AI-Modeling and Ethics in the COVID-19 Pandemic

Kevin Wang

"When you're talking about things with big magnitudes, you need precise data, because those make big differences." - Charlie Baker

Background

AI and Machine Learning

Artificial intelligence, in extremely broad terms, is the ability for machines to demonstrate capabilities such as logic, reasoning, and creativity. In pop culture, AI has been portrayed as malicious robots and androids with consciousness and emotions. Modern technology has yet to reach this point. Instead, AI manifests in many everyday parts of our lives, such as Gmail's spam filter, Facebook ads, and Amazon recommendations.

Machine learning is a subset of AI and refers to the idea that machines can learn, adapt, and grow better through experience. In each of the above examples, AI works by training itself on pre-existing data and applying its "knowledge" to new examples. The program is given a labeled dataset, such as a set of emails that have been classified as spam or not-spam, or the purchase history of millions of Amazon users. It then learns to recognize patterns in the data and applies them to new situations.

AI has been able to automate mundane, time-consuming tasks such as these with near-perfect accuracy. As both hardware and software grow more advanced, it has become more powerful, with programs that can caption images and compete in games from Go to Dota.

Techniques - The Neural Network

Several programming techniques are prevalent in artificial intelligence, but the most well-known and successful is the artificial neural network (ANN). The technique is most clearly understood with a video and animation — I used 3blue1brown [1] — but I'll try to provide a brief summary here.

The ANN is modeled after neurons in the brain. A biological neuron is an elongated cell with two ends: the dendrite and the synapse. To model it simply, a neuron receives signals from many neighboring cells at its dendrite. Once these signals are cumulatively powerful enough, the neuron sends its own signal out through its synapse, to signal other neurons. The brain is made up of these layers and layers of connections, which combine to produce our cognition.

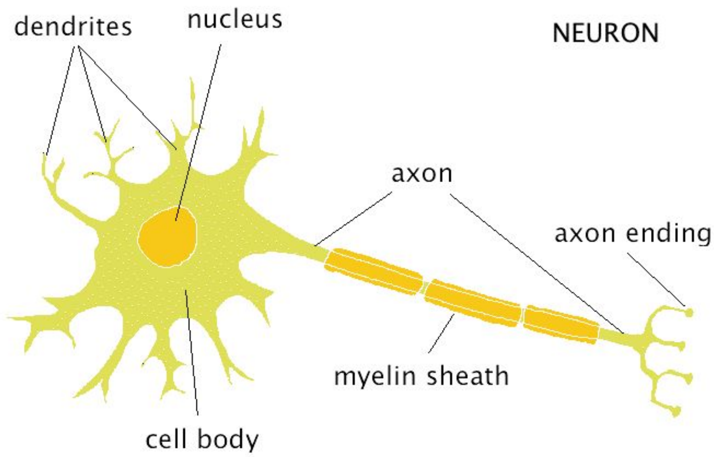


Diagram of a neuron [2]

