Title
Innovation at Intersections: Enhancing Public Systems with AI (Professor Michael Jordan)

Journal
Berkeley Scientific Journal, 28(1)

ISSN
1097-0967

Authors
Potgieter, Lara
Palayur, Ann
Delaney, Andrew

Publication Date
2023

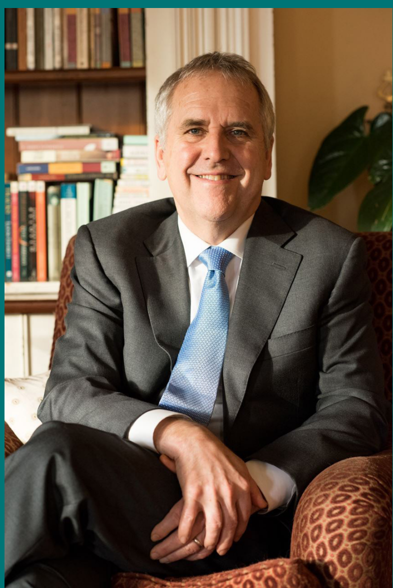
DOI
10.5070/BS328163626

Undergraduate

Innovation at Intersections: Enhancing Public Systems with AI

INTERVIEW WITH: PROFESSOR MICHAEL JORDAN

BY: LARA POTGIETER, ANN PALAYUR, & ANDREW DELANEY



Professor Michael Jordan teaches in the Department of Electrical Engineering and Computer Science and the Department of Statistics at UC Berkeley. He received the World Laureate's association prize in Computer Science or Mathematics in 2022 for his contributions to the understanding of machine learning and its applications. Jordan's research aims to connect the social sciences to computational, statistical, and cognitive science. Additionally in our interview we discussed the future of AI, from its ethics to its ability to uplift public systems across the globe.

BSJ: The research you are involved in ranges from economic and environmental to statistical and computing, where do you gain your inspiration to apply data analysis and computer science to such diverse fields?

MJ: Most academics' careers are a random walk. I started not knowing exactly what I wanted to do, and I believe that this is pretty common in the United States college system. The positive aspect of the U.S. college system is that it is open, and you can be very flexible. The slightly unfavorable side to this is that you are allowed to wander for a while and it is easy to get lost. I ended up doing a little bit of both. However, eventually, I became interested in psychology, philosophy, and neuroscience—all things connected with how the brain and mind work. So, I began my academic career more on the social and humanities side. This is what informs some breadth of my perspective. Although I learned about these different topics, a career in these fields was not a good fit for me due to my interest in more literal work. Eventually, I moved more into statistics, and from there, mathematics and its applications began to interest me.

By this point, I already had some sense in my mind about the applications of statistics and mathematics that I wanted to implement. I mainly focused on making human life easier, better, safer, etc. If there was something that I did not understand, I would work on it for a while, and begin to understand what things have an impact. I see my work as a back-and-forth pendulum of sorts, trying to do things that have a bigger and bigger impact. In my work, I would use the fundamentals to arrive at a new destination. All of this, in short, is about being open to new ideas, being curious about wanting to explore things, and building a career doing what I enjoy.

BSJ: AI, machine learning, and deep learning are large talking points in the current ever-evolving technological state. Many use these terms interchangeably, could you elaborate on the differences between them?

MJ: For outsiders to computer science, such as people in government or strategic levels in companies, I would tell them to think of them as the same. AI, machine learning, and deep learning are all composed of different networks that come together with data analysis to serve people. Whether it is a transportation system, commerce system, or a healthcare system, you can imagine that it has a plethora of data coming in, even if you are not that knowledgeable of the system. From experiments being done to services, all kinds of results are produced in the system. If you would like to make it better, you'd like to adapt it over time. In this regard, you do not want the system to "just be fixed." It should change if the data changes.

In the past, this has been done manually—people would gather a data set, they would analyze it, and would change some policies. Now with the advent of computers, the question arises: Why can this not be more automatic? Further still, why not have it be faster, automatic, and maybe even less subject to human biases?

This kind of exercise was termed "machine learning," as the machine was learning from data. Inside of it is the merging of statistics and computer science, just a blending of two traditions. In regards to artificial intelligence (AI), I see it, frankly, as a public relations buzzword that makes a lot of people excited. AI was an older term that had aspirations of putting thought into computers and making them intelligent. At the time, blended machine learning was having great success and was able to be useful in a variety of applications. In turn, it got called AI because, I think, this sounded more tantalizing and appealing to the public. However, as of now, the old AI idea of thought within a computer has not happened, it has been the machine learning

