



HHS Public Access

Author manuscript

Anesthesiology. Author manuscript; available in PMC 2020 December 01.

Published in final edited form as:

Anesthesiology. 2019 December ; 131(6): 1346–1359. doi:10.1097/ALN.0000000000002694.

Artificial Intelligence and Machine Learning in Anesthesiology

Christopher W Connor, MD PhD

Assistant Professor, Department of Anesthesiology, Perioperative and Pain Medicine, Brigham and Women's Hospital, Boston, MA, USA

Research Associate Professor, Department of Physiology and Biophysics, Boston University, Boston, MA, USA

Abstract

Commercial applications of artificial intelligence and machine learning have made remarkable progress recently, particularly in areas such as image recognition, natural speech processing, language translation, textual analysis and self-learning. Progress had historically languished in these areas, such that these skills had come to seem ineffably bound to intelligence. However, these commercial advances have performed best at single task applications in which imperfect outputs and occasional frank errors can be tolerated. The practice of anesthesiology is different. It embodies a requirement for high reliability, and a pressured cycle of interpretation, physical action and response rather than any single cognitive act.

This review covers the basics of what is meant by artificial intelligence and machine learning for the practicing anesthesiologist, describing how decision-making behaviors can emerge from simple equations. Relevant clinical questions are introduced to illustrate how machine learning might help solve them – perhaps bringing anesthesiology into an era of machine-assisted discovery.

Summary Statement

Anesthesiologists synthesize data from disparate sources, of varying precision and prognostic value, making life-critical decisions under time pressure. This review describes the evolution of artificial intelligence and machine learning through application to this challenging environment.

Introduction

The human mind excels at estimating the motion and interaction of objects in the physical world, at inferring cause and effect from a limited number of examples, and at extrapolating those examples to determine plans of action to cover previously unencountered circumstances. This ability to reason is backed by an extraordinary memory that

Address: Dr. Christopher W Connor, Department of Anesthesiology, Perioperative and Pain Medicine, 75 Francis Street, CWN-L1, Boston, MA 02115, cconnor@bwh.harvard.edu.

Conflicts:

US Patent 8460215 Systems and methods for predicting potentially difficult intubation of a subject

US Patent 9113776 Systems and methods for secure portable patient monitoring

US Patent 9549283 Systems and methods for determining the presence of a person

subconsciously sifts events into those experiences that are pertinent and those that are not, and is also capable of persisting those memories even in the face of significant physical damage. The *associative* nature of memory means that the aspects of past experiences that are most pertinent to the present circumstance can be almost effortlessly recalled to conscious thought. However, set against these remarkable cerebral talents are fatigability, a cognitive laziness that presents as a tendency to short-cut mental work, and a detailed short-term working memory that is tiny in scope. The human mind is slow and error-prone at performing even straightforward arithmetic or logical reasoning¹.

In contrast, an unremarkable desktop computer in 2019 can rapidly retrieve and process data from 32 gigabytes of internal memory – a quarter of a trillion discrete bits of information – with absolute fidelity and tirelessness, given an appropriately constructed program to execute. The greatest progress in artificial intelligence (AI) has historically been in those realms that can most easily be represented by the manipulation of logic and that can be rigorously defined and structured, known as classical or symbolic AI. Such problems are quite unlike the vagaries of the interactions of objects in the physical world. Computers are not good at coming to decisions – indeed, the formal definition of the modern computer arose from the proof that certain propositions are logically undecidable² – and classical approaches to AI do not easily capture the idea of a “good enough” solution.

For most of human history, the practice of medicine has been predominantly heuristic and anecdotal. Traditionally, quantitative patient data would be relatively sparse, decision making would be based on clinical impression, and outcomes would be difficult to relate with much certainty to the quality of the decisions made. The transition to evidence-based practice³ and Big Data is a relatively recent occurrence. In contrast, anesthesiologists have long relied on personalized streams of quantified data to care for their unconscious patients, and advances in monitoring and the richness of that data have underpinned the dramatic improvements in patient safety in the specialty⁴. Anesthesiologists also practice at the sharper end of cause-and-effect: decisions usually cannot be postponed, and errors in judgment are often promptly and starkly apparent.

The general question of artificial intelligence and machine learning in anesthesiology can be stated as follows:

1. There is some outcome that should be either attained or avoided.
2. It is not certain what factors lead to that outcome, or a clinical test that predicts that outcome cannot be designed.
3. Nevertheless, a body of patient data is available that provides at least circumstantial evidence as to whether the outcome will occur. The data is plausibly, but not definitively, related.
4. The signal, if it is present in the patient data, is too diffuse across the dataset for it to be learned reliably from the number of cases that an anesthesiologist might personally encounter, or the clinical decision-making relies upon a subconscious judgment that the anesthesiologist cannot elucidate.

5. Can an algorithm, derived from the given data and outcomes, provide insight in order to improve patient management and the decision-making process?

This form of machine learning might be termed *machine-assisted discovery*.

This article takes the form of an integrative review⁵, defined as “a review method that summarizes past empirical or theoretical literature to provide a more comprehensive understanding of a particular phenomenon or healthcare problem.” The article therefore introduces the theory underlying classical and modern approaches to artificial intelligence and machine learning, and surveys current empirical and clinical areas to which these techniques are being applied. Concepts in the fundamentals of the artificial intelligence and machine learning are introduced incrementally:

1. Beginning with classical/symbolic AI, a logical representation of the problem is crafted and then searched for an optimal solution.
2. Model fitting of physiological parameters to an established physiological model is shown as an extension of search.
3. Augmented linear regression is shown to allow certain non-linear relationships between outcomes and physiological variables to be discerned, even in the absence of a defined physiological model. It requires sufficient expertise about which combinations of non-physiological transformations of the variables might be informative.
4. Neural networks provide are shown to provide a mechanism to establish a relationship between input variables and an output without defining a logical representation of the problem or defining transformations of the inputs in advance. However, this flexibility comes at considerable computational cost and a final model whose behavior may be hard to comprehend.

Numerous other theoretical and computational approaches do exist, and these may have practical advantages depending on the nature of the problem and the structure of the desired outputs.⁶

The literature search for an integrative review should be transparent and reproducible, comprehensive but focused and purposive. A literature search was performed using PubMed for articles published since 2000 using the following terms: “artificial intelligence anesthesiology” (543 matches), “computerized analysis anesthesiology” (353 matches), “machine learning anesthesiology” (91 matches) and “convolutional neural network anesthesiology” (1 match). Matches were reviewed for suitability, and augmented with references of historical significance. The specialty of anesthesiology features a broad history of attempts to apply computational algorithms, artificial intelligence and machine learning to tasks in an attempt to improve patient safety and anesthesia outcomes (Table 1). Recent significant and informative empirical advances are reviewed more closely.

Classical Artificial Intelligence and Searching

Creating a classical artificial intelligence algorithm begins with the three concepts of a bounded solution space, an efficient search, and termination criteria.

Firstly, using what is known about the problem, a set of possible solutions that the algorithm can produce is defined. The algorithm will be allowed to choose one of these possible solutions, and so the solution set must be created in such a way that it is reasonably certain that the best possible solution is among the choices available. The algorithm will never be able to think outside of this “box”, and in that sense the solution space is bounded. In the game of tic-tac-toe, for example, the set of solutions is those squares that have not yet been taken. The best solution is the one that most diminishes the opponent’s ability to win, ideally until victory is achieved (i.e. *minimax*)⁷. In real life problems, however, it can be difficult to define a bounded set of solutions or even say explicitly what “best” means.

Secondly, the possible solutions are progressively evaluated and searched, trying to find the best one. In designing and programming the search strategy, anything else of worth that is known about the problem should be incorporated, such as how to value one solution versus another, ways to search efficiently by focusing on areas of the solution space that are more likely to be productive⁸, and intermediate results that might allow certain subsets of the solution space to be excluded from further evaluation (i.e. *pruning*). Sometimes the knowledge and understanding of the underlying problem might be quite weak, and then in the worst case it may be necessary to fall back on an exhaustive and computationally-intensive brute-force search of all the possible solutions.

Thirdly, the algorithm must terminate and present a result. Given enough time, eventually the algorithm should ideally find and select the optimum solution. Depending on the structure of the problem and the search algorithm, it may be possible to guarantee through theory that the algorithm will terminate with the optimal solution within a constrained amount of time. A weaker theoretical guarantee would be that the algorithm will at least improve its solution with each search iteration. However, in the general case and if no such theoretical guarantee is possible, the algorithm might only select the best good-enough solution found within an allowed time limit, or perhaps issue an error message that no sufficiently satisfactory solutions were identified.

