi-Head Attention Model

- Google proposed the Attention is All you Need algorithm that introduces Multi-Head Attention.
- The Attention is All you Need is well known to perform sequence to sequence predictions.
- Sean utilized the Multi-Head Attention neural network to build a deep learning model to predict Total Disasters.

Forward-Forward Continuous Model

- Gandhi extended the Forward-Forward Algorithm to classify sentiment of words.
- Gandhi's method extended the Forward-Forward algorithm by concatenating predicted activation layers.
- The Forward-Forward Continuous model uses concatenation of predicted activation layers to predict continuous values.

T.-F. Hu, I. G. Luja, H. Su, and C.-C. Chang, "Forecasting inflation under globalization with artificial neural network-based thin and thick models," in World Congress on Engineering and Computer Science, USA, Citeseer, 2007, pp. 909–914.

A. Vaswani, N. Shazeer, N. Parmar, et al. Attention is all you need, 2017. arXiv: 1706.03762 [cs.CL].

Y. Liu, D. Li, S. Wan et al. "A long short-term memory-based model for greenhouse climate prediction," International Journal of Intelligent Systems , vol. 37, no. 1, pp. 135–151, 2022

S. Gandhi, R. Gala, J. Kornberg, and A. Sridhar, Extending the forward forward algorithm, 2023. arXiv: 2307.04205 [cs.LG].

Amendments.

- A simple RNN model was added to help with the tuning process for the LSTM and Attention models.
- A sequencing step was added to the training of the LSTM and Attention models to better represent the shape of the problem.
- Forward-Forward Extension:
  - Kevin extended the Forward-Forward algorithm to predict continuous values.
- The Attention Neural Network was originally designed sequentially, but was switched to concatenating the multi-headed attention layer to the LSTM model.
  - This lead to the attention layer having more of an affect on training and yielded better results.

Solution Design.

Original Dataset.

- **Dataset Overview:**
  - Names (Total = 6) -
    - Surface Temperature, Forest and Carbon, Land Cover Accounts, Sea Level Change, and Frequency of Climate Disasters.
  - Source -
    - International Monetary Fund.
  - Purpose -
    - Mostly used for inflation, poverty and growth, and outlining changes of natural resources or human imprint.
- **Data Size:**
  - Volume -
    - 930 rows and 30 columns (216 KB).
  - Features (Total = 23)-
    - 13 features are land cover, 6 features are forest and carbon, 1 feature is seas/oceans, 1 feature is surface temperature, Country, and Year.
  - Targets -
    - Drought, Extreme Temperature, Flood, Landslide, Storm, Wildfire, Total.
- **Data Collection:**
  - The data is downloadable from the I.M Fund's website as 5 downloadable csv files. All csvs had to be consolidated for machine learning.

Data Pre-Processing.

- **Data Quality and Cleanliness:**
  - Missing Values were replaced with the column’s mean (for a particular country) except for frequency of natural disasters were replaced with zeros.
  - Outliers and duplicated records were removed from the dataset.
- **Preprocessing & Transformations:**
  - We filtered on the top 30 countries to extend the dataset from 29 samples to 930 samples.
  - Data was sorted by Country, Year and normalized with z-score for continuous value prediction and MinMax for time-series predictions.
  - Time series predictions grouped the data by Country, Year in sequences.
- **Train/Test Split:**
  - Time-series predictions used 70% train and 30% test split.
  - Continuous value prediction used 80% train and 20% test split.

Models Chosen.

1. **LSTM:**
    - A deep learning sequential neural network that allows information to be persistent during the learning process.
    - This model was chosen because of its capability to use long and short term memory to do sequence to sequence learning.
2. **Attention with LSTM:**
    - A deep learning sequential neural network that allows multi-head attention applied to sequential data during the learning process.
    - This model was chosen because its capability to use long and short term memory with attention to do sequence to sequence learning.
3. **Forward-Forward Continuous:**
    - A new neural network concept that uses activation layers as inputs to train a linear layer for continuous value prediction.
    - This model was chosen because of its ability to predict continuous values.

LSTM Model - Workflow.

- Initial LSTM Model (MSE- 34.0):

Objective: Establish baseline with climate indicators from 1992-2020.
Architecture: Single LSTM layer with 3-year input sequence.
Challenges: Limited feature set, resulting in moderate prediction accuracy.