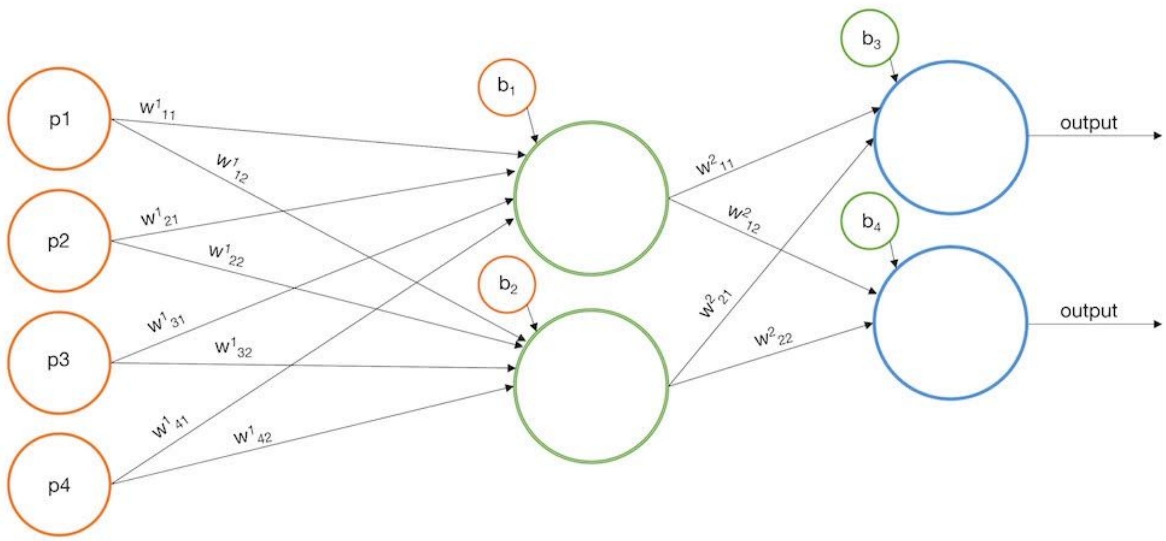


Diagram of an Artificial Neural Network [3]

The ANN models after this biological structure. Let's use the example of spam classification. The ANN begins with a set of input parameters (x_i), which are the words in the email. These parameters are the inputs to the next layer of "cells." The next layer consists of many cells, each of which takes in the parameters as inputs and outputs a single value. Each input is the one of the parameters multiplied by a certain weight (w_i). Each cell sums up these inputs with ($x_1w_1 + x_2w_2 + ...$) to create this output. At the end, the layer of cells all have unique outputs, which then become the inputs to the next layer of cells, and so on.

The final output of the ANN is one neuron, whose value signifies a certain meaning: in this example, how certain the algorithm is that an email should be classified as spam.

The ANN learns through training, where the weights begin as randomized values. A dataset with many, many emails (of known spamminess) is fed through the neural network. If the network correctly classifies the email, then the weights are reinforced. If it is incorrect, the weights are adjusted. The process may sound random, but over thousands of iterations, the network becomes increasingly accurate at classifying emails.

To clarify, the initial parameters are never adjusted by the neural network. These parameters are the input: an email and the words it contains. The model adjusts the weights between each cell, which in turn modifies the final output.

What makes an ANN so unique and powerful is the amount of complexity and nuance it allows for. The first layer of the network simply detects for the prevalence of certain words. Knowing that the words “trial,” “friends,” “free,” and “send” appear in the email doesn't mean the email is necessarily spam. Seeing the words “free” and “trial” next to each other may raise some red flags. And seeing the sentence “send to 10 friends and receive a free trial” means you probably have a match.

Ethics

Ethics, in the general sense, are principles that guide what is right vs what is wrong. As AI first emerged, the vast majority of research was done on hard science — algorithms and hardware that could make machines more advanced. Milestones were set in games, such as Deep Blue’s defeat against world chess champion Gary Kasparov in 1977. Today, AI is becoming increasingly prevalent in our daily lives, from Google search to algorithms that analyze peoples’ credentials to determine whether they should get home loans. As it has material consequences, there is a strong need to analyze the social implications and ethics of AI.

There are three main ethical concerns:

1. **Bias.** AI appears to be the solution to many societal issues: instead of having lawmakers and politicians, who may have conscious or unconscious biases, make policy decisions, we could have objective machines do so. Algorithms aren't subject to bias or emotions, but make decisions based on hard math. However, it’s easy for AI to become biased. Most training data comes from human-made examples that we want machines to emulate. If this data is biased, then the algorithms that are trained will be biased. For example, if bankers have consistently given minorities and women smaller loans, algorithms that are trained on this data will learn to be discriminatory as well. Bias may also arise from missing data and over/underrepresentation. For example, mood-recognition algorithms

detect one’s emotion from pictures of their face. If these are trained with mostly pictures of white people, they will have trouble with minorities’ faces.