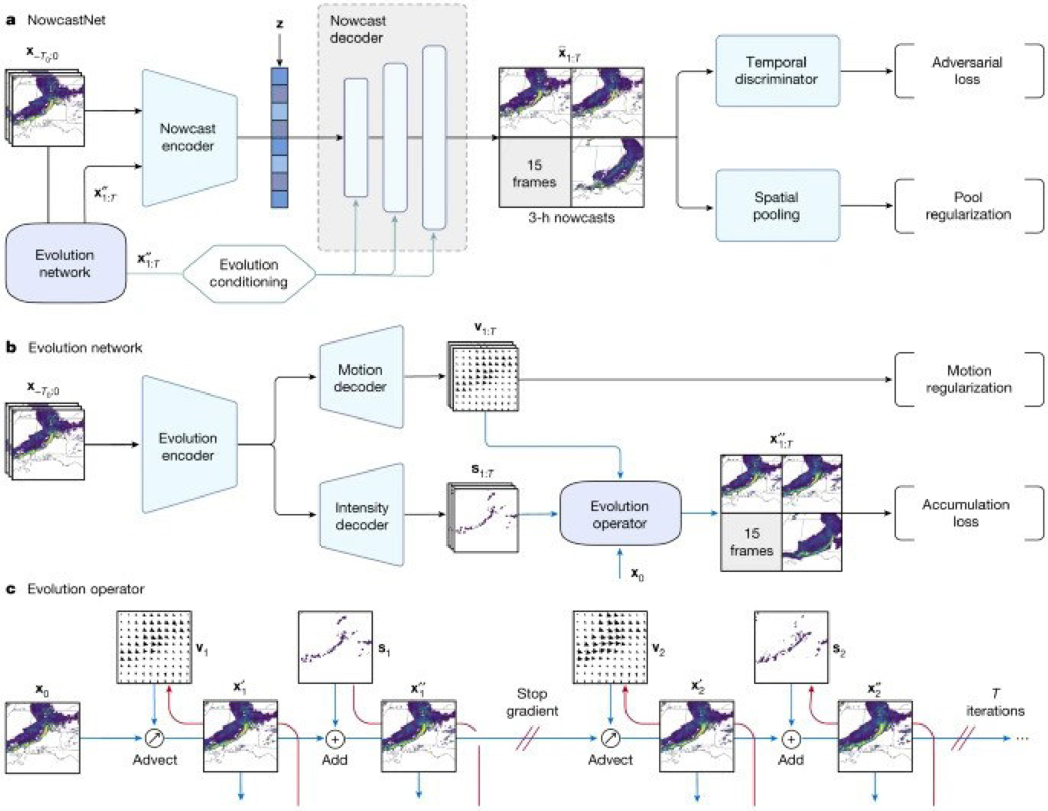


Figure 1: Architecture of NowcastNet. The nowcast encoder learns contextual representations of weather data, while the decoder conditions on the physics-informed evolutions. From here the nowcasts are generated and further separated and analyzed.

that has come to the fore.

Deep learning is a particular kind of machine learning, which operates within layered systems that are adaptive. If you make an error, it adjusts things a little bit. The "deep" just refers to those many layers. In some ways, it is just another buzzword for a particularly scalable, large, and successful part of machinery.

BSJ: In the NowcastNet project where deep-learning was combined with physical first principles, what are the advantages of implementing both of these structures simultaneously to a model? Were the findings of the NowcastNet project different than you initially expected?

MJ: Well, by using both deep learning and physical principles, it works better. Since the systems have to extrapolate, you want to distinguish interpolation and extrapolation. Let us say I have two observations of some state of nature, and with this, I am trying to make a prediction somewhere in between these two points. If it is close enough to the ones observed, and somewhere in between the observations, the system should be able to find their similarities and predict accordingly.

This process is where deep-learning and machine-learning systems excel. However, in many situations, such as limited data relative to a large system, you will nearly always have new phenomena that you have not seen before. Such phenomena will be outside of the training set, and you have to extrapolate. This extrapolation is, in some senses, dangerous and not something you want to execute carelessly. However, it is still something you want to be able to do, after all, the goal is to predict an outcome. Now to extrapolate, you can not just take the data points and put a smooth curve through them, you want to extend the range of outcomes using physical first principles. This is because physical principles like fluid flow, conservation of energy, and conservation of mass can be applied anywhere.

The deep-learning architecture used here is what is called generative AI nowadays. It takes in what is basically arbitrary and random white noise and can turn it into something new—in a similar manner to how image generators operate. If the generative AI can find enough structure in the 'white noise' input, then it will turn that into something potentially useful for analysis. It is a very natural thing to instead of using white noise in the generative system, to use noise that respects physical principles. This is really all that is done in the NowcastNet project paper—a very natural combination of physics with the deep learning system.

As for the results of the NowcastNet Project, it is natural to say that the results were different than what we expected. However, because weather is a very complicated domain you should almost expect that your predictions will not do well. With this in mind, our results were quite proficient, but at the same time there are cases where we want to improve upon and, due to this, we do not think the project is ready to be immediately deployed.