Search-based classical AI has obvious applications to practical problems such as wayfinding on road maps, in which a route must be chosen that is connected by legal driving maneuvers and arrives in the shortest time. Less obviously, this same logic can be applied to real-world problems such as locating a lost child in a supermarket. According to the order of operations above, the first step is to create a bounded solution set: by covering the exit, the location of the child is reasonably bounded to be somewhere within the supermarket. Secondly, a search is begun. A naive approach might be to walk up and down every aisle in turn until the child is found but, from insight, far better search strategies for this problem can be easily identified. The most efficient search strategy is clearly to walk along the ends of the aisles: this allows whole aisles to be scanned and excluded (i.e. *pruned*) rapidly. Thirdly, the search

terminates either on finding the child, or on determining that additional resources must be employed if the child cannot be found within a certain time.

Designing classical AI algorithms is not a turnkey mathematical task; it is heavily dependent on the human expertise of the designer. In classical AI, the role of the computer is to contribute its immense power of calculation to evaluate the relative merits of a large number of possible solutions, which the designer provides. This division of labor can be dated back to Ada Lovelace's 1843 description of the conception of the modern computer: "The Analytical Engine has no pretensions whatever to *originate* anything. It can do whatever we *know how to order it* to perform. It can *follow* analysis; but it has no power of *anticipating* any analytical relations or truths. Its province is to assist us in making *available* what we are already acquainted with."⁹

In 1997, the IBM supercomputer Deep Blue defeated the then world champion, Garry Kasparov, at chess. It had been a longstanding goal for a machine to be able to play chess at levels unattainable by humans¹⁰. The rules of chess are clear and unambiguous, and the actions take place within the confines of the board. The state of play is completely apparent and known to both players. It is possible to list all the available legal moves, all responses to those moves, all responses to those responses, and so on – the solution space is bounded. In principle, it is not even particularly difficult to write a program that can play chess flawlessly. The program simply tries out (i.e. *searches*) all possible moves, and all possible responses until the game is either won or lost. However, a computer program that tried to evaluate every possible move and all of its consequences would not be able to make its first move, so immense is the search space.¹¹ Deep Blue's success rested on two pillars. Firstly, its search algorithm possessed an evaluation function to approximate the relative value of a position. This function was crafted from the distilled, programmed, strategic wisdom of human chess experts, and allowed the search algorithm to ignore choices that were likely to be unproductive. Secondly, this search algorithm was supported by brute-force computing power capable of evaluating two hundred million moves per second. These techniques proved sufficient for Deep Blue to achieve superhuman mastery of a game with approximately 10^{47} possible board positions - an immense but bounded space. In many ways, however, mastery of chess was classical AI's triumph but also its swansong. The division of labor remains the same as in Lovelace's original description, and the human strategic understanding of the game was not outdone but instead overwhelmed by the indefatigability of the machine's tactical evaluation of millions and millions of positions. The computer did what it was told, but it did not learn.

In anesthesiology practice, the closest example is open-loop Target Controlled Infusion (TCI). Pharmacokinetic (PK) models describe the forwards relationship from a drug administration schedule $D(t)$ to an effect site concentration $\epsilon(t)$. However, it is the inverse solution that is required: for a requested $\epsilon(t)$, some $D(t)$ should be produced, perhaps subject to limits on administration rate or plasma concentration¹². An open-loop TCI pump will perform a search for a drug administration schedule that brings the predicted concentration of the medication within the body towards this goal, subject to the given constraints. The underlying equations are concise and effective¹³, but the device cannot become more proficient at its task. It follows the algorithms that it is given.

Model Fitting as a Form of Searching

Anesthesiologists take particular interest in objective patient outcomes, and whether these good or bad outcomes can be predicted from the data that are available. Lacking a direct test for the desired outcome (in which case prognosis would be straightforward), the research question becomes whether the patient's outcome is in some way imprinted upon and foreshadowed by the imperfectly informative data that are available. An approach is to seek to fit models to the data in order to try to make more reliable predictions, and therefore potentially discover previously unappreciated but useful relationships within the data. This approach requires a large enough body of data and patient outcomes on which model fitting can be performed, and this large body of data cannot reasonably be analyzed by hand.

A model is created using an example set for which the data and outcomes are known – i.e. the *training data*. The essence of a useful model is that it should be able to make useful predictions about data it has not previously seen, i.e. that it is *generalizable*. An overly complicated model may become over-fit to its training data, such that its predictions are not generalizable. Figure 1 shows examples of this. Each figure shows a population of green circles, representing notionally favorable outcomes, and red crosses, representing notionally unfavorable outcomes. The question is whether the two available items of data, Feature X and Feature Y, can predict the outcome. Three models are fit to the exact same training data, to produce a black line (known as the *discriminant*) that separates the figure into prediction regions, shaded either green or red, accordingly. An item of training data is correctly *classified* when its symbol falls in a region that is shaded the same color, and is *misclassified* when it does not (i.e. when it lies on the wrong side of the discriminant).

Figure 1A shows a model that is *underfit*. Although most of the symbols are correctly classified, there are several misclassified red crosses on the left of the figure and the simple linear discriminant has no way to capture these. A decision algorithm based on this model would have high specificity (green circles are almost all correctly classified), but a lower sensitivity (several bad outcomes are erroneously predicted as good). The decision performance is therefore somewhat reminiscent of the Mallampati test¹⁴ which also demonstrates high specificity but low sensitivity¹⁵. The discriminant in Figure 1A would function better if it could assume a more complex form. Figure 1B, in contrast, shows a model that is overfit. Although all the training data is correctly classified, the unwieldy discriminant is governed too much by the satisfaction of individual data points rather than the overall structure of the problem. This model is unlikely to generalize well to new data, as it is overly elaborate. Figure 1C shows a model that is appropriately fit to the data (indeed, the data were created to illustrate this point). The discriminant is complex enough to capture the distribution of the outcomes, but it is also *parsimonious* in that the shape of the discriminant is described by only a few parameters. In practice, of course, the true underlying distribution is not known in advance, so the performance of the discriminant must be tested statistically. The discriminant in Figure 1C has fewer *degrees of freedom* than the discriminant in Figure 1B, so its performance is statistically more likely to represent the true nature of the underlying process even though it has more misclassifications than the overfit discriminant. Model fitting is therefore a form of search in which the choices are the parameters admitted to the model and their relative weights, in order to find models that are

statistically most likely to represent the underlying process based upon the training data that are available.

Discovering Non-Linear Relationships in Clinical Medicine

Many judgments in anesthesiology are based on absolute thresholds or linear combinations of variables. A patient with a heart rate above 100 is tachycardic, one with a temperature above 38°C is febrile. A man whose EKG features an R wave in aVL and an S wave in V₃ that combined exceed 28 mm has left-ventricular hypertrophy¹⁶.

Logistic regression is useful when fitting a weighted combination of variables to an outcome. Logistic regression defines an *error function* that measures the extent to which the current weighted combination of variables tends to misclassify outcomes. These weights are subsequently modified in order to improve the classification rate. The regression algorithm determines the changes in the weightings that would most improve the present classification, and then repeats the process until an optimum weighting is settled upon. The regression algorithm therefore performs a *gradient descent* on the error function; one can picture an imaginary ball rolling down a landscape defined by the error function until a lowest, optimum, point is reached such that the best linear combination of the input variables is determined. Logistic regression is a powerful machine learning technique that works quickly and is also *convex*, meaning essentially that the “ball” can roll down to the optimum solution from any starting point (for almost all well-posed problems). However, many outcome problems in anesthesiology and critical care clearly do not depend on linear criteria. For example, ICU outcomes that may depend on a patient’s potassium level, or glucose level, or airway PEEP are more likely to be *Goldilocks problems*: the best outcomes require an amount that is neither too big, nor too small, but just right. Figure 2A illustrates such a situation, in which the good outcomes are clustered around a point in the feature space, and deviations from that point result in poor outcomes. As a clinical correlate, one might imagine that the outcomes are timely ICU discharges¹⁷ and the data K and G represent well-controlled levels of potassium and glucose, although the data shown here are purely artificial and created solely for this example. The data show a clear clustering of the outcomes, but an algorithm that is only capable of producing a discriminant based on a linear combination of K and G would not be able to capture that separation. Rather than performing a non-linear regression over the two variables K and G, a solution lies in transforming the data by calculating the squares of K and G (i.e. K², G²) and their cross-term KG, and then performing an augmented linear logistic regression over the five variables K, G, K², G² and KG.