2. **Black box.** Compounding the problem of bias is the fact that algorithms are difficult to interpret. A neural network can have dozens of layers, each with hundreds of cells. Analyzing the way this algorithm makes decisions involves looking at thousands of weights between each connection. This process is complicated and little meaning can be obviously discerned from these numbers. Most companies don't even make these algorithms open for inspection, citing privacy concerns and trade secrets, making them even harder to regulate. Even if AI appears to work, it is crucial to understand its mechanisms to ensure that they don't have hidden biases or side effects.
3. **What can AI decide?** Supposing that we have the fairest, most accurate AI that is better than any human. Do we let it decide whether someone on trial is guilty? Do we let it determine our war strategies? Some may say yes, but many people have an aversion to this idea. In any scenario, an exception may arise that is completely unlike any of the training data, throwing off the algorithm. Also, in situations of crime and punishment, of life and death, most people don't adopt a purely utilitarian calculus, and instead use innate, purely human forms of judgment. Whether this is ideal or not is up for debate, but it means that AI will not be single-handedly making these decisions anytime soon.

Each of these concerns prompts the need for regulation of AI. As this powerful force becomes increasingly prevalent, we need to be able to detect problems and find solutions.

Philosophy

Another topic that isn't the core of my research but that I found interesting is the ethical standing of AI. Machines haven’t reached the point of consciousness or emotion, but with the development of technology such as automated chat bots, it doesn't seem out of reach. With this innovation begs the question: can AI be living? Can we create a program that is completely digital but is still living, that has cognition and logic; in short, a true artificial intelligence?

Biologists may answer no, but many philosophers have debated the issue. A popular branch of philosophy is computationalism: in essence, the belief that our brains are essentially computers. ANNs themselves are modeled off of the neural networks in our brains. Are we just blobs of flesh, a few pounds of tissue with electrical impulses firing in just the right places at just the right time? Can we simulate, or even create life in a computer? Or is there some other life essence, a “soul,” if you will, that is required for consciousness?

Now, the project.

The Project

Background of Modeling it

For my project, I modeled the growth of the coronavirus in different US states given live data, and predicted its growth over the next several weeks.

I decided to do this project in late February-early March, when the coronavirus was beginning to spread beyond mainland China and to the United States. Throughout the course of this project, the situation has progressed from a two-week lockdown to indefinite school closings to state-wide stay-at-home orders around the US. As the pandemic devastates communities, scientists have provided widely varying estimates on the scope of the tragedy. Some believe that it will be over by the end of the calendar year, while others estimate that we'll be living with the virus for the next 3-5 years.

For my project, I decided to create my own model for the coronavirus spread using AI and analyze it through an ethical lens. While it will be nowhere near as robust as existing professional models, I hope to learn about AI concepts and pose questions that may be important to the field of statistical modeling at large.

The Model

I used the New York Times Github dataset, which is published for free online and updates the amount of cases and deaths in US counties every day. This allows me to pull from the live dataset and not have to manually update a file everyday. I coded in Python, which is a popular language for AI given its flexibility and many libraries.

The modeling process involved two main steps: finding an equation to model the growth and finding the parameters for the equation.

I browsed many forums and research papers on the coronavirus and pandemics. Most of the basic functions wouldn't work for clear reasons. A pandemic begins with rapid, almost exponential growth as infected people spread the virus to more and more people without detection or prevention. As prevention efforts increase and/or fewer people are left to be infected, the growth rate tapers off.

Linear, exponential, and logarithmic functions would be too restrained and wouldn't satisfy all of these conditions. Something like a logistic function would be closer, as it has rapid then slower growth:

$$f(x) = \frac{L}{1+Ae^{Bx}}$$

However, this function is also somewhat restrictive, as it must be horizontally symmetrical across its midpoint. In other words, the growth rate of a function exactly determines the growth rate at a future date.

After more research, I found that many people were recommending a function suggested in an earlier paper [4]. Interestingly enough, in addition to being written in French, it was published in 1995 by two economists. They were trying to model the spread of new innovations throughout a population:

$$f(x) = N * (1 - e^{-at})^{\alpha}$$

This has a similar growth to the logistic function, but allows for more flexibility and variability. Even though it wasn't designed for a pandemic, the model has many similarities. Technological growth is spread by word of mouth and contact, and eventually tapers out as everyone who may want the technology has received it.

I would use this model for both the confirmed cases count and death count in each US state.

Regression

After choosing the model, the next step is to determine the constants in the function. This basically entails running a regression and trying to find the parameters (the N, a, α values) that most closely model the spread with the least error. In terms of AI, it is finding the weights in each connection (though, in this case, they don't relate linearly).