A vast number of people focus on weather prediction, and our goal is to go beyond state-of-the-art extreme weather prediction, by also identifying rainfall on a certain time and space scale. The aim is to be able to get smaller timescales, potentially within an hour or two and a few kilometers so that the data becomes actionable. Once these scales are reached, you want the system to work wherever you deploy it in the world, so we deploy with particular datasets based on particular conditions. This begs a question regarding extrapolation: Will this work

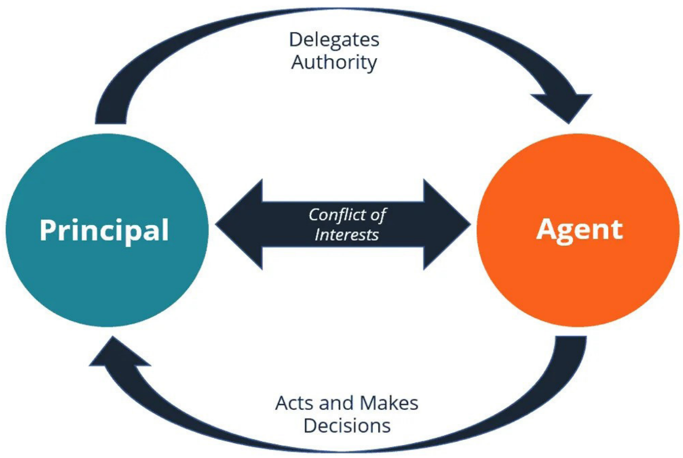


Figure 2: Principal Agent Relationship. The principal can be a person or entity which hires the agent to perform a desired task, while the agent is a person or entity that executes this. This relationship can exist in a plethora of instances, in this instance the regulator of pharmaceutical companies (the principal) and the pharmaceutical companies (the agent).

in India, or will this work in Africa? In short, the answer depends on the various local conditions. We thought NowcastNet was a good idea that would push the state of the art of weather forecasting, and I would say we did, but there is still much to be improved.

BSJ: Your paper introduces an interesting relationship between an agent (pharmaceutical company) and the principal (the regulator), where the agent wants to profit off of their drug and the principal only wants to approve drugs that benefit the public. In this context, how do e-values outperform p-values in informing the construction of contracts?

MJ: Most machine learning assumes you gather all of your data in one place, and then you make a big model of it, which is not true in many real-world situations. In reality, data is distributed in many places, and, even more importantly, the holder of the data does not want to just transfer it to the central site because they own the data. These are the economic reasons why data tends to be distributed.

As a response to this, we want machine learning systems that will respect this information asymmetry and respect the distinction between someone who is trying to do a task (the principal), and someone who will execute the task (the agent). So, we developed a way to do that, by using contract theory from economics and making it into a statistical contract theory where data is gathered and observations made.

Consider hypothesis testing, where you say "yes" or "no," but you do not want to make false-positive errors or false-negative errors. In a statistics class, p-values are used to quantify the strength of evidence against a null hypothesis. However, there are some difficulties in using p-values: it is hard to combine p-values, and it is hard to stop an experiment when a p-value is small because they turn out to be flawed ways of reasoning. Using e-values tends to fix those problems.

What we discovered is that if you bring in the economic side, e-values are actually "if and only if," meaning they are the only right way to do the contracts, and p-values are just not. So, there is a sharper criticism of p values for this task with contracts; where there are principles, agents, and incentives. I hope this research will help e-values to become more well-known in science. There are other reasons to use e-values, but whenever there is distributed data, they become even more highly recommended.

BSJ: To those untrained in machine-learning, it may seem like bad practice for conclusions to be drawn based on predictions by machine learning. So, how would you explain Prediction-Powered Inference to those weary about its accuracy?

MJ: Concretely, alpha fold was one of our motivations for the project. Alpha fold is the world's best predictor of protein structure. If given a new sequence of amino acids, for which a crystal structure is not known, alpha fold will predict a structure. And if you compare the predicted structure to lab-obtained ground truth, it is more accurate than any previous system. It is even more accurate than an individual human in many cases.

That begs the question: if alpha fold is so accurate, why not just use its predictions in place of data if you do not have enough data? That is not necessarily a terrible idea, but it is unsafe. To test this idea, we studied it in particular problems. In one instance, it was assessing whether there was an association between the intrinsic disorder of proteins (where there were strands forming, and not just folded structures), and protein activity (meaning that it is doing something in the cell). After this assessment, you do statistics with the data, such as doing a hypothesis test (ie. "is there an association or not?"), and you get a confidence interval on some statistic. If you do that using the output of alpha fold, you get a very narrow confidence interval, meaning you are very sure of yourself. We also got the lab-obtained ground truth of this problem, and the alpha fold turned out to be not at all close to covering the ground truth. This meant that we were very sure of ourselves, but we were very wrong.

