

# Artificial Intelligence and Machine Learning Based Decision-Making Tools: A Seminal Advancement for Neonatology?

Fu-Sheng Chou, MD, PhD, Monalisa Patel, MD

Last month, I had the pleasure of reading Dr. Gilbert Martin's article on the role of artificial intelligence (AI) in the NICU. (1) Numerous questions were raised in that article. I believe these are the exact questions that are on everyone's mind, including us.

Earlier this month, many of us attended the annual Pediatric Academic Society (PAS) conference. We were surprised yet delighted to see that there were more than a handful of presentations that used machine learning (ML) algorithms for analysis. A study compared linear and non-linear ML models for growth failure prediction and another one that used the Hidden Markov Model on serial microbiome data to predict growth failure. A multinational group is using unsupervised ML algorithms to redefine bronchopulmonary dysplasia (BPD). ML was also used to learn from MRI and clinical features to predict outcomes in newborns with hypoxic-ischemic encephalopathy. A study showed that ML-based models have greater sensitivity and specificity than the modified Bell criteria for predicting necrotizing enterocolitis. There was another study that compared various ML algorithms in critical congenital heart disease screening.

*“A study showed that ML-based models have greater sensitivity and specificity than the modified Bell criteria for predicting necrotizing enterocolitis.”*

Moreover, there was also a meta-analysis of studies that used various ML-based models for neonatal mortality prediction. The authors concluded that ML-based modeling is a feasible approach. Additionally, in a live Q&A session, one of the presenters for a California Perinatal Quality Care Collaborative study on active resuscitation of periviable infants hinted that the analysis is underway to use ML algorithms to predict outcomes of periviable infants based on maternal and perinatal clinical factors. Our apologies to the authors whose ML projects were not mentioned here.

The field of newborn medicine has seen the least progress in ML and Artificial Intelligence (AI) thus far. However, based on recent PAS presentations, it is likely that studies using ML or deep learning (DL) algorithms in our field will explode in the next few years. As we anticipate reading these articles very soon, it is probably a good time to learn some basic language about AI.

## AI and its Clinical Utility:

AI is math, computer science, and statistics in a nutshell. ML and DL are both components of AI. While human neonatologists learn how to manage transient tachypnea of newborns after seeing three cases, or five at most, machines need hundreds if not thousands of observations to delineate a pattern to correlate with the outcome. On the other hand, machines can gather all kinds of relevant or irrelevant information, sort through AI algorithms to identify hidden patterns that are not visible to human eyes, especially non-linear. Machines can pay attention to the speed of the car in front of you without getting tired, but humans can only do that for hours at most. Machines may not perform well in what humans can do well, such as abstract reasoning or creative problem solving or end-of-life discussions (based on current AI progress

in medicine). Machines can perform exceedingly well on specific tasks where human brains simply cannot, whether because of overwhelming amounts of information, mental exhaustion, or dogma passed down from generations ago that may not be applicable anymore.

Machines were initially created to assist humans in performing tasks. It is thought that there are three different strategies for how humans can utilize machines – the 3 A's: assisted, augmented, and autonomous. Assisted learning is when machines can perform basic repetitive tasks with no human involvement. An example of that is an alert system of the electronic record system (for example, prompting clinicians to consult an infectious disease specialist if a specific antibiotic is being ordered). On the other end of the spectrum, there is autonomous, where machines can complete the task at hand with little or no human involvement, including writing their own codes to instruct themselves. An example of that could be the FDA-approved device to screen for diabetic retinopathy without physician input. This autonomy is what human clinicians are scared of: I do not want a computer to tell me what to do; I do not trust the computer's decision; physicians will be out of the job market because of AI.