Each state will have two sets of parameters — one for the confirmed cases count and one for the death count. There will be a total of 100 sets of parameters, each of which calculates the count of confirmed cases or deaths in one of the 50 states.

With the three parameters that relate exponentially, it appears difficult to determine them. However, there are many machine learning techniques that can be used. The one that I used was the Nelder-Mead method. [5]

The Nelder-Mead method uses the concept of a simplex, which is the shape that corresponds to a triangle in multiple dimensions. For example, in two dimensions, the simplex is a triangle, in three dimensions, a tetrahedron, and so on. The amount of dimensions corresponds to the number of parameters.

To explain, let’s attempt to model a linear growth, $g(x)$, using an equation with two parameters:

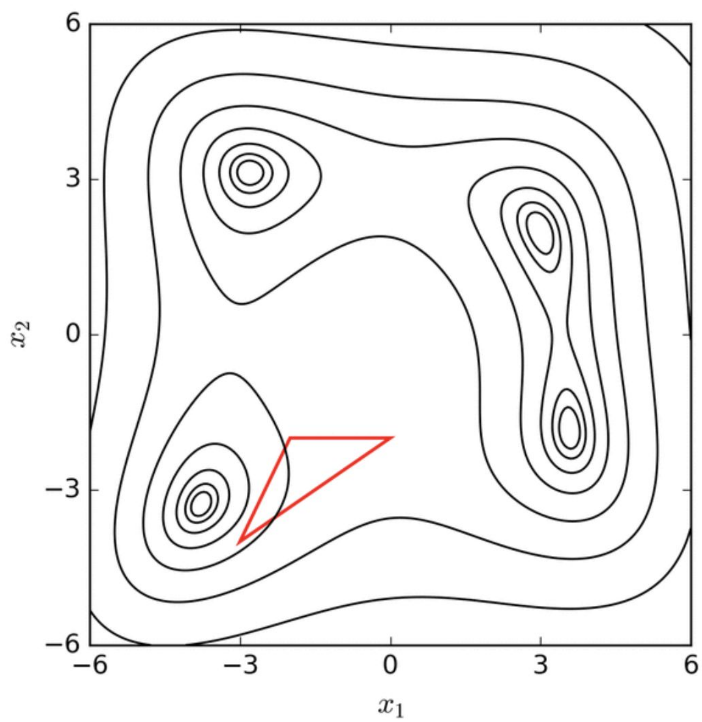
$$f(x) = ax + b$$

We are trying to solve for the equation that best fits $g(x)$, and are looking for the values of a and b in $f(x)$. To solve this, we use a two-dimensional coordinate plane. Each of the axes represents one of the parameters -- the horizontal axis may be the a value, and the vertical axis the b value. Every point on the grid, therefore, corresponds to a value of (a, b).

Each coordinate pair would also have a loss value. This is generally calculated by taking the sum of the squared residuals:

$$\sum_{a}^{b}(g(i) - f(i))^2$$

The closer $f(x) = ax + b$ matches the preexisting line, $g(x)$, the smaller this loss would be. If each point on the grid has a loss value, we can imagine the entire grid as a topographical map, where higher elevations correspond to the worst losses. Then, we are looking for the lowest elevation — the point with the least loss.



A Diagram of the Nelder-Mead Method [6]

This is where the simplex comes in. We begin with three points on the coordinate grid that form a triangle — three coordinate pairs of (a, b). Each of these three points has a loss value. We take the point with the worst loss value and reflect it across the midpoint of the other two points to create a new triangle. This is an attempt to push the points away from the peaks and towards the valleys — towards the points with less loss. The algorithm and procedure is more complicated and involves several steps, but this is the basic concept.

Python has a built in Nelder-Mead function, but in the spirit of this project, I coded it myself. For my algorithm, which uses three parameters, we need a tetrahedron in a three-dimensional space. Each of the points in this space has a loss value as well, and we use the same procedure to obtain the parameters.

I ran the algorithm to determine the parameters for both confirmed cases and deaths in all 50 states. Below are the parameter values I obtained for New Jersey.