- Second LSTM Model (MSE- 12.2):

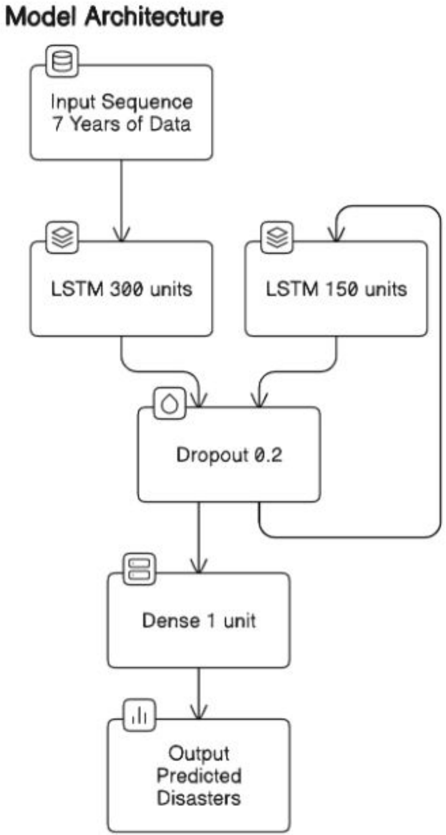
Enhancements: Added complexity with additional LSTM layers and dropout regularization.
Hyperparameter Tuning: Implemented L2 regularization, adjusted dropout rates.
Outcome: Higher MSE indicated need for further refinement despite increased sophistication.

- Final and Best-Performing LSTM Model (MSE- 7.7):

Innovation: Extended input sequence to 7 years for richer temporal learning.
Simplification: Streamlined to two LSTM layers, optimized for predictive strength.
Results: Significantly improved MSE and MAE, achieving robust and reliable predictions.

LSTM Architecture.

- **Temporal Depth (7-year sequence):** Captures long-term climate patterns for enhanced prediction accuracy.
- **First LSTM Layer (300 units):** Understands complex dependencies, providing a strong foundation for data analysis.
- **Second LSTM Layer (150 units):** Refines predictions, focusing on the most relevant features for disaster forecasting.
- **Dropout Regularization, after each LSTM layer (20%):** Prevents overfitting, ensuring model reliability on unseen data.
- **Output Layer (Single unit):** Condenses learning into precise disaster predictions.
- **Optimizer (Adam):** Dynamically adjusts the learning rate, ensuring efficient convergence to the optimal solution without overshooting, leading to faster and more stable training.
- **Activation Function (tanh):** Used in the LSTM layers, helps in managing the gradient flow, allowing the model to effectively learn from both short-term and long-term dependencies within the data.



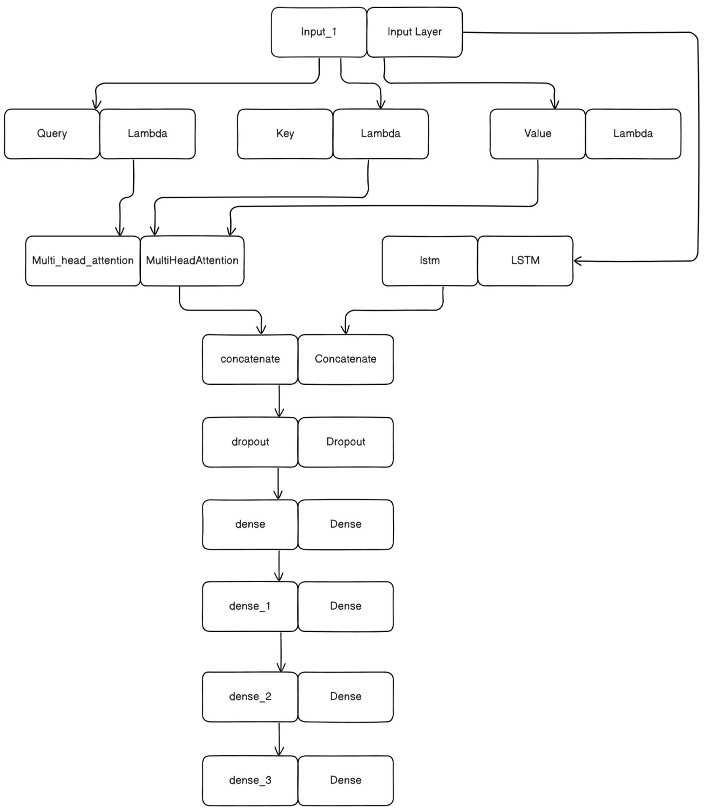
RNN Model.