As shown in Figure 2B, a linear discriminant in the K², G² plane will perfectly separate the outcomes. This discriminant, given by $(K^2) + (G^2) - 9 = 0$, is exactly the same as a circle of radius 3 in Figure 2A, demonstrating that non-linear boundaries can be discovered. Although it may seem clinically bizarre to talk about the squared value of the serum potassium (K²), it is easy to write a *quadratic* function that has clinical meaning. For example, the function $35 - 17K - 2K^2$ is positive if the value of K lies between 3.5 and 5.0, but turns negative for any more hypokalemic or hyperkalemic value outside that range. This simple example underscores the ways in which the outputs from machine learning algorithms can seem

inscrutable or *black box*. To a computer, the two definitions are essentially equivalent: one is no more meaningful or better than the other. It takes human effort to explain numerical results in a clinically meaningful way¹⁸.

Augmenting the variable space with quadratic terms allows a linear algorithm to define non-linear features like islands (as in Figure 2A) and open curves (as in Figure 1B), but the technique can be extended further by using higher polynomial terms. Augmentation can also be performed with reciprocal powers such as K^{-1} (i.e. $1/K$), which would in principle allow the machine-learning algorithm to discern useful relationships based on ratios. Common such clinical examples include:

- the Shock Index¹⁹ ($SI = HR / \text{Systolic BP}$), which rises in response to the combined increase in heart rate and fall in blood pressure associated with hypovolemia.
- the Rapid Shallow Breathing Index²⁰ ($RSBI = f / V_T$) which rises in response to the fast, small, panting breaths associated with respiratory failure.
- the Body Mass Index ($BMI = \text{Weight} / \text{Height}^2$), which represents obesity as excess weight distributed over an insufficiently-sized physical frame.

A downside to augmenting the variable space is that the number of input variables can increase dramatically, which can overwhelm the size of the available training data and lead to a significant risk of overfitting. One challenge is that the input variables and the augmented combinations that are to be considered must be fully defined in advance. Only non-linear relationships that can be approximated from a linear combination of the variables that are supplied can be found. For real-world problems in medicine and biology, considerable expertise is required in order to define a meaningful and informative set of inputs. Human insight must also be applied to determine what problem should be solved and what outputs are useful. When only limited knowledge is available about the best way to frame a problem numerically, modern AI and neural networks provide an alternative approach.

Modern AI

The limitations of classical AI were particularly apparent in attempts to produce programs capable of playing Go. Go is, at least in terms of its rules, a simpler game than chess. Two players take it in turns to place a stone, white or black respectively, on a 19×19 board. Plays in Go take place at the intersections of the grid lines, rather than on the squares as in chess. Once a stone is placed, it does not subsequently move. Briefly, the game is won by whoever manages to corral the largest total space on the board. However, in play, Go is a much more complex game than chess, with approximately 10^{170} board positions compared to 10^{47} . It is hard to overstate the magnitude of these combinatoric numbers. There are about 10^{80} atoms in the universe, so if each individual atom were actually another universe in its own right, then that would still represent only a total of 10^{160} atoms. Even the best classical AI approaches to Go seemed unable to accomplish anything better than amateurish play.

In March 2016, a computer program, AlphaGo, defeated a 9-dan rated human player, the world number-two Lee Sedol, in a head-to-head five-game series of Go. This was the first time that a computer program had beaten a player of that level of skill without handicaps. Although AlphaGo featured an algorithm to choose moves that was somewhat guided by the design of its developers, its evaluation function was composed of a neural network that had been trained against a database of recorded games and outcomes²¹. In chess, a crude valuation of a position can be made from the strength of the pieces remaining to the player and their ability to move freely. In Go, the stones do not have an equivalent individual worth and the valuation of a Go position instead depends on the relative spatial interplay of the player's stones and the opponent's stones. Where classical AI algorithms were unable to discern this strategic posture, AlphaGo's neural network approach was successful.

Neural Networks

Figure 3A illustrates the simplest feasible fully-connected feed-forward neural network, taking two inputs and returning one outcome. The network is composed of the inputs, an input layer, a hidden layer, an output layer, and the output. Each layer is fully connected to the next, meaning there is path from each *node* (i.e. *neuron*) to every node in the following layer. Each path has an associated *weight*, which describes how much the signal traveling along that path is amplified or attenuated or inverted. At each node, the weighted inputs are added together and then applied to an activation function. Each of the nodes illustrated here uses a sigmoid activation function, which is the most basic of the standard activation functions as shown in Figure 3B. The general idea is that a node, in a manner reminiscent of a biological neuron, will remain 'off' until a suitable degree of excitation is reached, at which point it will quickly turn 'on'. The first node in the input layer, for example, receives inputs from Features S and T. These inputs are weighted by $w_{fs,il}$ and $w_{ft,il}$ respectively, so the total input z to the first input node is given by $z = w_{fs,il}S + w_{ft,il}T$. The total input is then applied to the sigmoid function, producing an output from this node of $\frac{1}{(1 + e^{-z})}$. This output feeds forward to the next, hidden layer along with the other weighted contributions from the input layer, and so forth until an output is produced. The output of the sigmoid activation function is always between 0 and 1, so if the outcomes are classified as 0 (e.g. red crosses) and 1 (e.g. green circles), the performance of the network can be assessed from how closely it predicts the various outcomes in the training data.

The behavior of the network depends on the values of the various weights w , and so the general idea of machine learning in a neural network is to adjust these weights until satisfactory performance is achieved. To begin, the weights are set to random values and so the initial performance of the network will usually be poor. However, for each error in prediction that is output, a degree of blame can be apportioned over the weights that contributed to it, and these weights can then be adjusted accordingly. This process is called *back propagation*²², and it is the process by which the network learns to improve from its mistakes. Data is fed forward through the weights and nodes to produce output predictions, and then the errors in these predictions are propagated backwards through the network to readjust the weights. This process is continued until, hopefully, the network settles to some form in which it is able to model the outputs satisfactorily based upon the input data.

Beyond this basic description, of course, there are extraordinary implementational details and subtleties. For example, even in the rudimentary neural network shown in Figure 3A, there are already ten different weights that can be adjusted. The number of parameters in any practical network will be very large, and a great deal of care is required in the handling of the training data in order to avoid immediate overfitting. Additionally, the error function for neural networks is not globally convex, so there is no guarantee that the learning process will converge upon the optimum solution and it may instead settle on some less ideal solution. In the gradient descent analogy described earlier, this would be like the imaginary ball becoming stuck in a small divot and failing to roll down to the valley below. Two ways around this problem are either to survey the landscape by starting from a selection of different locations, or to occasionally give the ball (or the landscape) some sort of shake (i.e. *stochastic gradient descent*²³). Nevertheless, the process remains very computationally intensive and slow, despite technical advances in repurposing the hardware of 3D graphics cards (i.e. *GPUs*) to parallelize the calculations²⁴.

The primary reason to take on the burden of training neural networks is that they possess the new property of *universality*²⁵. Universality means that, given an adequately large number of nodes in the respective layers, the weights of a neural network can be configured to approximate any other continuous function to within any desired level of accuracy²⁵. This leads to two immediate and important benefits.

1. The property of universality stipulates that the neural network can, in principle, represent any continuous function to any desired degree of accuracy. The idea of a *function* is very broad – it does not just mean the transformation of one numerical value into another. It incorporates any transformation of input data into an output, such as a Go board position into a verdict into whether that position is winning or losing²¹, or determining the location of a lesion on a 3D brain MRI²⁶. A function can be any transformation, even if its mathematical form is not known in advance.
2. The behavior of the network is dependent on the weights. The network learns the appropriate weights solely from the training data that it is given. Therefore, the network can learn the functional relationship between the outcome and the data *even if there is no pre-existing knowledge about what that relationship might be*. However, it can be extraordinarily difficult to reverse this process to determine an efficient statement of the functional relationship that is described by the fitted weights. This leads to the well-known criticism that the operation of a neural network is particularly hard to characterize and therefore hard to validate.

