

**Utilizing reinforcement learning and deep neural networks to optimize non-pharmaceutical
COVID-19 interventions in Florida**

Megan Yang, Leya Joykutty

Abstract

Under the umbrella of artificial intelligence, machine learning is a technique that allows a system to improve through experience without any explicit programs. One approach to machine learning used in this study is neural networks, in which the computer learns to finish a task by analyzing training samples. The other approach used is reinforcement learning which manipulates a programmed environment and discovers errors and rewards to output the optimal solution.

This study created a deep neural network and used reinforcement learning to develop a model that could predict whether the number of COVID-19 cases would increase or decrease. Then using that information, the model predicted which actions would most effectively cause a decline in cases while keeping things like economy and education in mind for a better long-term effect. These models were made based on eight different Florida counties' data, including mobility, temperature, dates of government actions, etc. Based on this information, data exploration and feature engineering were conducted to add dimensions that would further the accuracy of the neural network.

The reinforcement learning model's actions consisted of a shutdown for about two months before reopening schools and allowing things to return to normal. Then interestingly, the model decided to keep schools operating in a hybrid model, with some students going back to school while others continue to study remotely.

These models can be generalized to train agents for taking action against policies in areas beyond COVID-19. Using neural networks and reinforcement learning in government decision-making could push for better measurement and consideration for education, local economy, and other essential factors that impact people's day-to-day lives. Additionally, there is an extra precaution against harmful short-sighted decisions with long-term negative impacts.

Background Research

Machine learning is a subset of artificial intelligence (AI) that allows a system to learn and improve through experience without explicit instructions and programs. Machine learning algorithms find patterns in huge amounts of digitally stored data ranging from images to numbers and words using statistics. Whatever platform the machine learning is applied to, it collects as much data and information as possible to find the most accurate patterns and applies those patterns. These can be found on websites like Netflix, Youtube, Spotify, etc. as it makes recommendations based on the user's history (What is Machine Learning? 2020).

Neural networks are an approach to machine learning where the computer learns to complete a task by analyzing training samples. For example, thousands of images are fed into the system to train the computer to recognize the object to identify objects in images. They are based on the human brain, both containing thousands of connected processing nodes. These networks have several layers of nodes; an individual node would be connected to several nodes beneath from which it receives data, and above that it sends data to. These nodes are assigned a "weight." When the neural network is in use, each node receives different data items and multiplies them by the weight. These products are then added to produce a sum that is compared to a threshold value. When the value is above the threshold, it continues to pass the data to the next layer; if the value is below, it stops and passes on no data. Initially, the weights and thresholds are randomized. As training continues, they get modified to become more and more accurate (Hardesty 2014).

Reinforcement learning (RL) manipulates its environment to discover errors or rewards. It is a trial and error approach that aims to get as many rewards as possible. This required reward feedback is known as the reinforcement signal. This type of learning allows systems to determine the best behavior to maximize the rewards (Hao 2020). Through this seeking of rewards, the computer is susceptible to finding unexpected solutions to doing so. Human involvement is very limited to making slight modifications to the environment or rewards and penalties. Thus, the computer is ensured not to act expectedly and provides a new way to do things more optimally than the usual. The main component that distinguishes reinforcement learning from others is how it is trained. It manipulates and interacts with the environment instead of analyzing and observing the dataset provided. (Osiński and Budek 2018). Proximal Policy Optimization (PPO) has become one of the most popular classes of reinforcement learning methods since its introduction by OpenAI in 2017. It is an "on-policy" approach since it updates its decision-making policy based on a small collection of experiences; this small batch of experiences only slightly alters and improves the policy once. As a result, there is less variance in training and overall smoother training, although there is a downside to introduced bias. Despite the bias, the PPO approach allows the agent more guidance and makes sure it does not go off-track too much (Trivedi 2019). Stable

baselines are important reinforcement learning algorithms to implement and make training the PPO agent significantly easier. On the other hand, for the environment, OpenAI Gym plays a large role in providing important algorithms (Raffin, 2018).

Hypothesis

If a neural network is modeled using the COVID-19 Florida counties' data, then the neural network would accurately classify whether cases increase or decline based on given circumstances.

If an agent and environment are created based on the circumstances due to COVID-19, then reinforcement learning can decide what would be the best course of action.

The purpose of this experiment is to design a neural network and use reinforcement learning to design a plan to deter the spread of COVID-19 while taking into account not only the number of cases but economy and education.

Materials and Methods

Datasets

- Florida Department of Health - Florida COVID19 Case Line Data
- Google mobility
- National Oceanic and Atmospheric Administration (NOAA) - Monthly Climate Data
- Ballotpedia - Florida's path to discovery from COVID-19

Procedure

1. Collect datasets about 8 counties, clean up data, get rid of empty rows, and merge relevant datasets together.
2. Next, use a multilayered perceptron to process the data.
 - a. Starting off with the number of layers equal to the number of variables involved.
 - b. The first layer will start off with a fairly large number of perceptrons the rest will have a smaller number until the output layer will have 1.
 - c. The activation functions will be ReLU for hidden layers and Sigmoid for the output layer.
 - d. Batch and epoch sizes must be chosen then the data would be put through to test for accuracy and precision.
 - e. The number of layers and perceptrons per layer, batch sizes, and epoch sizes will be adjusted until a satisfactory level of accuracy and precision are achieved.
 - f. Dropout layers are added to prevent overfitting

3. Put the data through PPO reinforcement learning.
 - a. Design the environment by including dimensions such as economy, number of cases, and education.
 - b. Design an agent that takes actions.
 - c. Create functions that cause the environment to be affected by the actions the agent takes.