*“This autonomy is what human clinicians are scared of: I do not want a computer to tell me what to do; I do not trust the computer's decision; physicians will be out of the job market because of AI.”*

When we know more about AI and where we are with its development that is applicable to healthcare, immediately, we know that we are far from having fully autonomous AI to complete the “SOAP” process (well, there is no S in the eyes of AI). We likely are not going to have such a humanized machine for clinical decision-making in the next three decades or a century. One of the reasons is that five neonatologists sitting at the round table may have seven different ideas when discussing a complex case. If human neonatologists cannot agree with each other, how do we begin to instruct the machines to learn? AI will be implemented when humans do not perform well, but it does not take over human decision-making processes. The optimal way of using AI in medicine is to use it synergistically to take away a lot of repetitive and trainable work like documentation and billing to focus more time on direct patient care and communication.

Where we are with AI is well put by Dr. Anthony Chang of the Children's Hospital Orange County. Dr. Chang is a pediatric cardiologist specializing in cardiac intensive care. He is the founder of the medical intelligence society (2) and the editor-in-chief of Intelligence-Based Medicine (3), an Elsevier journal. Dr. Chang said he chose the title “Intelligence-Base” as opposed to “Artificial Intelligence-Base” because human intelligence is an integral part of the new era of medicine. What Dr. Chang meant precisely is how we should perceive AI in medicine, that machine intelligence should “augment” human intelligence to inform the best decision-making process. It should be symbiotic between human and machine, not additive, not in tandem.

Dr. Chang suggested that intelligence-based medicine will replace evidence-based medicine, which appears to be a positive health-care change. Evidence is at best based on findings humans can interpret, not necessarily the facts. For example, linear regression models and odds ratios are easy to interpret and understand, but not all variables are additive and linear, so when we say a particular risk factor is an independent one after adjusting for x, y, and z, is it an independent risk factor? All models are prone to bias; some are more useful than others, but are we learning the useful ones? Is that part of the reason we also say medicine is not pure science?