As shown in Figure 4 therefore, it is possible, at least in the abstract, for a neural network to take a pre-operative image of a patient and produce a prediction about how difficult that patient's intubation might be. The proposed function is a transformation from the pixel values of the image to an estimated Cormack-Lehane view²⁷, but it is hard to intuit in advance what the form of that underlying function might turn out to be. While the picture alone is very unlikely to contain sufficient information to produce a reliable prediction, it is plausible that it is in some way informative as to the outcome. Although a fully-connected neural network is shown, the universality theorem only demonstrates that a fully-connected

network with a single hidden layer can represent any function. It does not guarantee that the network contains a reasonably tractable number of nodes, nor that the inputs are informative as to the output, nor that convergence to a satisfactory answer can occur within a feasible amount of time. The current state of the art in computer science therefore involves finding *network topologies* that use a more efficient number of nodes and can be trained in a reasonable period of time. Two examples of these alternative network connection patterns are deep convolutional neural networks²⁸, in which there are several hidden layers but many weights are constrained to have the same set of values, and residual neural networks²⁹, in which additional paths with a weight of 1 skip over intervening hidden layers. Both of these approaches derive plausible justification from analogous arrangements of neurons in the mammalian brain, such as visual field maps for convolutional networks and pyramidal projection neurons for residual networks.

Evolution in Time

Each of the feed-forward neural network tasks illustrated so far make their predictions based solely on the input data available at that immediate time. They are *stateless*, i.e. they have no temporal relationship to any measurement taken before or after. If a neural network is intended to make intraoperative decisions about patient management, then the network will require some way to base its decision-making on memories of evolving trends. Figure 5 illustrates an Elman network³⁰, with three inputs and one output. In this topology, weighted paths project from the outputs of the hidden layer to a context layer, and then further weighted paths return from the context layer to the inputs of the hidden layer. An Elman network is the simplest example of a *recurrent* neural network (RNN) that can evaluate changes in data over time. For example, the output might be a decision whether to transfuse or not³¹, and the inputs might be the clinically observable parameters heart rate (R), blood pressure (P) and estimated blood loss (B). The context layer would allow the network to discern and respond to trends in these inputs.

Practical Approaches to Machine Learning in Anesthesiology

Advances in technology and monitoring can change the impetus for machine learning. For example, a neural network developed to detect esophageal intubation from flow-loop parameters³² is obviated by continuous capnography^{33,34}. In this instance, a reliable clinical test has made readily apparent what was once an insidious and devastating complication. A machine-learning model to predict difficult intubation from patient appearance³⁵ must now be tempered by the convenience and ubiquity of video laryngoscopy. Advances in airway management technology have broadened the range of outcomes of laryngeal visualization that can be accepted. Anesthesiologists have long considered the possibility of an algorithm that might autonomously control depth of anesthesia based on EEG recordings^{36,37} since the 1950s – yet, this concept remains very much a topic of current research³⁸.

Two papers from 2018 illustrate the theoretical concepts covered. The first paper by Hatib *et al*³⁹ uses a very highly augmented data set in conjunction with logistic regression to produce an algorithmic model that can, in *post hoc* analysis, detect the incipient onset of hypotension up to 15 minutes before hypotension actually occurs. For model training, the authors

employed a database of 545,959 min of high-fidelity (100Hz) arterial waveform recordings acquired from the records of 1,334 patients, internally validated against the records of 350 additional patients that were held back. The training dataset included 25,461 episodes of hypotension. The model itself is derived from 51 base variables obtained from the processing of arterial waveforms by the Edwards FloTrac device⁴⁰. Each variable was augmented with its squared term and their reciprocals (i.e. X , X^2 , X^{-1} , X^{-2}) and then every combination of these variables was generated to produce an overall input set of 2,603,125 parameters. The authors chose two clearly separated outcomes: hypotension defined as MAP < 65 mmHg (e.g. notionally red crosses), and non-hypotension defined as MAP > 75 mmHg (e.g. notionally green circles), but did not consider the “gray zone” between these outcomes. Despite the large number of available parameters and the risk of overfitting, the authors were nevertheless able to use a parsimonious parameter selection process to produce a final model that depended on only 23 of the 2.6 million available inputs (M. Cannesson, *personal communication*). The study did have some limitations: notably that it did not include any episodes in which hypotension was caused by surgical intervention, all model-fitting and assessments were retrospective and off-line, and the algorithm made no recommendations as to whether an intervention should be performed. Nevertheless, the authors demonstrated an algorithm that was apparently able to foresee episodes of hypotension in operative and ICU patients up to 15 minutes in advance of the onset of the event itself with an area-under-the-curve (AUC) of 0.95.

The second paper by Lee *et al*⁴¹ describes a neural network approach to predicting the bispectral index (BIS) based upon the infusion history of propofol and remifentanyl. This paper is particularly noteworthy because a strongly theoretical approach to this question already exists in the TCI literature. The classical approach is to model the pharmacokinetics of propofol⁴² and remifentanyl⁴³ in the body independently, based upon the infusion history. The effect site concentration of each drug is then combined in a response surface model⁴⁴, producing an estimate of the BIS. These classical pharmacokinetic models are well-established and have been used to demonstrate closed-loop TCI control of anesthetized patients⁴⁵. In contrast, Lee *et al* created a neural network comprised of two stages. The first stage receives the infusion history of propofol and remifentanyl over the preceding 30 minutes with a resolution of 10 seconds (i.e. 180 inputs for each medication). The inputs for each medication are fed to 2 separate 8-node RNNs. Rather than using an Elman³⁰ arrangement, as seen in Figure 5, the paper made use of a newer configuration known as a Long Short Term Memory (LSTM)⁴⁶. Simple RNNs such as Elman have difficulty recalling or learning events that happen over a long timeframe as their training error gradients become too small to be adaptive. The LSTM is a more robust memory topology that also includes pathways that explicitly cause the network to reinforce or forget remembered states. The output from the LSTM layer is applied directly to a simple fully-connected feed-forward neural network with 16 nodes of the type shown in Figure 2 and Figure 3. A single output node emits a scaled BIS estimation. The network was developed from a database of 231 patient cases (101 cases used for training, 30 for validation, and 100 for final testing), and comprised a total of around 2 million data points. In *post hoc* analysis, the classical PK/PD models were able to predict the BIS value with a root-mean-square error (RMSE) of 15 over all phases of the anesthetic. Despite being naïve to all existing theory, the neural network

comfortably outperformed the best current models with an RMSE of 9 – a remarkable victory for modern AI over existing classical PK/PD expert systems⁴⁷ that might lead us to question the ongoing utility of classical response surface models⁴⁸.

Future Directions

The most exciting recent advance in machine learning has been the development of AlphaGo Zero⁴⁹, a system capable of learning how to play board games without any human guidance, solely through self-play alone. It performs at a level superior to all previous algorithms and human players in chess, go and shogi. This learning approach requires that the system be able to play several lifetimes worth of simulated games against itself. Though anesthesia simulators exist, they do not presently simulate patient physiology with the fidelity with which a simulated chess game matches a real chess game.

The most plausible route to the introduction of AI and machine learning into anesthetic practice is that the routine intraoperative management of patients will begin to be handed off to closed-loop control algorithms. Maintaining a stable anesthetic is a good first application because the algorithms do not necessarily have to be able to render diagnoses, but rather to detect if the patient has begun to drift outside the control parameters that have been set by the anesthesiologist⁵⁰. In this regard, such systems would be like an autopilot, maintaining control but alarming and disconnecting if conditions outside the expected performance envelope were encountered – hardly a threat to the clinical autonomy of the anesthesiologist⁵¹. A closed-loop control system need not necessarily have any learning capability itself, but it provides the means to collect a large amount of physiological data from many patients with high fidelity and this is an essential precursor for machine learning. Access to large volumes of high quality data will enable more machine learning successes, such as the off-line *post hoc* prediction of bispectral index⁴¹ and hypotension³⁹ discussed above. For now, finding algorithms that provide good clinical predictions in real time should be emphasized. Management of all the parameters of a stable anesthetic is not a simple problem⁵², but embedding⁵³ the machine in the care of the patient is a good way to begin⁵⁴.

For further reading, the following books provide accessible introductions to decision making by humans¹ and algorithms⁵⁵, neural networks⁵⁶, information theory⁵⁷ and the early history of computer science⁵⁸.