Results

After graphing the data, running the neural network, and using reinforcement learning, the following results were obtained.

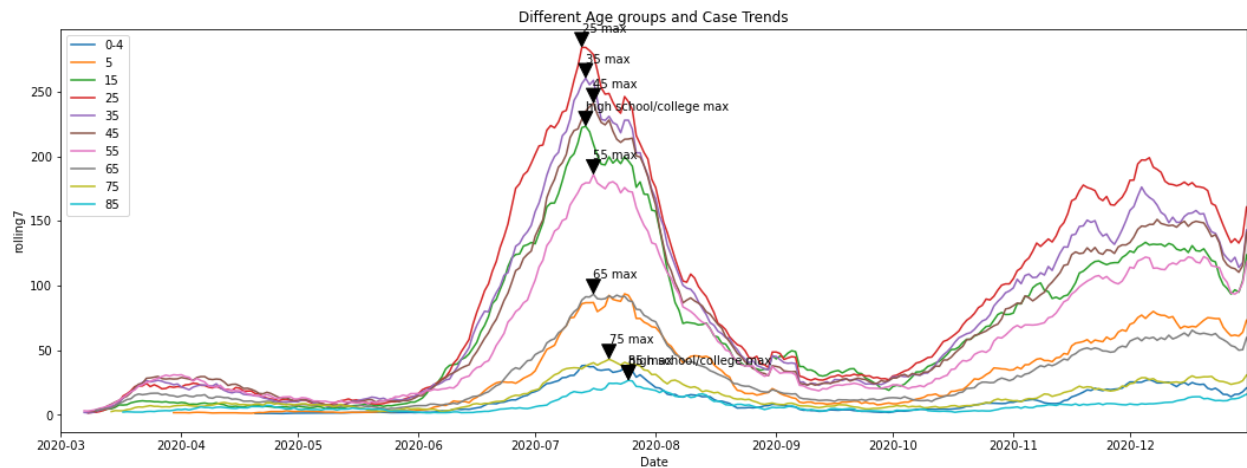


Figure 1. The graph depicts the different peaks in COVID-19 cases while using the average number of cases for the week called a rolling 7 average for various age groups at their respective times.

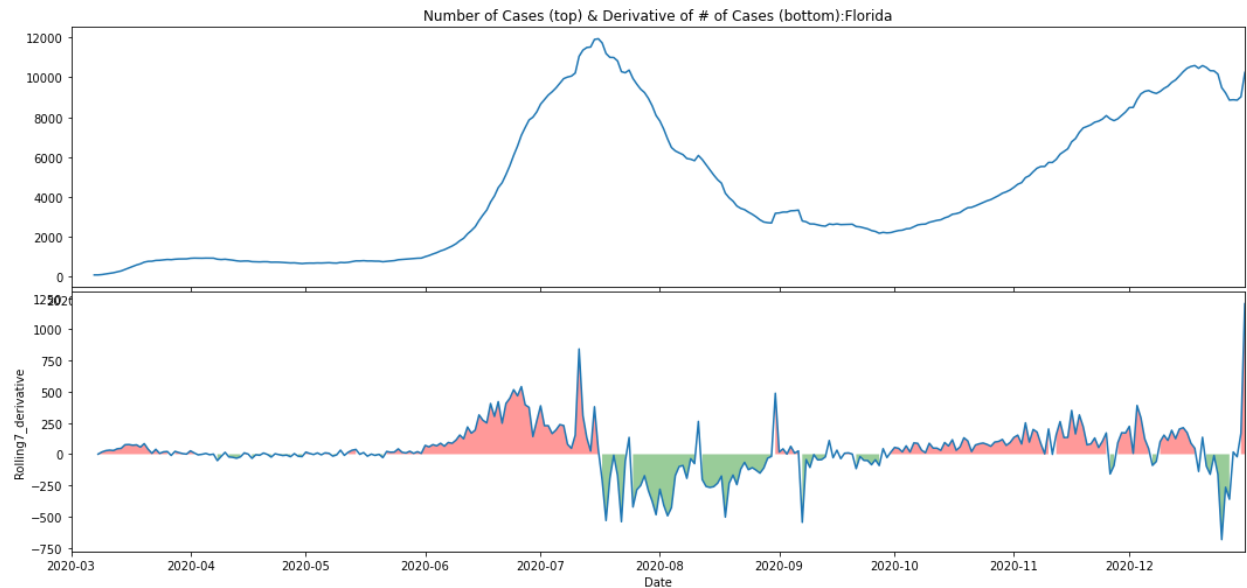


Figure 2. The graph shows the number of cases as well as the derivative of each point for the cases in

Florida to display its differences compared to the rest of the individual counties.