- 125 units, ReLu Activation, Adam Optimizer, and MSE Loss
- A simple Recurrent Neural Network was used as a baseline model for figuring out how to properly structure the dataset and hyperparameter tuning.
- This was done to help improve the LSTM and Attention Neural Networks.
- Its performance helped discern data preprocessing needs and paved the way for the LSTM's enhanced temporal analysis capabilities.
- Specifically for the Attention model, this was a helpful tool in when there was issues with it performance and we did not know whether it was the model or if it was the way the data was being preprocessed.

Attention Model.

- The Attention model was built using Tensorflow Keras.
- 125 Epochs with batch size of 30
- 31 units, Adam Optimizer, and MSE Loss
- The attention transformer was implemented onto an LSTM neural network through concatenation.
    - The attention transformer is a neural layer that solely uses attention mechanisms for tuning.
    - This is done to help the model learn sequencing better.
- It also contains 5 additional layers:
    - 1 dropout layer with a rate of 0.1
    - 3 Dense Layers with ReLu activation.
    - 1 final dense layer but with softmax activation.

ATTENTION MODEL ARCHITECTURE.



Forward-Forward Model.

- Uses two forward passes with positive and negative goodness to classify inputs.
- Training of goodness function is performed before neural activities are passed to next layer.
- Each layer is trained independently where positive goodness is maximized while negative goodness is minimized.
- Hot encoding is used to create positive and negative data as signals.
- Training the goodness function is independent from forward predictions.

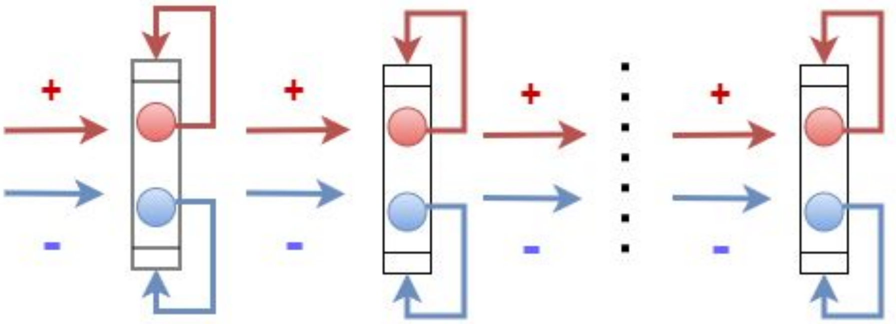


Figure 1: *This visual showcases the training architecture for the Forward-Forward Algorithm where (red) represents the positive goodness and (blue) represents negative goodness that is either trained inside the layer itself or passed to the next fully connected layer*

Forward-Forward Model (Continued).

- The Forward-Forward network only needs forward predictions for layer activities instead of using training sequence.
- Forward-Forward network is fully trained before continuing with the training process.
- Layer activations from Forward-Forward network are concatenated before training traditional linear layer.
- Traditional linear layer is trained without backward propagation, which uses function derivatives to correct errors.

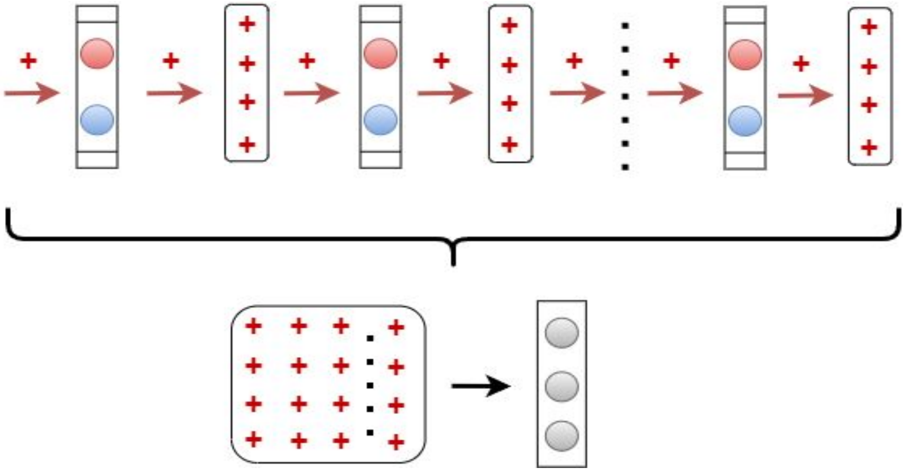


Figure 2: *The top architecture is the prediction of the Forward-Forward algorithm that yields positive (+) goodness from each fully connected layer. Each positive goodness' are concatenated to make a matrix that is passed into a singular linear layer for training and predictions.*

Tool List.