Acknowledgments

Funding: NIH R01 GM121457, Departmental

References

1. Kahneman D: Thinking, fast and slow, 1st pbk edition. New York, Farrar, Straus and Giroux, 2013
2. Turing AM: On computable numbers, with an application to the Entscheidungsproblem. Proceedings of the London mathematical society 1937; 2: 230–265
3. Evidence-Based Medicine Working G: Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992; 268: 2420–5 [PubMed: 1404801]

4. Eichhorn JH, Cooper JB, Cullen DJ, Maier WR, Philip JH, Seeman RG: Standards for patient monitoring during anesthesia at Harvard Medical School. *JAMA* 1986; 256: 1017–20 [PubMed: 3735628]
5. Whittemore R, Knafl K: The integrative review: updated methodology. *J Adv Nurs* 2005; 52: 546–53 [PubMed: 16268861]
6. Kendale S, Kulkarni P, Rosenberg AD, Wang J: Supervised Machine Learning Predictive Analytics for Prediction of Postinduction Hypotension. *Anesthesiology* 2018
7. Maschler M, Solan E, Zamir S, Hellman Z, Borns M: *Game theory*, 2013
8. Hart PE, Nilsson NJ, Raphael B: A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics* 1968; 4: 100–107
9. Lovelace A: ‘Notes on L. Menabrea’s ‘Sketch of the Analytical Engine Invented by Charles Babbage, Esq.’’. *Taylor’s Scientific Memoirs* 1843; 3: 1843
10. Shannon CE: XXII. Programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1950; 41: 256–275
11. Newell A, Shaw JC, Simon HA: Chess-playing programs and the problem of complexity. *IBM Journal of Research and Development* 1958; 2: 320–335
12. Van Poucke GE, Bravo LJ, Shafer SL: Target controlled infusions: targeting the effect site while limiting peak plasma concentration. *IEEE Trans Biomed Eng* 2004; 51: 1869–75 [PubMed: 15536889]
13. Shafer SL, Gregg KM: Algorithms to rapidly achieve and maintain stable drug concentrations at the site of drug effect with a computer-controlled infusion pump. *J Pharmacokinet Biopharm* 1992; 20: 147–69 [PubMed: 1629794]
14. Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiburger D, Liu PL: A clinical sign to predict difficult tracheal intubation: a prospective study. *Can Anaesth Soc J* 1985; 32: 429–34 [PubMed: 4027773]
15. Shiga T, Wajima Z, Inoue T, Sakamoto A: Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. *Anesthesiology* 2005; 103: 429–37 [PubMed: 16052126]
16. Carey MG, Pelter MM: Cornell voltage criteria. *Am J Crit Care* 2008; 17: 273–4 [PubMed: 18450685]
17. Desautels T, Das R, Calvert J, Trivedi M, Summers C, Wales DJ, Ercole A: Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: a cross-sectional machine learning approach. *BMJ Open* 2017; 7: e017199
18. Connor CW, Segal S: The importance of subjective facial appearance on the ability of anesthesiologists to predict difficult intubation. *Anesth Analg* 2014; 118: 419–27 [PubMed: 24361845]
19. Birkhahn RH, Gaeta TJ, Terry D, Bove JJ, Tloczkowski J: Shock index in diagnosing early acute hypovolemia. *Am J Emerg Med* 2005; 23: 323–6 [PubMed: 15915406]
20. Yang KL, Tobin MJ: A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation. *N Engl J Med* 1991; 324: 1445–50 [PubMed: 2023603]
21. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M: Mastering the game of Go with deep neural networks and tree search. *nature* 2016; 529: 484 [PubMed: 26819042]
22. LeCun YA, Bottou L, Orr GB, Müller K-R: *Efficient backprop*. *Neural networks: Tricks of the trade*, Springer, 2012, pp 9–48
23. Mei S, Montanari A, Nguyen PM: A mean field view of the landscape of two-layer neural networks. *Proc Natl Acad Sci U S A* 2018; 115: E7665–E7671 [PubMed: 30054315]
24. Peker M, Sen B, Guruler H: Rapid automated classification of anesthetic depth levels using GPU based parallelization of neural networks. *J Med Syst* 2015; 39: 18 [PubMed: 25650073]
25. Cybenko G: Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems* 1989; 2: 303–314

26. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 2017; 36: 61–78 [PubMed: 27865153]
27. Cormack RS, Lehane J: Difficult tracheal intubation in obstetrics. *Anaesthesia* 1984; 39: 1105–11 [PubMed: 6507827]
28. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR: Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016; 316: 2402–2410 [PubMed: 27898976]
29. He K, Zhang X, Ren S, Sun J: Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp 770–778
30. Elman JL: Finding structure in time. *Cognitive science* 1990; 14: 179–211
31. Etchells TA, Harrison MJ: Orthogonal search-based rule extraction for modelling the decision to transfuse. *Anaesthesia* 2006; 61: 335–8 [PubMed: 16548951]
32. Leon MA, Rasanen J: Neural network-based detection of esophageal intubation in anesthetized patients. *J Clin Monit* 1996; 12: 165–9 [PubMed: 8823638]
33. American Society of Anesthesiologists: Standards for Basic Anesthetic Monitoring. Park Ridge, IL American Society of Anesthesiologists, 1986.
34. Cheney FW: The American Society of Anesthesiologists Closed Claims Project: what have we learned, how has it affected practice, and how will it affect practice in the future? *Anesthesiology* 1999; 91: 552–6 [PubMed: 10443619]
35. Connor CW, Segal S: Accurate classification of difficult intubation by computerized facial analysis. *Anesth Analg* 2011; 112: 84–93 [PubMed: 21081769]
36. Bickford RG: Use of frequency discrimination in the automatic electroencephalographic control of anesthesia (servo-anesthesia). *Electroencephalogr Clin Neurophysiol* 1951; 3: 83–6 [PubMed: 14822927]
37. Bellville JW, Attura GM: Servo control of general anesthesia. *Science* 1957; 126: 827–30 [PubMed: 13467283]
38. Pasin L, Nardelli P, Pintaudi M, Greco M, Zambon M, Cabrini L, Zangrillo A: Closed-Loop Delivery Systems Versus Manually Controlled Administration of Total IV Anesthesia: A Meta-analysis of Randomized Clinical Trials. *Anesth Analg* 2017; 124: 456–464 [PubMed: 28099320]
39. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, Rinehart J, Cannesson M: Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. *Anesthesiology* 2018; 129: 663–674 [PubMed: 29894315]
40. Pratt B, Roteliuk L, Hatib F, Frazier J, Wallen RD: Calculating arterial pressure-based cardiac output using a novel measurement and analysis method. *Biomed Instrum Technol* 2007; 41: 403–11 [PubMed: 17992808]
41. Lee HC, Ryu HG, Chung EJ, Jung CW: Prediction of Bispectral Index during Target-controlled Infusion of Propofol and Remifentanyl: A Deep Learning Approach. *Anesthesiology* 2018; 128: 492–501 [PubMed: 28953500]
42. Schnider TW, Minto CF, Shafer SL, Gambus PL, Andresen C, Goodale DB, Youngs EJ: The influence of age on propofol pharmacodynamics. *Anesthesiology* 1999; 90: 1502–16 [PubMed: 10360845]
43. Minto CF, Schnider TW, Egan TD, Youngs E, Lemmens HJ, Gambus PL, Billard V, Hoke JF, Moore KH, Hermann DJ, Muir KT, Mandema JW, Shafer SL: Influence of age and gender on the pharmacokinetics and pharmacodynamics of remifentanyl. I. Model development. *Anesthesiology* 1997; 86: 10–23 [PubMed: 9009935]
44. Bouillon TW, Bruhn J, Radulescu L, Andresen C, Shafer TJ, Cohane C, Shafer SL: Pharmacodynamic interaction between propofol and remifentanyl regarding hypnosis, tolerance of laryngoscopy, bispectral index, and electroencephalographic approximate entropy. *Anesthesiology* 2004; 100: 1353–72 [PubMed: 15166553]
45. Liu N, Chazot T, Hamada S, Landais A, Boichut N, Dussaussoy C, Trillat B, Beydon L, Samain E, Sessler DI, Fischler M: Closed-loop coadministration of propofol and remifentanyl guided by