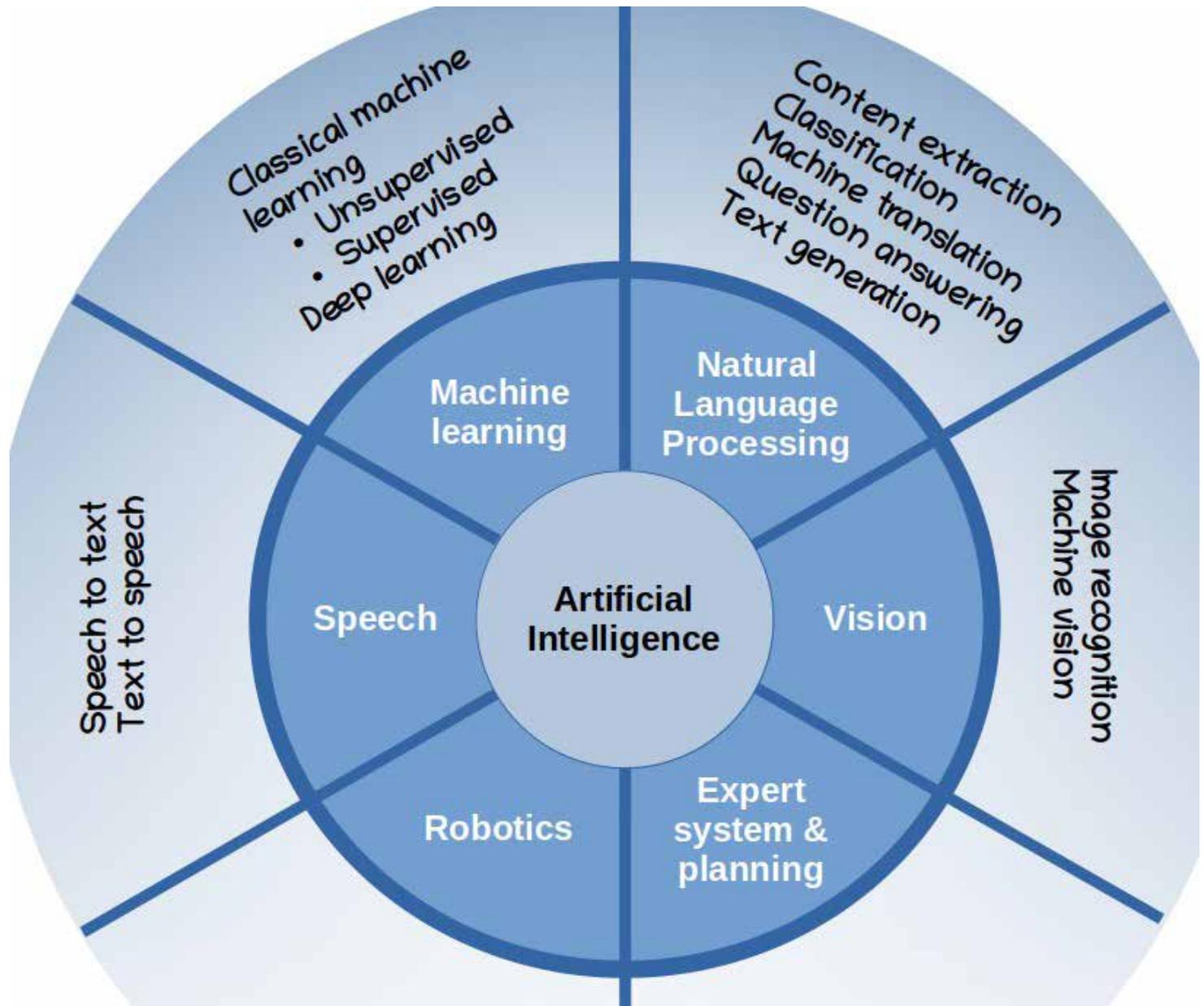
**Statistical Modeling vs. AI:**

AI is a branch of statistical modeling with an emphasis on generalizability. AI aims to learn from the present and the past to predict the outcome of future events. Both statistical and AI models begin with data. Statistical modeling is to use mathematical approaches to characterize and summarize existing data. On the other hand, AI models characterize existing data and use additional mathematical techniques to fine-tune how the model characterizes the data. Hence, it is “less” perfect but “better” generalizable. The fine-tuning part of model development is the spirit of the “learn-

ing” process by the machine, turning data into information and knowledge. A commonly used technique for model tuning is called k-fold cross-validation, where data is split into k parts, followed by training the model using the random (k-1) parts of the data and assessing model performance using a holdout fold of the data. This process can be repeated numerous times with a different holdout fold, and can also be nested, where n-fold cross-validation for hyperparameter tuning is performed within the (k-1) parts of data to find the optimal hyperparameters.

**Machine learning vs. deep learning:**

ML, or classic ML, takes a data set and runs through an algorithm to obtain predicted outcomes. The ML algorithms are categorized into linear and non-linear and based on the availability of outcome data, ML algorithms are also categorized into unsupervised (no outcome correlation) and supervised (with outcome correlation). A supervised linear model is like a linear regression model, where there are independent variables (called *features*) and dependent variables (called *outcome*). Additional techniques are built into the algorithm for post-processing to compress or eliminate features that are less likely to influence the generalized outcome. There are multiple approaches to non-linear modeling, which we will dis-



cuss in greater detail in the coming months. Unsupervised learning is complicated, and many algorithms are still under development. An easy way to visualize it mentally is to think about taking a bunch of cardiopulmonary data and allow the machine to cluster the data based on learned patterns. These clusters do not have any labels attached to them but may allow clinicians to redefine disease severity, BPD as an example. The powerfulness of unsupervised learning is it could revolutionize disease classification based on hidden pathophysiology or discover new phenotypes.