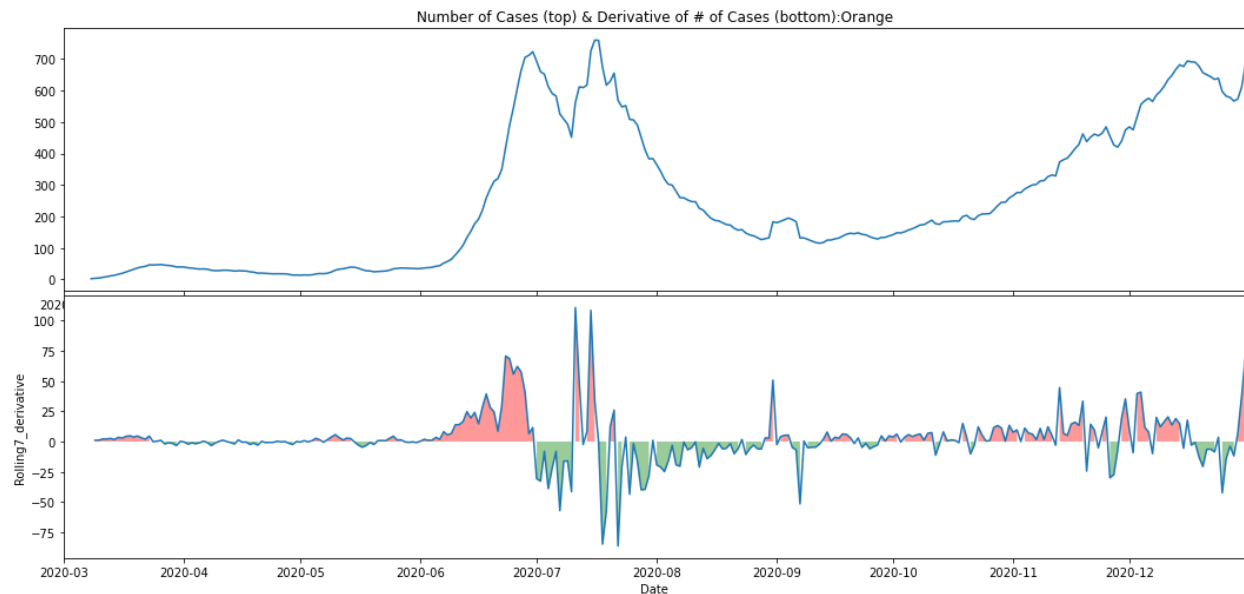


Figure 3. The graph shows the number of cases as well as the derivative of each point for the cases in Orange county to display its differences compared to the rest of the counties and Florida as a whole.

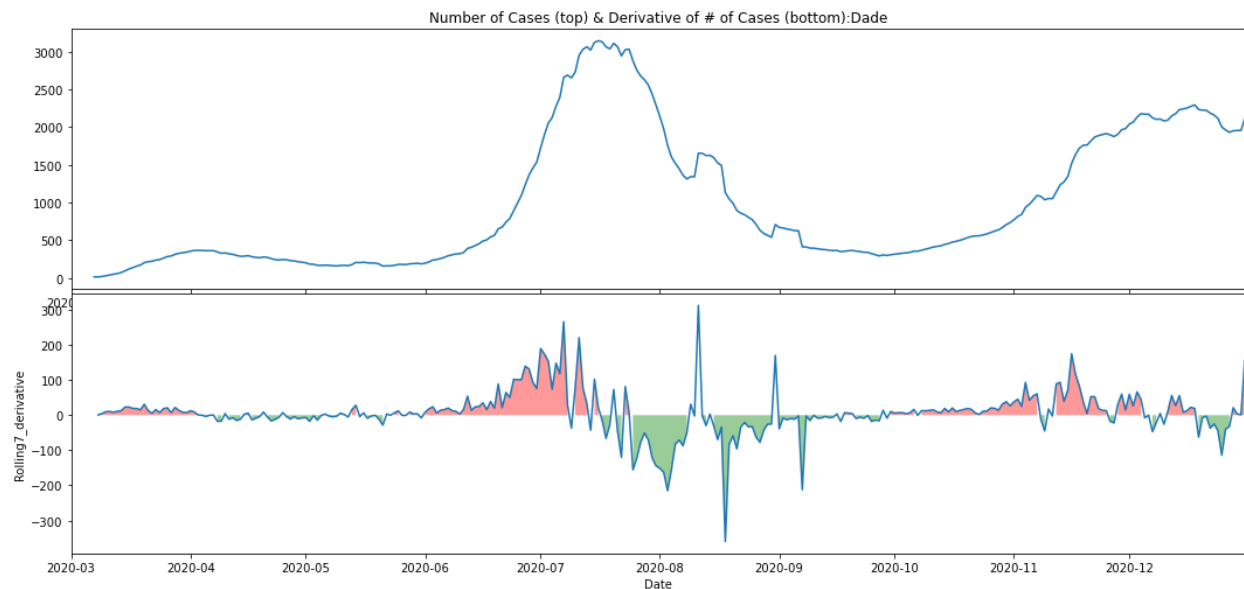


Figure 4. The graph shows the number of cases as well as the derivative of each point for the cases in Miami Dade county to display its differences compared to the rest of the counties and Florida as a whole.