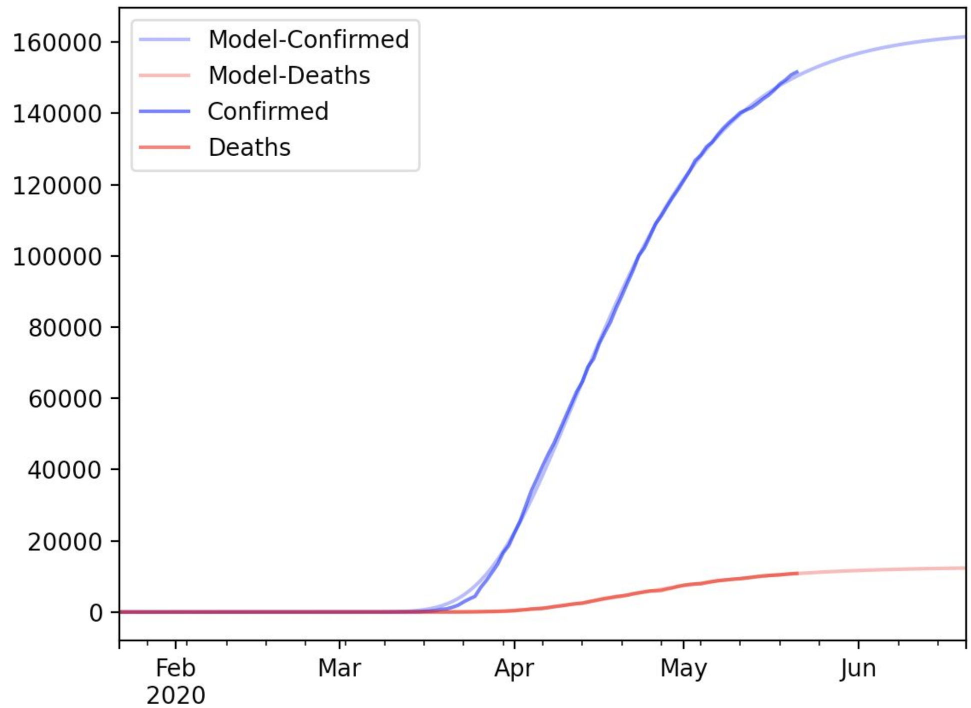
```
Confirmed Parameters from my function:
[163548.06615231058, 0.0632670383261013, 165.7238257272004]
Confirmed Parameters from python function:
[1.63548067e+05 6.32670385e-02 1.65723828e+02]
Death Parameters from my function:
[12678.143461294283, 0.06015833623791009, 219.67147708723854]
Death Parameters from python function:
[1.26781435e+04 6.01583360e-02 2.19671476e+02]
```

The parameters for N, a, and alpha, respectively, obtained. As you can see, the values obtained from my function were the same as those from Python's

Results

The model seemed to perform fairly well in all of the states it was tested in. Some of the states had sharp irregularities in the graphs that the model could not predict, but these are outliers and seem impossible to account for.

I also included the future predictions for each statistic. The values appear to tail off and increase a marginal amount each day, as expected.



Predictions and Results for New Jersey

Of course, there are several weaknesses that the model fails to account for. Besides the irregular spikes in some cases, there is also the possibility of a “second wave.” When social distancing measures are relaxed or the virus mutates, there will likely be a smaller surge in cases and deaths, possibly in the fall. The model is unable to account for this, as we aren't finished with the first wave, and it is impossible to predict the magnitude of a second wave without guessing.

Overall, though, given the limited information about the coronavirus spread, I’m satisfied with the strength of my model.

Ethics Analysis

The Suitability of the Model

As explained earlier, one potential flaw of AI is that it may be suited for certain situations, but not others. This specific model is based on human contact, which may work better in certain situations, such as urban settings.

To test this, I ranked the success of the model in each state, measured by the amount of error. To clarify, the states were not ranked by how contained the spread of the virus was, but by how well the model predicted the spread. Even states with horrible numbers of cases may have been predicted well. I calculated this by taking the total error of each final model and dividing it by the amount of cases to account for differences in population size.