- Python
- Github
- Jupyter-Lab and Jupyter Notebook
- Pytorch
- Tensorflow and Tensorflow Keras
- Numpy
- Pandas
- SKLearn
- MatplotLib’s Pyplot

Time Schedule (using Jira).

Sample Video/Demo.

Code JupyterLab Python 3 (ipykernel)

[1]:
```python
import pandas as pd

# Load the dataset
dataset_path = "C:/Users/ATAISH NEHRA/Downloads/merged_natural_disaster_dataset_1992_2020.csv"
data = pd.read_csv(dataset_path)
```

[2]:
```python
# Display the first few rows of the dataset to verify it's loaded correctly
print(data.head())
```

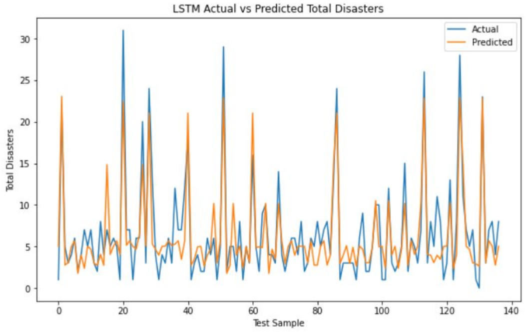
```
       Country  Year  Temperature  Drought  Extreme temperature  Flood  \
0  Afghanistan  1992       -0.294        0                    0      2
1  Afghanistan  1993        0.220        0                    0      0
2  Afghanistan  1994        0.430        0                    0      0
3  Afghanistan  1995        0.359        0                    0      2
4  Afghanistan  1996       -0.116        0                    0      1

   Landslide  Storm  Total Disasters  Wildfire  ...  \
0          0      0                2         0  ...
1          2      1                3         0  ...
2          0      0                0         0  ...
3          1      0                3         0  ...
4          0      0                1         0  ...

   Sparsely natural vegetated areas: Climate neutral  \
0                                         4275.2553
1                                         4275.2243
2                                         4275.2243
3                                         4227.8536
4                                         4220.7930

   Terrestrial barren land: Climate neutral  \
0                                24963.0009
1                                24963.0009
2                                24963.0009
3                                25088.5751
4                                25108.3556
```

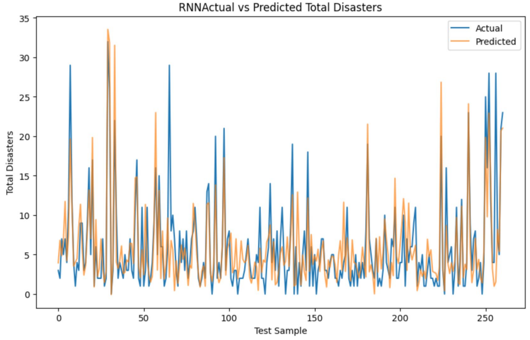
Postmortem.

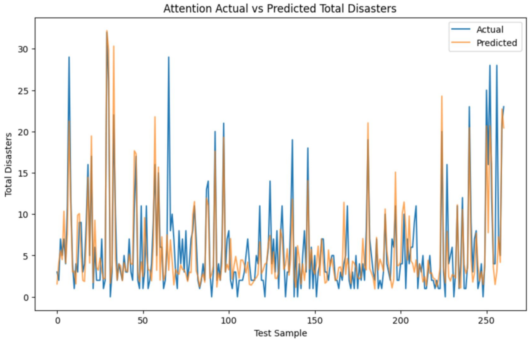
- Original goal of expanding to more granular data points for dataset.
- Expand the dataset to include more countries than the 30 chosen.
- Expand prediction to include the specific type of disasters instead of just broad total disasters.
  - Learn what specific climate change indicators are better predictors for each type of disaster.
- Forward-Forward RNN
  - Kevin decided to discontinue exploration of the Forward-Forward RNN architecture since predictions of continuous values would need to be researched and extended before building the Forward-Forward RNN architecture. Not enough time permitted research of both methodology along with proper experimentation and documentation.
  - This allowed Kevin to focus on one extension instead of two, which allowed the proper amount of time to experiment and document of one extension.