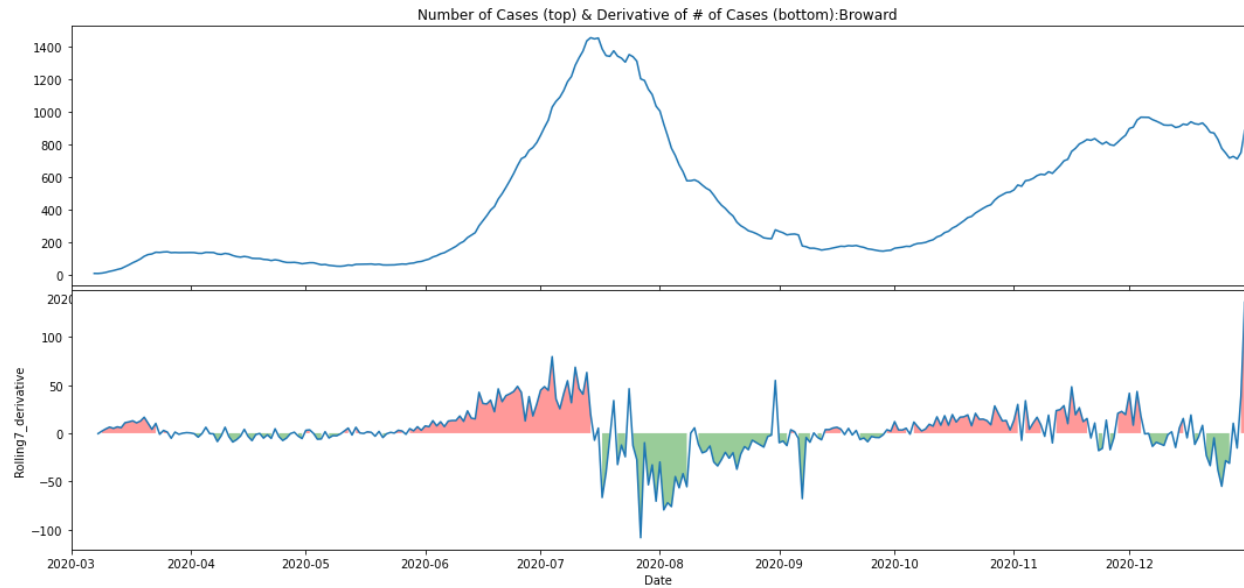


Figure 5. The graph shows the number of cases as well as the derivative of each point for the cases in Broward county to display its differences compared to the rest of the counties and Florida as a whole.

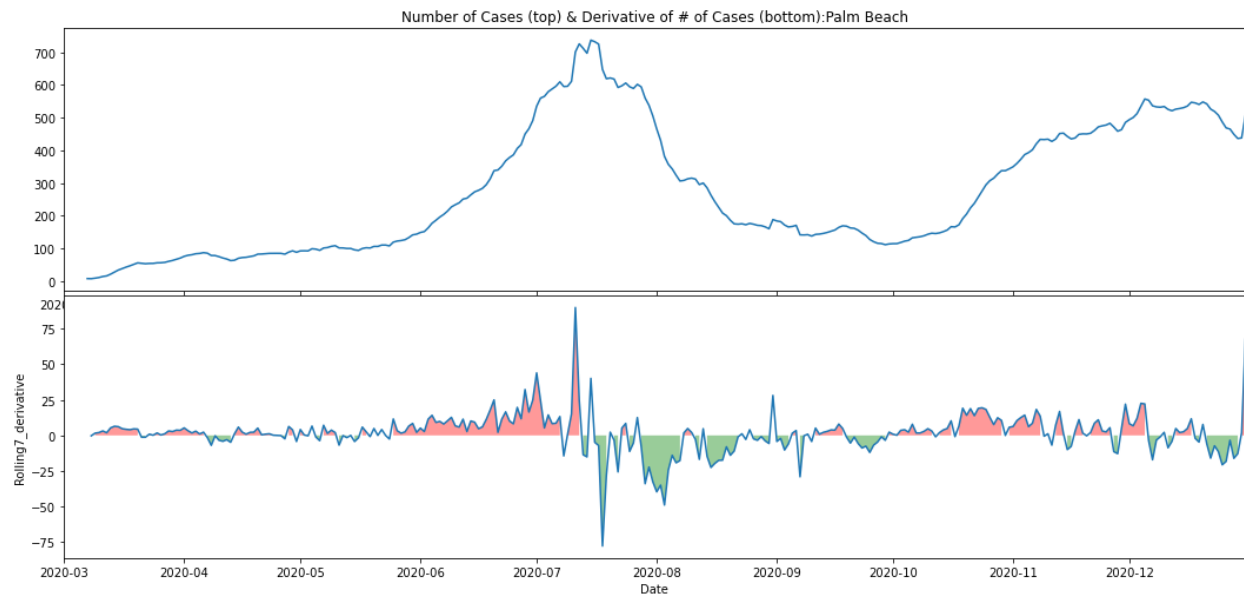


Figure 6. The graph shows the number of cases as well as the derivative of each point for the cases in Palm Beach county to display its differences compared to the rest of the counties and Florida as a whole.

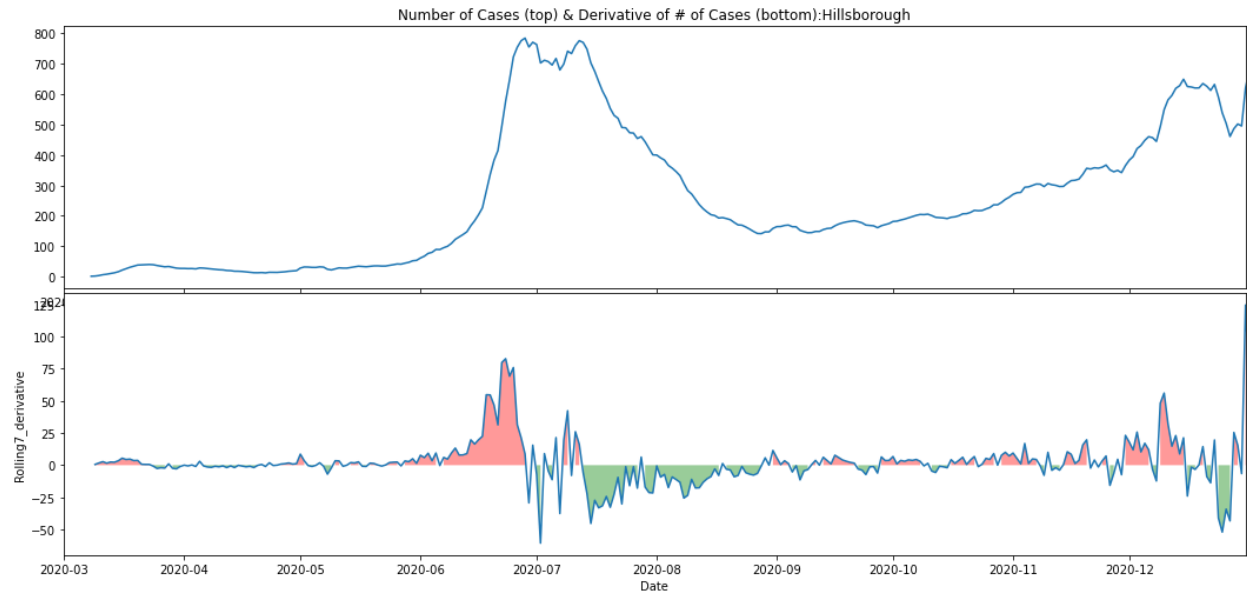


Figure 7. The graph shows the number of cases as well as the derivative of each point for the cases in Hillsborough county to display its differences compared to the rest of the counties and Florida as a whole.