DL is a subset of ML. The structure of DL is to mimic how the human brain functions: environmental input (data input) travels through multiple interconnected neurons (nodes) arranged hierarchically (artificial neural network) to produce a nerve impulse to instruct behavior (outcome prediction). Nodes within a layer are interconnected, and each node contains a linear regression algorithm that receives input from upstream nodes and provides output to downstream nodes. Multiple layers were developed by human design to decode and process the information for pattern development. Deep learning is used for image recognition (convoluted neural networks, or CNN) or temporal data recognition such as murmur detection or continuous vital sign monitoring (recursive neural networks, or RNN). Developing a model to recognize pneumatosis on abdominal x-ray would require CNN; developing a model to recognize flow directions of a patent ductus arteriosus would probably require both CNN and RNN and a lot of computing power.

---

***“The powerfulness of unsupervised learning is it could revolutionize disease classification based on hidden pathophysiology or discover new phenotypes.”***

---

In general, data for supervised ML are less complex and smaller in scale when compared to DL. DL is powerful but also consumes significant computing power. Choosing between ML and DL and which specific algorithms for training largely depend on the clinical question asked.

#### **Cognitive Computing:**

Cognitive computing deploys AI methodologies (like reinforcement learning, deep learning, natural language processing, machine learning, neural networks, sentiment analysis, and contextual awareness) to simulate a human-like cognition characterized by self-learning behavior with intelligence. This strategy is going to be the future wave of AI.

#### **Limitations:**

Algorithms are very prone to bias. After all, most of the learning so far is still supervised, which requires human labeling of data during the training process. An infamous example of this is the major flaw in the Google algorithm that could not differentiate an African American from a gorilla. Another big issue with AI is the black box. In particular, with DL systems, the lack of explainability of how the AI program can reach a particular conclusion or make a particular recommendation can be daunting for clinicians and perhaps patients as well. Many AI systems lack interpretability, and again, this will be a hindrance to the clinician buy-in.

Moreover, it takes enormous amounts of data, training, programming, and funding to create a well-trained, reliable algorithm that can produce results that will augment human performance. And

as medicine advances, newer testing and treatment options are discovered, the algorithms will need to be constantly updated, which can be a gargantuan task. Lastly, abstract reasoning, creative problem solving, and complex decision-making remain limitations of AI as we know it.

#### **Summary:**

Similar to statistical inference and the famous line in statistics, “all models are wrong, some are useful,” AI models are being developed to augment human decision; some of them will become useful in clinical settings, others will be abandoned quickly as prospective validation fails to prove its value. We will discuss bias vs. variance, model performance, and what to pay attention to when reviewing manuscripts that report AI-based models in the coming months.

#### **References:**

1. *Martin GI. Will Artificial Intelligence Have a Place in the NICU? Are we there yet or should we be? Neonatology Today [Internet]. 2021 [cited 2021 May 6];16(4):53–5. Available from: [https://neonatologytoday.org/Archive/html/NeonatologyToday\\_2021\\_4\\_16\\_4\\_53\\_55.html](https://neonatologytoday.org/Archive/html/NeonatologyToday_2021_4_16_4_53_55.html)*
2. *Medical Intelligence Society [Internet]. [cited 2021 May 6]. Available from: <https://www.misociety.org/>*
3. *Chang AC. Intelligence-based medicine: Artificial intelligence and human cognition in clinical medicine and healthcare [Internet]. San Diego, CA: Academic Press; 2020 [cited 2021 May 7]. 550 p. Available from: <https://www.journals.elsevier.com/intelligence-based-medicine/>*

*Disclosure: The author identifies no conflict of interest*

**NT**

*Corresponding Author*



*Fu-Sheng Chou, MD, PhD -  
Senior Associate Editor,  
Director, Digital Enterprise  
Neonatology Today  
Assistant Professor of Pediatrics  
Division of Neonatology, Department of Pediatrics  
Loma Linda University Children's Hospital  
Email: [FChou@llu.edu](mailto:FChou@llu.edu)*



*Monalisa Patel, MD  
Division of Neonatology,  
Department of Pediatrics,  
Children's Hospital Orange County  
Orange, CA  
Email: [dr.monalisa.patel08@gmail.com](mailto:dr.monalisa.patel08@gmail.com)*