- bispectral index: a randomized multicenter study. *Anesth Analg* 2011; 112: 546–57 [PubMed: 21233500]
46. Hochreiter S, Schmidhuber J: Long short-term memory. *Neural Comput* 1997; 9: 1735–80 [PubMed: 9377276]
 47. Gambus P, Shafer SL: Artificial Intelligence for Everyone. *Anesthesiology* 2018; 128: 431–433 [PubMed: 29166324]
 48. Liou JY, Tsou MY, Ting CK: Response surface models in the field of anesthesia: A crash course. *Acta Anaesthesiol Taiwan* 2015; 53: 139–45 [PubMed: 26321504]
 49. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, van den Driessche G, Graepel T, Hassabis D: Mastering the game of Go without human knowledge. *Nature* 2017; 550: 354–359 [PubMed: 29052630]
 50. Harrison MJ, Connor CW: Statistics-based alarms from sequential physiological measurements. *Anaesthesia* 2007; 62: 1015–23 [PubMed: 17845653]
 51. Cannesson M, Rinehart J: Closed-loop systems and automation in the era of patients safety and perioperative medicine. *J Clin Monit Comput* 2014; 28: 1–3 [PubMed: 24306322]
 52. Yang P, Dumont G, Ford S, Ansermino JM: Multivariate analysis in clinical monitoring: detection of intraoperative hemorrhage and light anesthesia. *Conf Proc IEEE Eng Med Biol Soc* 2007; 2007: 6216–9 [PubMed: 18003441]
 53. Brooks RA: Elephants don't play chess. *Robotics and autonomous systems* 1990; 6: 3–15
 54. Alexander JC, Joshi GP: Anesthesiology, automation, and artificial intelligence. *Proc (Bayl Univ Med Cent)* 2018; 31: 117–119 [PubMed: 29686578]
 55. Koerner TW: Naive decision making : mathematics applied to the social world. Cambridge, Cambridge University Press, 2008
 56. Rashid T: Make your own neural network : a gentle journey through the mathematics of neural networks, and making your own using the Python computer language, CreateSpace, Charleston, SC, 2016
 57. MacKay DJC: Information theory, inference, and learning algorithms. Cambridge, Cambridge University Press, 2004
 58. Padua S: The Thrilling Adventures of Lovelace and Babbage. New York City, NY, USA, Pantheon Books, 2015
 59. Sobrie O, Lazouni MEA, Mahmoudi S, Mousseau V, Pirlot M: A new decision support model for preanesthetic evaluation. *Comput Methods Programs Biomed* 2016; 133: 183–193 [PubMed: 27393809]
 60. Zhang L, Fabbri D, Lasko TA, Ehrenfeld JM, Wanderer JP: A System for Automated Determination of Perioperative Patient Acuity. *J Med Syst* 2018; 42: 123 [PubMed: 29846806]
 61. Ortolani O, Conti A, Di Filippo A, Adembri C, Moraldi E, Evangelisti A, Maggini M, Roberts SJ: EEG signal processing in anaesthesia. Use of a neural network technique for monitoring depth of anaesthesia. *Br J Anaesth* 2002; 88: 644–8 [PubMed: 12067000]
 62. Ranta SO, Hynynen M, Rasanen J: Application of artificial neural networks as an indicator of awareness with recall during general anaesthesia. *J Clin Monit Comput* 2002; 17: 53–60 [PubMed: 12102250]
 63. Jeleazcov C, Egnér S, Bremer F, Schwilden H: Automated EEG preprocessing during anaesthesia: new aspects using artificial neural networks. *Biomed Tech (Berl)* 2004; 49: 125–31 [PubMed: 15212197]
 64. Lalitha V, Eswaran C: Automated detection of anesthetic depth levels using chaotic features with artificial neural networks. *J Med Syst* 2007; 31: 445–52 [PubMed: 18041276]
 65. Shalhaf R, Behnam H, Sleigh JW, Steyn-Ross A, Voss LJ: Monitoring the depth of anesthesia using entropy features and an artificial neural network. *J Neurosci Methods* 2013; 218: 17–24 [PubMed: 23567809]
 66. Lin CS, Li YC, Mok MS, Wu CC, Chiu HW, Lin YH: Neural network modeling to predict the hypnotic effect of propofol bolus induction. *Proc AMIA Symp* 2002: 450–3 [PubMed: 12463864]

67. Laffey JG, Tobin E, Boylan JF, McShane AJ: Assessment of a simple artificial neural network for predicting residual neuromuscular block. *Br J Anaesth* 2003; 90: 48–52 [PubMed: 12488378]
68. Santanen OA, Svartling N, Haasio J, Paloheimo MP: Neural nets and prediction of the recovery rate from neuromuscular block. *Eur J Anaesthesiol* 2003; 20: 87–92 [PubMed: 12622489]
69. Nunes CS, Mahfouf M, Linkens DA, Peacock JE: Modelling and multivariable control in anaesthesia using neural-fuzzy paradigms. Part I. Classification of depth of anaesthesia and development of a patient model. *Artif Intell Med* 2005; 35: 195–206 [PubMed: 16019196]
70. Nunes CS, Amorim P: A neuro-fuzzy approach for predicting hemodynamic responses during anesthesia. *Conf Proc IEEE Eng Med Biol Soc* 2008; 2008: 5814–7 [PubMed: 19164039]
71. Moore BL, Quasny TM, Doufas AG: Reinforcement learning versus proportional-integral-derivative control of hypnosis in a simulated intraoperative patient. *Anesth Analg* 2011; 112: 350–9 [PubMed: 21156973]
72. Moore BL, Doufas AG, Pyeatt LD: Reinforcement learning: a novel method for optimal control of propofol-induced hypnosis. *Anesth Analg* 2011; 112: 360–7 [PubMed: 21156984]
73. El-Nagar AM, El-Bardini M: Interval type-2 fuzzy neural network controller for a multivariable anesthesia system based on a hardware-in-the-loop simulation. *Artif Intell Med* 2014; 61: 1–10 [PubMed: 24703775]
74. Lin CS, Chiu JS, Hsieh MH, Mok MS, Li YC, Chiu HW: Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks. *Comput Methods Programs Biomed* 2008; 92: 193–7 [PubMed: 18760495]
75. Lin CS, Chang CC, Chiu JS, Lee YW, Lin JA, Mok MS, Chiu HW, Li YC: Application of an artificial neural network to predict postinduction hypotension during general anesthesia. *Med Decis Making* 2011; 31: 308–14 [PubMed: 20876347]
76. Reljin N, Zimmer G, Malyuta Y, Shelley K, Mendelson Y, Blehar DJ, Darling CE, Chon KH: Using support vector machines on photoplethysmographic signals to discriminate between hypovolemia and euvoolemia. *PLoS One* 2018; 13: e0195087 [PubMed: 29596477]
77. Aleksic M, Luebke T, Heckenkamp J, Gawenda M, Reichert V, Brunkwall J: Implementation of an artificial neuronal network to predict shunt necessity in carotid surgery. *Ann Vasc Surg* 2008; 22: 635–42 [PubMed: 18761224]
78. Peng SY, Peng SK: Predicting adverse outcomes of cardiac surgery with the application of artificial neural networks. *Anaesthesia* 2008; 63: 705–13 [PubMed: 18582255]
79. Thottakkara P, Ozrazgat-Baslanti T, Hupf BB, Rashidi P, Pardalos P, Momcilovic P, Bihorac A: Application of Machine Learning Techniques to High-Dimensional Clinical Data to Forecast Postoperative Complications. *PLoS One* 2016; 11: e0155705 [PubMed: 27232332]
80. Fritz BA, Chen Y, Murray-Torres TM, Gregory S, Ben Abdallah A, Kronzer A, McKinnon SL, Budelier T, Helsten DL, Wildes TS, Sharma A, Avidan MS: Using machine learning techniques to develop forecasting algorithms for postoperative complications: protocol for a retrospective study. *BMJ Open* 2018; 8: e020124
81. Kuo PJ, Wu SC, Chien PC, Rau CS, Chen YC, Hsieh HY, Hsieh CH: Derivation and validation of different machine-learning models in mortality prediction of trauma in motorcycle riders: a cross-sectional retrospective study in southern Taiwan. *BMJ Open* 2018; 8: e018252
82. Kim WO, Kil HK, Kang JW, Park HR: Prediction on lengths of stay in the postanesthesia care unit following general anesthesia: preliminary study of the neural network and logistic regression modelling. *J Korean Med Sci* 2000; 15: 25–30 [PubMed: 10719804]
83. Traeger M, Eberhart A, Geldner G, Morin AM, Putzke C, Wulf H, Eberhart LH: [Prediction of postoperative nausea and vomiting using an artificial neural network]. *Anaesthesist* 2003; 52: 1132–8 [PubMed: 14691625]
84. Knorr BR, McGrath SP, Blike GT: Using a generalized neural network to identify airway obstructions in anesthetized patients postoperatively based on photoplethysmography. *Conf Proc IEEE Eng Med Biol Soc* 2006; Suppl: 6 765–8
85. Tighe P, Laduzenski S, Edwards D, Ellis N, Boezaart AP, Aygtug H: Use of machine learning theory to predict the need for femoral nerve block following ACL repair. *Pain Med* 2011; 12: 1566–75 [PubMed: 21899717]