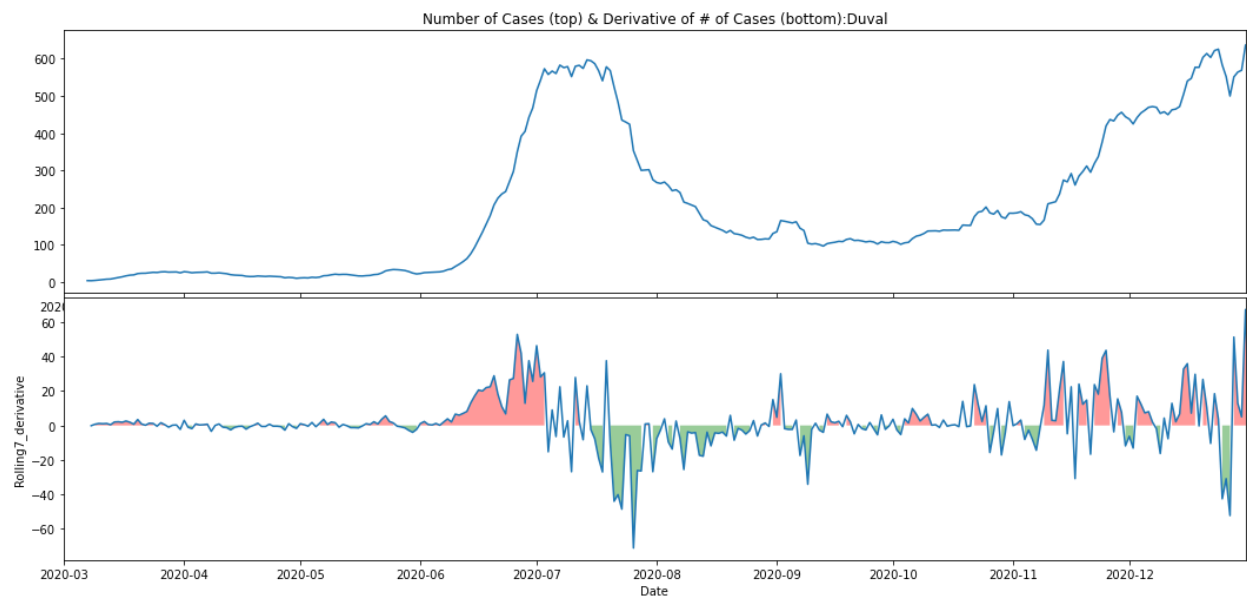


Figure 8. The graph shows the number of cases as well as the derivative of each point for the cases in Duval county to display its differences compared to the rest of the counties and Florida as a whole.

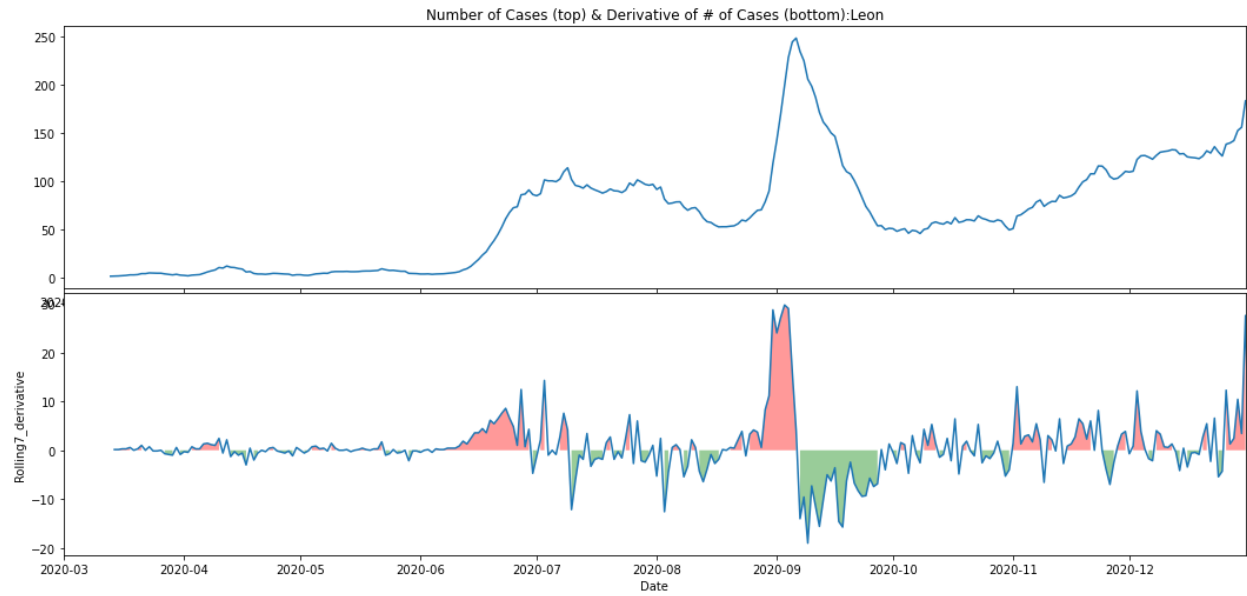


Figure 9. The graph shows the number of cases as well as the derivative of each point for the cases in Leon county to display its differences compared to the rest of the counties and Florida as a whole.