Then, I found other lists of the 50 states in order of median income, urbanization, healthcare quality, and whiteness. I was curious to see if there would be a correlation between success of the model and these other factors. For this, I used the Mann-Whitney U Test, which tests rank correlation to see if two rankings are correlated or independent. The formula used for this is:

$$r = 1 - \frac{6\sum_{0}^{n}(a_i - b_i)^2}{(n^3 - n)}$$

where a_i and b_i are the placements of a state in each ranking, n is the number of states, and r is the correlation coefficient. The closer the placements are of the states in each ranking, the greater the r value will be.

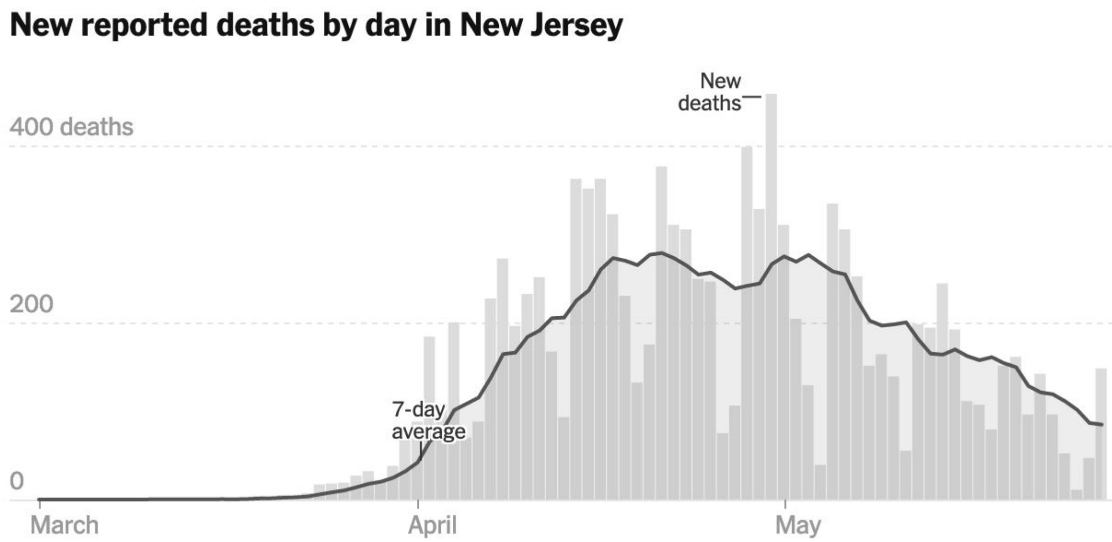
The correlation between Whiteness-Confirmed Success, Income-Deaths Success, Internet Quality-Deaths Success, and Urbanization-Deaths Success appeared to be the greatest. They had r-values of 0.328, 0.393, 0.335, and 0.413, respectively. However, using a critical values chart [7], these correlations are significant. They return P-values of $P<0.025$, $P<0.005$, $P<0.025$, and $P<0.01$, respectively. This means that if there was no relationship at all between the factors, there would be less than a 2.5% chance of obtaining these kinds of results by chance. In other words, there is convincing evidence that every one of these factors have a relationship with stronger predictions.

This suggests that the model is more accurate in states with higher incomes, greater urbanization, better healthcare, and fewer minorities. This has disturbing implications. More urban states may be more accurately modeled because they have denser populations. In other words, they have more predictable and set behavior, as opposed to rural, distanced states. States with higher income and better healthcare may also have more predictable treatment patterns, rather than erratic ones. Perhaps the most disturbing is the strong correlation with whiteness. This has two possibilities, neither of which is encouraging. First, the model may not be accurate at modeling the movement and transmission of the coronavirus in non-white groups — for example, many holidays, such as Eid Al-Fitr, have fallen during quarantine. Second, treatment or healthcare may not be as strong in states with greater minority populations, which serves to confirm the systematic inequalities in the healthcare system.

AI can help us expose biases in healthcare and social justice. However, we must be wary of the biases in AI itself.

Missing Data

Another concern raised is that missing data could lead to devastating consequences. I check the New Jersey times graphs of the coronavirus spread everyday. One day, I noticed something peculiar about the death counts.