86. Hu YJ, Ku TH, Jan RH, Wang K, Tseng YC, Yang SF: Decision tree-based learning to predict patient controlled analgesia consumption and readjustment. *BMC Med Inform Decis Mak* 2012; 12: 131 [PubMed: 23148492]
87. Tighe PJ, Lucas SD, Edwards DA, Boezaart AP, Aytug H, Bihorac A: Use of machine-learning classifiers to predict requests for preoperative acute pain service consultation. *Pain Med* 2012; 13: 1347–57 [PubMed: 22958457]
88. Tighe PJ, Harle CA, Hurley RW, Aytug H, Boezaart AP, Fillingim RB: Teaching a Machine to Feel Postoperative Pain: Combining High-Dimensional Clinical Data with Machine Learning Algorithms to Forecast Acute Postoperative Pain. *Pain Med* 2015; 16: 1386–401 [PubMed: 26031220]
89. Hetherington J, Lessoway V, Gunka V, Abolmaesumi P, Rohling R: SLIDE: automatic spine level identification system using a deep convolutional neural network. *Int J Comput Assist Radiol Surg* 2017; 12: 1189–1198 [PubMed: 28361323]
90. Dawes TR, Eden-Green B, Rosten C, Giles J, Governo R, Marcelline F, Nduka C: Objectively measuring pain using facial expression: is the technology finally ready? *Pain Manag* 2018; 8: 105–113 [PubMed: 29468939]
91. Lotsch J, Sipila R, Tasmuth T, Kringel D, Estlander AM, Meretoja T, Kalso E, Ultsch A: Machine-learning-derived classifier predicts absence of persistent pain after breast cancer surgery with high accuracy. *Breast Cancer Res Treat* 2018
92. Wong LS, Young JD: A comparison of ICU mortality prediction using the APACHE II scoring system and artificial neural networks. *Anaesthesia* 1999; 54: 1048–54 [PubMed: 10540093]
93. Gottschalk A, Hyzer MC, Geer RT: A comparison of human and machine-based predictions of successful weaning from mechanical ventilation. *Med Decis Making* 2000; 20: 160–9 [PubMed: 10772354]
94. Ganzert S, Guttman J, Kersting K, Kuhlen R, Putensen C, Sydow M, Kramer S: Analysis of respiratory pressure-volume curves in intensive care medicine using inductive machine learning. *Artif Intell Med* 2002; 26: 69–86 [PubMed: 12234718]
95. Perchiazzi G, Giuliani R, Ruggiero L, Fiore T, Hedenstierna G: Estimating respiratory system compliance during mechanical ventilation using artificial neural networks. *Anesth Analg* 2003; 97: 1143–8, [PubMed: 14500172]
96. Nagaraj SB, Biswal S, Boyle EJ, Zhou DW, McClain LM, Bajwa EK, Quraishi SA, Akeju O, Barbieri R, Purdon PL, Westover MB: Patient-Specific Classification of ICU Sedation Levels From Heart Rate Variability. *Crit Care Med* 2017; 45: e683–e690 [PubMed: 28441231]
97. Perchiazzi G, Rylander C, Pellegrini M, Larsson A, Hedenstierna G: Monitoring of total positive end-expiratory pressure during mechanical ventilation by artificial neural networks. *J Clin Monit Comput* 2017; 31: 551–559 [PubMed: 27067075]

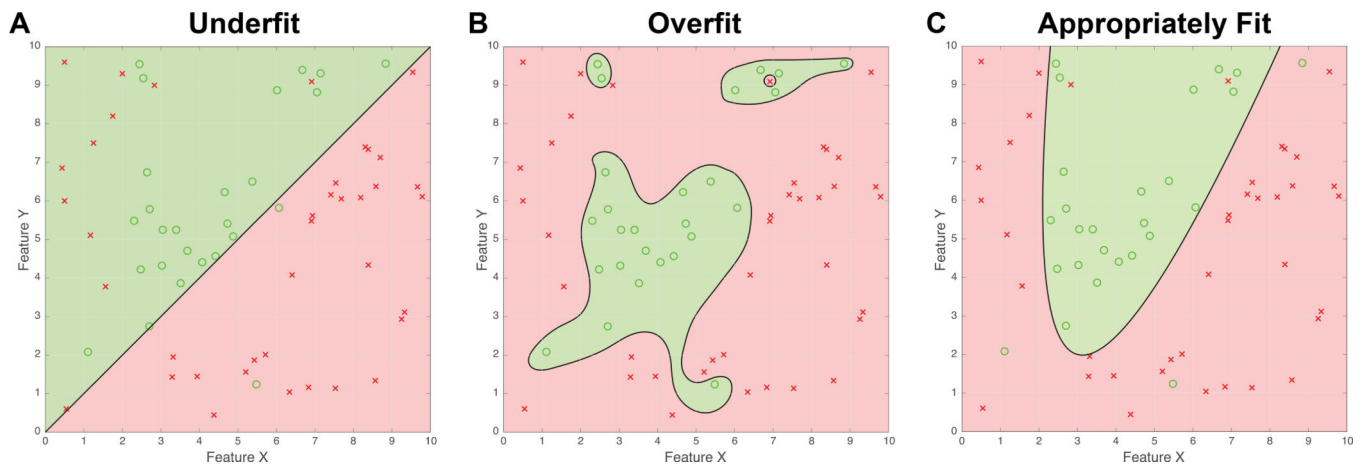


Figure 1: Examples of model fitting to data. The data are synthetic, for the purposes of illustration only.

(A) An under-fit representation of the data. Although the linear discriminator captures most of the green circles, numerous red crosses are misclassified. The linear model is too simple.

(B) The discriminator is over-fit to the data. Although there are no classification errors for the example data, the model will not generalize well when applied to new data that arrives.

(C) A parabola discriminates the data appropriately with only a few errors. This is the best parsimonious classification.

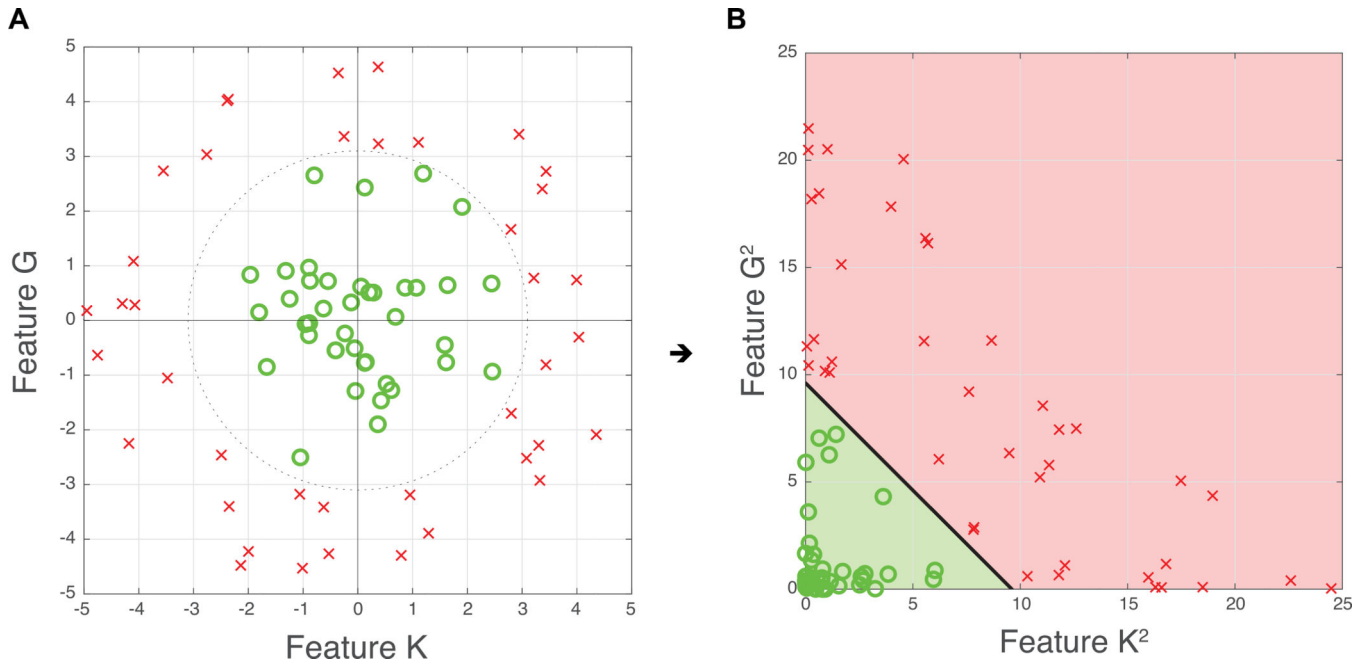


Figure 2:

Examples of model fitting using augmented variables. The data are synthetic, for the purposes of illustration only.