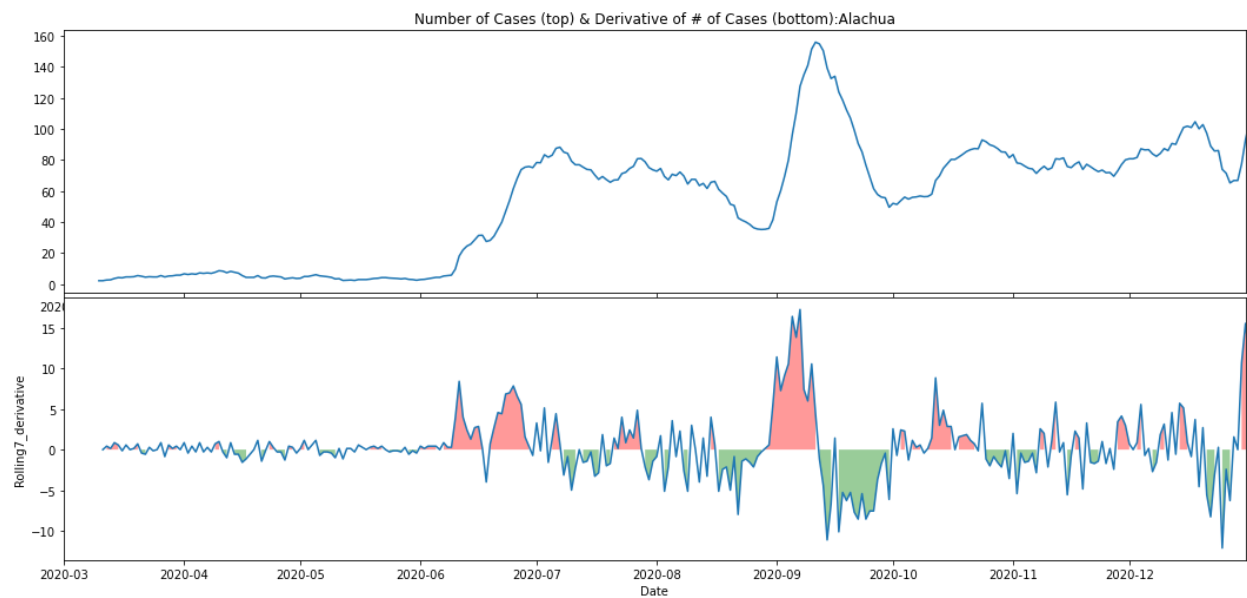


Figure 10. The graph shows the number of cases as well as the derivative of each point for the cases in Alachua county to display its differences compared to the rest of the counties and Florida as a whole.

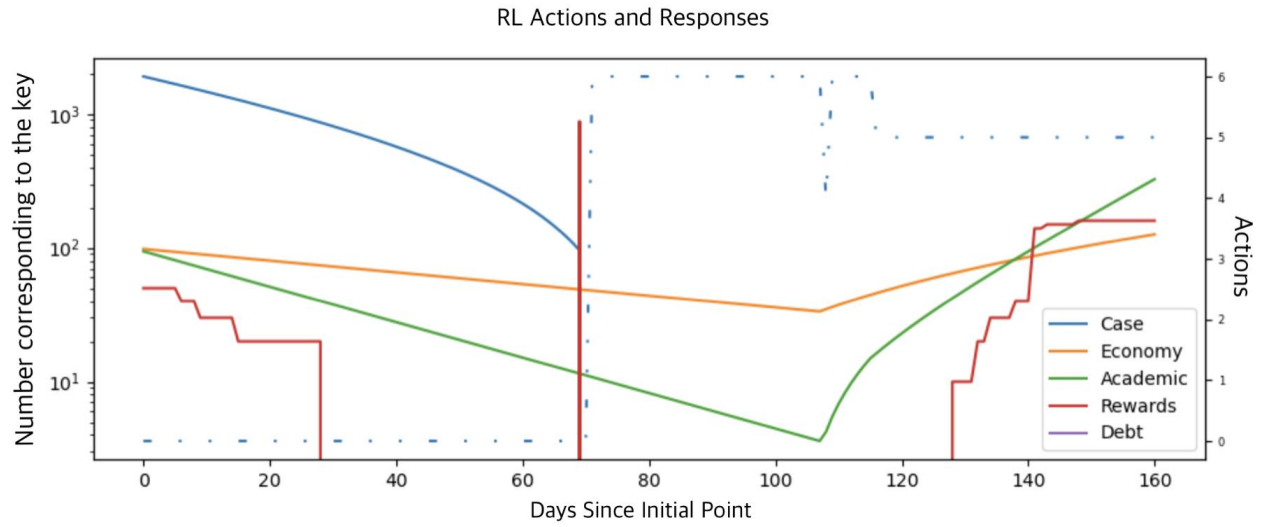


Figure 11. The dotted line shows the action taken and the colored lines show the results of those actions generated by reinforcement learning. The actions correspond so that Stay Home = 0, Phase 1 = 1, Phase 2 = 2, Phase 3 = 3, Normal = 4, School Hybrid = 5, and School Normal = 6.