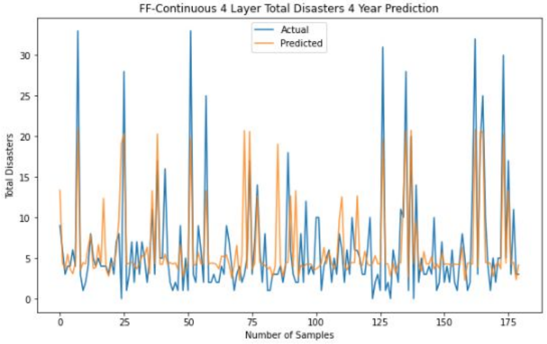
Results and Deliverables.

Actual vs Prediction graphs.







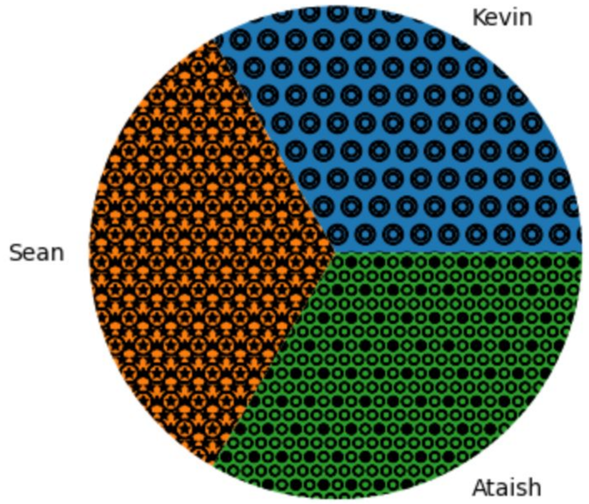


Comparison.

Models	MSE	MAE
LSTM	7.78	2.13
RNN	15.205	2.58
Attention	13.606	2.39
Forward Forward	20.45	3.09

Contributions.

- **Kevin Russell:**
    - Forward-Forward Continuous Model.
    - Report Writer.
    - Pre-processing of Data.
    - Torch Batch DataLoader.
    - Handling of Total Disasters and Surface Temperature.
    - Project Management in Jira.
- **Sean Klein:**
    - Attention Neural Network.
    - Simple RNN model.
    - Pre-processing of Data.
    - Report Writer.
    - Initial Handling of Sea Levels and Co Concentration datasets.
    - Project Leader.
- **Ataish Nehra:**
    - LSTM model.
    - Report Writer.
    - Pre-processing of Data.
    - Initial Handling of Forest & Carbon and Land Cover Accounts datasets.
    - Sequencing the data set.
    - Lead Programmer.



Future Enhancements.

- Extend the database to include more countries to make the models more global.
    - This will also make a larger dataset which will help improve accuracy.
- Find other datasets that can be joined with ours to increase feature space.
- More feature selection and feature construction.
- More experimentation on the forward forward model with time-series problems.
- Further hyper-parameter tune attention model to make it more competitive.
    - Increase number of units per layer, try different activation functions, apply multi-year sequencing, etc.

https://spendmatters.com/2019/05/24/dont-forget-the-big-4-questions-to-ask-during-any-mega-acquisition/

Questions?

DEMO.

```
[1]: import pandas as pd

     # Load the dataset
     dataset_path = "C:/Users/ATAISH NEHRA/Downloads/merged_natural_disaster_dataset_1992_2020.csv"
     data = pd.read_csv(dataset_path)

[2]: # Display the first few rows of the dataset to verify it's loaded correctly
     print(data.head())
```

```
       Country  Year  Temperature  Drought  Extreme temperature  Flood  \
0  Afghanistan  1992       -0.294        0                    0      2
1  Afghanistan  1993        0.220        0                    0      0
2  Afghanistan  1994        0.430        0                    0      0
3  Afghanistan  1995        0.359        0                    0      2
4  Afghanistan  1996       -0.116        0                    0      1

   Landslide  Storm  Total Disasters  Wildfire  ...  \
0          0      0                2         0  ...
1          2      1                3         0  ...
2          0      0                0         0  ...
3          1      0                3         0  ...
4          0      0                1         0  ...

   Sparsely natural vegetated areas: Climate neutral  \
0                                          4275.2553
1                                          4275.2243
2                                          4275.2243
3                                          4227.8536
4                                          4220.7930

   Terrestrial barren land: Climate neutral  \
0                                24963.0009
1                                24963.0009
2                                24963.0009
3                                25088.5751
4                                25108.3556
```

Conclusion.

Explicitly say which model succeeded

Say how this (potentially?) shows that climate change has an explicit effect on the rise in Natural Disasters that occur

Why our project/climate change research matters

Future Work

What can we or people who are interested in this project can help to expand on this project to either generalize it or expand the scope