We did this in many problems, including some problems in astronomy, some problems in ecology, and some problems in demography. Many times, the hypothesis testing confidence interval was very small (meaning we were very sure of ourselves), because you have a huge amount of pseudo data, but it did not cover the truth at all. As a result, you are making a very inaccurate decision. Why is this happening if the model is so accurate? The model is overall accurate, averaged over billions of instances. However, in some instances, it can be quite inaccurate. Those inaccurate instances might be the instance that matters for your hypothesis test, which was exactly what was happening here. This is not surprising in science, where you are often trying to push the envelope and test something that has not been looked at before. You are necessarily going to be doing things that are not in the training set. In the process, you are most likely going to make some uncontrolled errors. This is what motivated our problem, and we started thinking about how to solve it.

BSJ: Many people are scared about the future of AI. You have talked about how you hope to use AI to uplift communities. How do you believe we can best implement AI modeling to benefit and empower people around the world?

MJ: This is a big question. First of all, we have to understand what AI is and what it is not. Much of the dialogue around AI has nothing to do with what it actually is. Secondly, AI is being applied to collaborative systems where different groups jointly gather data to make conclusions–like health care, transportation, commerce, etc. AI's effect on these systems has obvious societal implications which should be discussed.

These questions are different from questions like: "can individual human beings be replaced by a computer?", and "do we have AI that is as smart as a human?" These questions are easier to think about, like a chess-playing robot replacing humans, or an AI surgeon that can perform as well as a human. However, it is not the right level of analysis. More important questions are: what if we took out a few surgeons and put in some robots? That would change things a bit. We would lose a few jobs, so we would have to think about it from a labor market point of view. But even then, you have to think about what the overall system is here. For example, we would have a healthier transportation system if things were more automated. The air traffic control system is an example of this; it got safer over time as it got more automated. An additional question to ask is, can we create new kinds of interactions, or new markets if we start to incorporate AI at the right time and the right level of analysis? I think we can.

For example, I talk about the application of AI in music. Many people are making music, and mainly 16- and 20-year-olds are making good music that people are willing to listen to, it gets streamed and it is listened to, but there is no market, meaning they do not make money. However, if so many people are listening to that music there should be a way to create a market. There are now companies whose job is to create a platform for producers, consumers, and brands. When actual value is created, if a certain musician is playing a large number of songs that a certain demographic listens to, and some brands are associated with that demographic, the brand can associate itself to the musician and have them write some songs for the brand. Thus, it becomes a three way market. Then, the musician could actually start to have a job as well. I work in an up-and-running company called UnitedMasters in the outside world, it has 3 million musicians and is really creating value in money for people. To me, that is what technology and AI can do. You can create new markets that are not just local, but large scale planetary wide. They can create more social value, better transportation, healthcare, new kinds of jobs, and so on. I do not think there is a panacea; AI creating new jobs does not mean that the problems of some jobs disappearing from the market are solved. It is just a broader perspective on what the overall goal of these things are, the goal is not to have super robots that replace humans. There is some truth to these fears, but they are very unlikely to happen in the near term. In a few 100 years, there may be a super robot that is more smart than any human. However, by the time we reach that stage, I think humans will know best how to control those kinds of super robots. However, these sorts of super fears are science fiction and not really worth thinking about for most of us. Unfortunately, these sorts of fears have taken over most journalists, who write solely with these fears in mind in regards to AI. Most journalists go directly to: "is this gonna kill all humans?" I actually think that the state of AI is moving ahead somewhat, and I am worried more about journalism. There is just a lack of conveying complex thoughts to people.

BSJ: Is there anything you are looking forward to, or is important, in your projects and future projects regarding AI?

MJ: I think this is a big international effort, not just a US effort. For this reason, I put all my work openly on the web. My students come from all over the world, they are a very diverse group of people. I travel, and I make sure my ideas are talked about by others around the world;I try to have an impact internationally. I think that that is not just true of me, but particularly true of me, perhaps. I think that it is critical that the way we are going to solve hard problems is not by creating little, separate entities that work on technology, but by bringing everyone's concerns and skills together–throughout the world.

IMAGE REFERENCES

1. 7 May 2019. PRINCIPAL AGENT RELATIONSHIP, https://studycafe.in/principal-agent-relationship-schedule-i-of-cgst-act-66904.html.
2. Bates, Stephen, et al. "Principal-Agent Hypothesis Testing." – arXiv Vanity, 24 May 2023, www.arxiv-vanity.com/papers/2205.06812/.
3. Zhang, Yuchen, et al. "Skilful nowcasting of extreme precipitation with NowcastNet." Nature, vol. 619, no. 7970, 2023, pp. 526–532, https://doi.org/10.1038/s41586-023-06184-4.