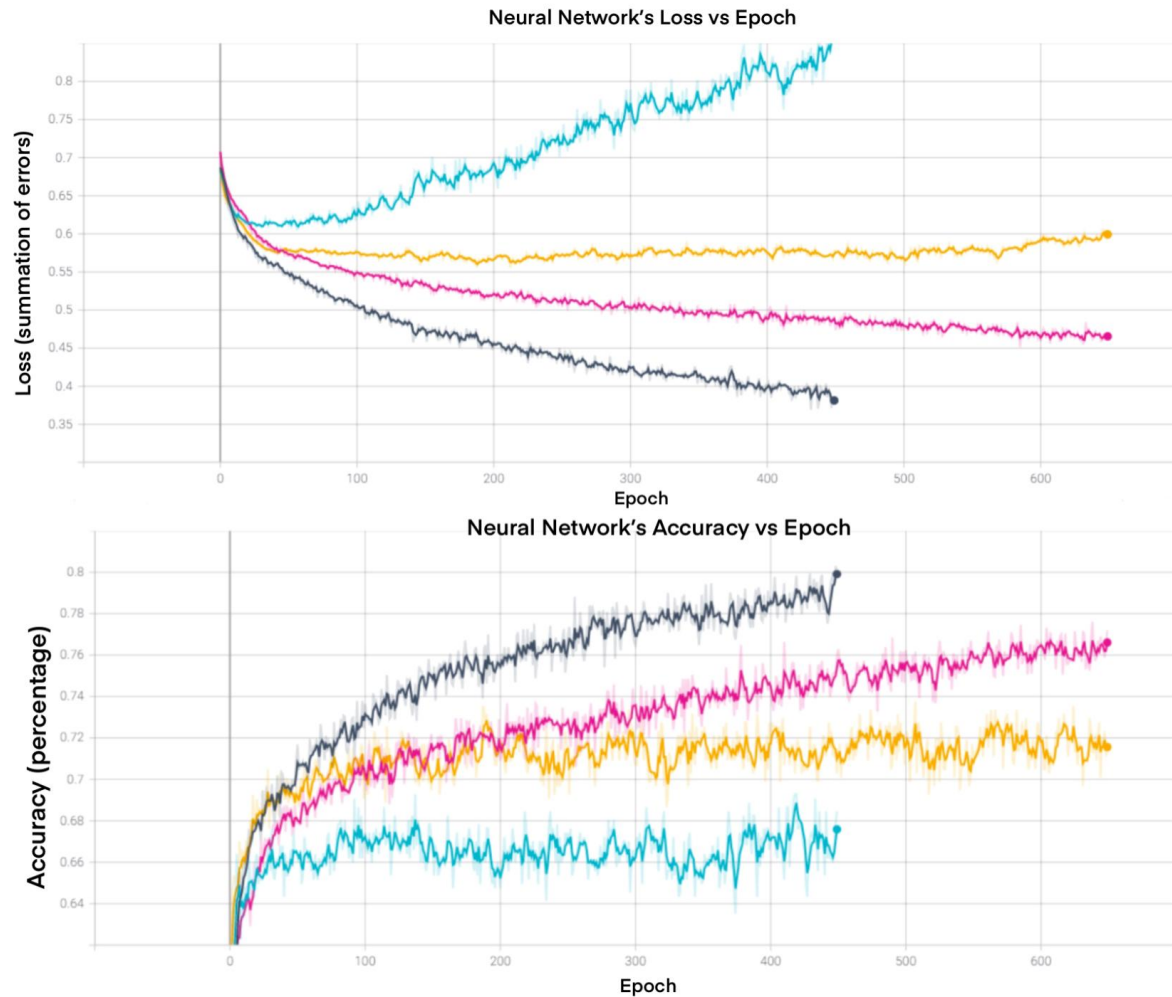


Figure 12. The dark blue and light blue lines show the neural networks loss, validation loss, accuracy, and validation accuracy before including dropout nodes. The pink and yellow lines show the same thing but includes the dropout nodes making the neural network more accurate and have less loss as time passed. Validation accuracy and loss are based on the training part of the dataset.

The classification deep learning neural network had five layers, two of which had the dropout feature, a layer that sets random edges of neurons to zero to help prevent overfitting. The data collected was used to predict whether the number of cases increased or decreased with binary cross-entropy as the loss function and accuracy as the metric. I also calculated validation accuracy and loss to avoid overfitting. After 2000 epochs, the final model yielded a validation accuracy of 0.7152, an accuracy of 0.7720, a loss of 0.4640, and a validation loss of 0.5994 as seen in Figure 12. Additionally, the validation accuracy and loss are similar to the accuracy and loss in value, meaning the neural network does not overfit by too much.

Both feature engineering and data exploration expanded on and prepped the data for the neural network to help illustrate the necessary dimensions for the reinforcement learning environment. For example, looking at each counties' graph of cases, the peaks varied from county to county. Alachua's and Leon's cases peaked after school started due to college towns containing universities like the University of Florida or Florida State University. The neural network helped identify that the general age composition was a significant factor.

The reinforcement learning took steps that would lead to the decrease in COVID-19 cases to a threshold that deemed it low enough to be considered over while ensuring that the economy and education had not faltered. First, a shutdown was done for about two months before reopening schools allowing things to return to normal. Then interestingly, the model decided to keep schools operating in a hybrid model, with some students going back to school while others continue to study remotely.

Conclusion

The neural networks are shown to have been able to successfully predict the rise and decline of cases in a county, given data for each day about the county. This experiment utilized data about people's mobility related to retail, grocery, park, transit, workplaces, and residential, how many cases were recorded that day, the temperature, and the number of days since the implementation of government action in eight of the largest or most populated counties in Florida: Alachua, Broward, Miami Dade, Orange, Palm Beach, Leon, Hillsborough, and Duval. From these, several more dimensions were added such as the minimum, maximum, and average temperature for each day. Next, I averaged the number of cases over seven days to calculate the rolling seven average. Additionally, I took the derivative of both the rolling seven averages and the raw data for the number of cases to see whether the number of cases increased or decreased.