NYT Death Counts in NJ [8]

While there is an obvious rise and decline in the general trend of cases, there also appears to be significant dips at regular intervals in the death count. When I looked closer, these appeared to be every Monday, more or less. This is unlikely to be by chance - maybe certain hospitals or counties don't release a new death count every Monday, so the amount of deaths is artificially deflated. Regardless, I was curious if this could affect my model.

I ran the algorithm using data that ran up through Monday, May 4th, and again, on data that ran up through Tuesday, May 5th. Given that this was only a one day difference out of the 60+ days that have already been recorded, ideally there shouldn't be too much of a difference in the predictions.

Instead, there were very different results. The predictions of confirmed cases stayed similar, but the model that was cut on a Monday predicted that there would be 2000 fewer deaths.

```
            Model-Confirmed  Model-Deaths
0
2020-05-26    150911.020908  12263.400603
2020-05-27    151425.589527  12382.886897
2020-05-28    151909.181967  12497.574856
2020-05-29    152363.568642  12607.606453
2020-05-30    152790.430403  12713.124497
2020-05-31    153191.361593  12814.271994
2020-06-01    153567.873197  12911.191580
2020-06-02    153921.396063  13004.025020
2020-06-03    154253.284148  13092.912757
2020-06-04    154564.817782  13177.993533
            Model-Confirmed  Model-Deaths
0
2020-05-26    151686.729103  10658.228554
2020-05-27    152215.744047  10718.883765
2020-05-28    152713.195041  10776.009055
2020-05-29    153180.866808  10829.791185
2020-05-30    153620.455226  10880.409350
2020-05-31    154033.570220  10928.035184
2020-06-01    154421.738779  10972.832804
2020-06-02    154786.408046  11014.958908
2020-06-03    155128.948462  11054.562903
2020-06-04    155450.656921  11091.787066
>>>
```

Prediction for the next 30 days after May 4th (top) and May 5th (bottom)

The Significance

These findings may seem like interesting easter eggs — quirks in the data that have no real impact. However, when we're talking about the coronavirus, these effects can be very real.

Algorithms are created with a certain logic -— an idea of how people interact, how quickly and successfully measures are adopted, etc. Our results clearly show that these algorithms are designed using the profile of the richest, whitest states — a profile that isn't universally applicable. The fact that models are less successful in less privileged states means that they will have less accurate information. They will be forced to guess on their day-to-day situation, which could have significant consequences. They may not have accurate knowledge on future levels of spread, which influences decisions such as when to reopen and when to tighten measures. By using the wrong model or even ending it on a different day of the week, states may prematurely shut down or open their economy, either of which would have drastic economic and health consequences.

Conclusion

Artificial intelligence is an incredibly useful tool that is already becoming prevalent in our everyday lives. As it grows more powerful, ethical issues are being raised about its possible flaws, which could have drastic real-world consequences. However, no matter the criticism, AI will not be going away anytime soon. Thus, it's crucial to understand the programs we are creating and find ways to regulate them to prevent bias and error. Some solutions may include finding completely different models for different situations to take into account the flexibility and different cultures. Others include simply monitoring data to prevent quirks and detect outliers.

Regardless of the situation, AI can contribute to social good. However, we must keep it in check to make sure it doesn't do more harm than good.

References

[1] 3blue1brown Youtube Video About Neural Networks:
https://www.youtube.com/watch?v=aircAruvnKk

[2] Diagram of a neuron from Webspace: http://webspace.ship.edu/cgboer/theneuron.html

[3] Gorman Analysis Diagram of ANN:
https://www.gormanalysis.com/blog/neural-networks-a-worked-example/

[4] French Paper on Spread of Innovation: https://www.jstor.org/stable/40588987?seq=1

[5] Nelder-Mead Method:
https://codesachin.wordpress.com/2016/01/16/nelder-mead-optimization/

[6] Diagram for Nelder-Mead:
https://chejunkie.com/knowledge-base/simplex-nelder-mead-optimization-amoeba-search/

[7] Mann-Whitney Critical Value Chart: https://math.usask.ca/~laverty/S245/Tables/wmw.pdf

[8] NYT Death Counts in NJ:
https://www.nytimes.com/interactive/2020/us/new-jersey-coronavirus-cases.html