Instead of regression, I created a classification network to predict whether the cases would rise or fall. Because each county's data vary drastically from one another, each county had to be analyzed separately. For example, a network modeled to predict a largely populated county like Broward cannot be used to predict Alachua county which is more of a college town due to large differences in population composition. Unfortunately, there is not enough data to train a regression neural network. I had to concatenate the eight counties' data to create a large enough dataset and train appropriately for a classification network. Therefore, the only accurate possible model is a classification model. A universal model for predicting COVID-19 cases is not practical. For example, a model trained using the full Florida dataset will not work for counties like Alachua and Leon due to their dramatically different population. For models to work correctly, they need to be built and trained for each specific community.

Application

The neural network serves to predict whether the coronavirus cases would increase or decrease given the conditions of the current circumstances; this prediction leads to help the reinforcement learning model. The reinforcement learning model serves as a base to begin optimizing the possible actions the government can take to cause the number of coronavirus cases to decrease while maintaining a stable economy and education in Florida. With more massive amounts of detailed data specific to each county, these models can be even more accurate and consider the dimensions specific to each location. By doing so, the model would be able to more accurately approximate and find the best solution to the given circumstances.

Additionally, these models do not exclusively work just for the current pandemic, COVID-19. They can be used and slightly altered to fit past and future scenarios in which there is a contagious disease outbreak as long as the necessary data is inputted. Furthermore, there is extra precaution against harmful short-sighted decisions with long-term negative impacts and instead, prepare actions that will be better in the long run.

They can be generalized to train agents for taking action against policies in areas beyond COVID-19. Using neural networks and reinforcement learning in government decision-making pushed for better measurement and consideration for education, local economy, and other essential factors that should be considered as a factor that impacts people's day-to-day lives.

Future Research

If more data is collected for the most populated and largest counties throughout the country, this model can be applied to predict the rise and decline of COVID-19 cases throughout the country or even the world if the necessary data is available. Moreover, if more detailed information like the number of cases for people with immune deficiency or for different age groups was known, how many restaurants or other public spaces were opened or closed and how many employees tested positive, etc., then the network could account for significantly more features and provide an even more accurate form of prediction. The most up-to-date census data considers and has access to this sort of composition to the population in each county. Another important factor to account for when it comes to contagious diseases is population density since more people come in contact with each other more often. With the rise of COVID-19, the significance of masks also rose and should be accounted for, although the data on this is very limited and needs to be collected in order to be able to provide it as an additional dimension for input.

For further training, data is not limited to only to COVID-19. Data from diseases like the seasonal flu or other contagious diseases could be used in place of COVID-19 data to make up for the lack of

training for the neural network to better understand the important factors and have a large impact on different populations.

Furthermore, as cases increase towards an incredibly large number, neural networks can be implemented into the reinforcement learning model instead of manually entering each possible case. The neural network would take "action" and would use the environment data from the specific county as input and produce the predicted changes to the environment.

Limitations

There were numerous limitations in this experiment due to things like limited access to data and time constraints. Due to being a high school student with limited resources, it was extremely difficult to come up with enough data to do training. Moreover, when reaching out to companies or people to request for data, there was no response making only public databases available. To account for this lack in resources, I used eight counties in Florida to train the neural network instead of just one. Additionally, there is less than a year's worth of data available relating to COVID-19, causing the need for this concatenation of eight different counties even though they have unique and different makeups. Furthermore, even though it is evident that the different ages are impacted in different ways by the events that occur, data like the Google mobility data sets are not split up by age group and therefore make it impossible to be able to account for it in the neural network.

Error Analysis

The discrepancy between the test accuracy and accuracy displays a bit of overfitting, so random noise in the data makes it difficult to completely match up the accuracy and test accuracy. Moreover, it seems that the neural network, despite any alterations, cannot score a test accuracy of about 0.7152. This is largely due to the experiment's limitations, such as lack of data that made it difficult to allow the neural network to have a lot to work with.

Acknowledgements

I would like to thank Leya Joykutty and Kepa Oyarbide for guiding me through this process and mentoring me to help me get to where I am today in research.

References

- Brownlee, J. (2016, August 16). *Supervised and Unsupervised Machine Learning Algorithms*.
<https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/>
- Hao, K. (2018, November 17). *What is machine learning?* MIT Technology Review.
<https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/>
- Hardesty, L. (2017, April 14). *Explained: Neural networks*. MIT News.
<http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>
- Osiński, B., & Budek, K. (2018, July 5). *What is reinforcement learning? The complete guide*. deepsense.ai. <https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/>
- Raffin, A. (2018, August 21). *Stable Baselines: a Fork of OpenAI Baselines — Reinforcement Learning Made Easy*.
<https://towardsdatascience.com/stable-baselines-a-fork-of-openai-baselines-reinforcement-learning-made-easy-df87c4b2fc82>
- Society for Science and the Public (2017-18). International Science and Engineering Fair 2017-18: International Rules & Guidelines. Washington, DC: Society for Science and the Public.
- Trivedi, C. (2019, August 12). *Proximal Policy Optimization Tutorial (Part 1/2: Actor-Critic Method)*. Towards Data Science.
<https://towardsdatascience.com/proximal-policy-optimization-tutorial-part-1-actor-critic-method-d53f9affbf6>
- What is Machine Learning? A definition*. (2020, May 6). Expert System.
<https://expertsystem.com/machine-learning-definition/>