(A) An example of a model-fitting problem in which desirable outcomes (represented by green circles) are clustered around a mean point, and adverse outcomes (represented by red crosses) are associated with deviations from that point. For clinical correlation, one might imagine that the data represent favorable or unfavorable ICU outcomes based on rigorous control of potassium (Feature K) and glucose (Feature G).

(B) Rather than attempting to fit outcomes solely to the variables K and G, the variable space can be augmented by also fitting to K^2 and G^2 . This example demonstrates that the fitting of a perimeter around a mean value is easily accomplished by a linear fitting within the augmented space of K^2 and G^2 . The linear discriminant of $(K^2) + (G^2) - 9 = 0$ as shown produces a circular boundary of radius 3 in the K,G space.

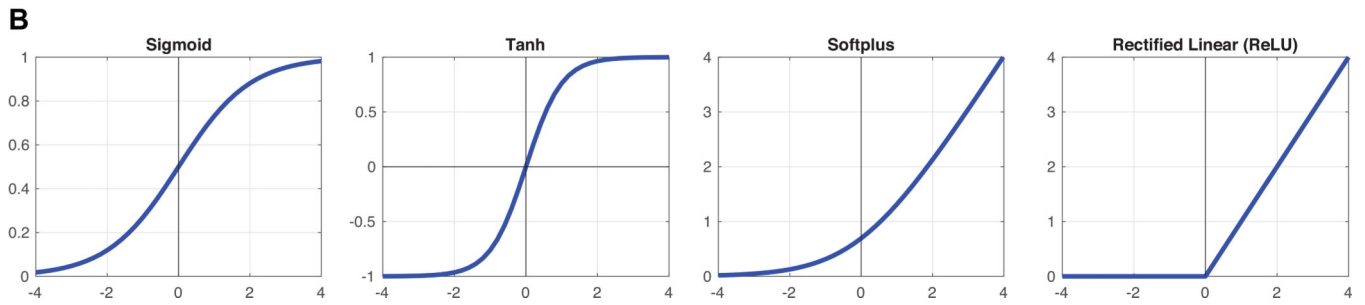
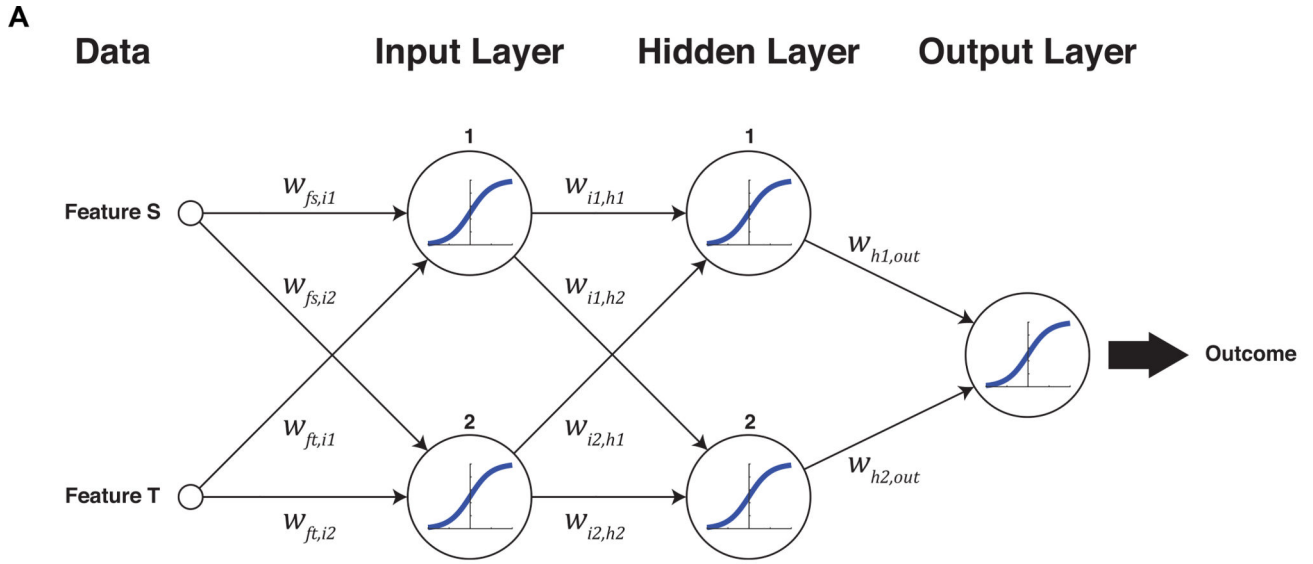


Figure 3:
(A) The simplest, fully-connected neural network from two input features to one output. The *weights* for each connection are illustrated, and each neuron in the network uses the sigmoid activation function to relate the sum of its weighted inputs, z , to its output. The sigmoid function is $\sigma(z) = (1 + e^{-z})^{-1}$
(B) Other biologically-inspired activation functions are possible and have practical benefits beyond the original **sigmoid**. Further evolutions are the **tanh** function (essentially two sigmoids arranged symmetrically), the **softplus** (the integral of the sigmoid), and the **rectified linear unit** (a non-smooth variant of the softplus).

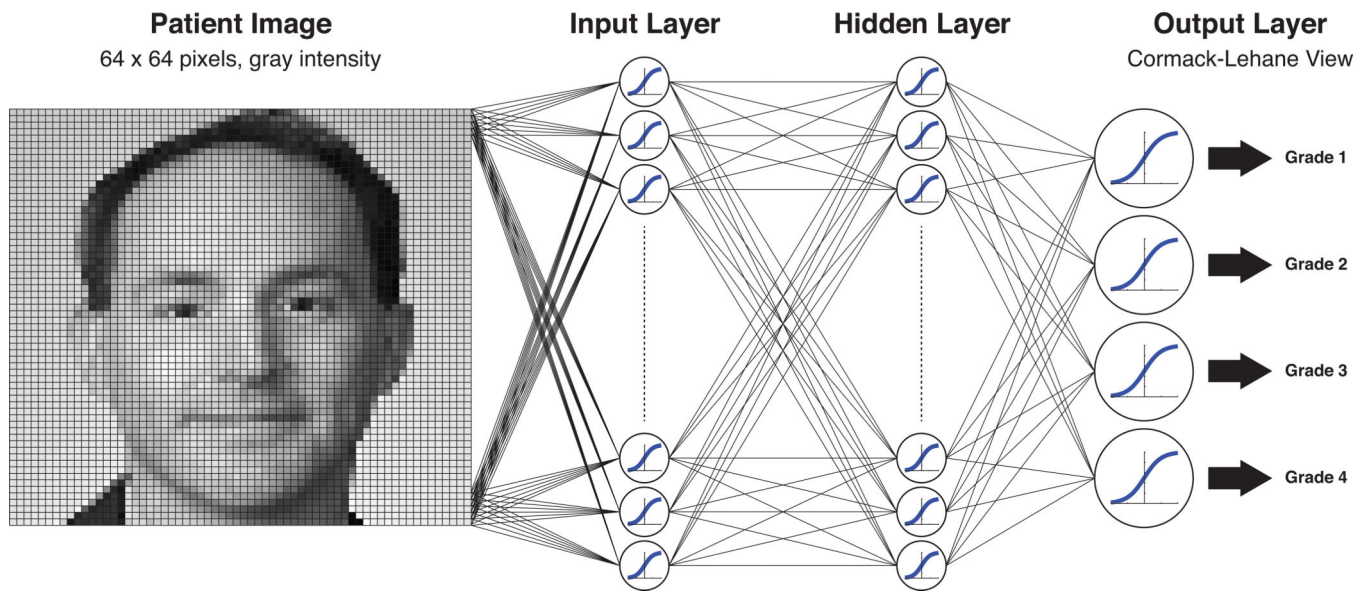


Figure 4:

The property of *universality* means that neural networks can represent any continuous function. The neural network shown here represents a hypothetical system to take a photographic image of a patient and render a prediction of their Cormack-Lehane view at intubation. (Not all nodes and connections are illustrated, as the input and hidden layers would each contain several thousand nodes. More pragmatic network topologies can be applied to visual recognition problems than the general case shown here.)

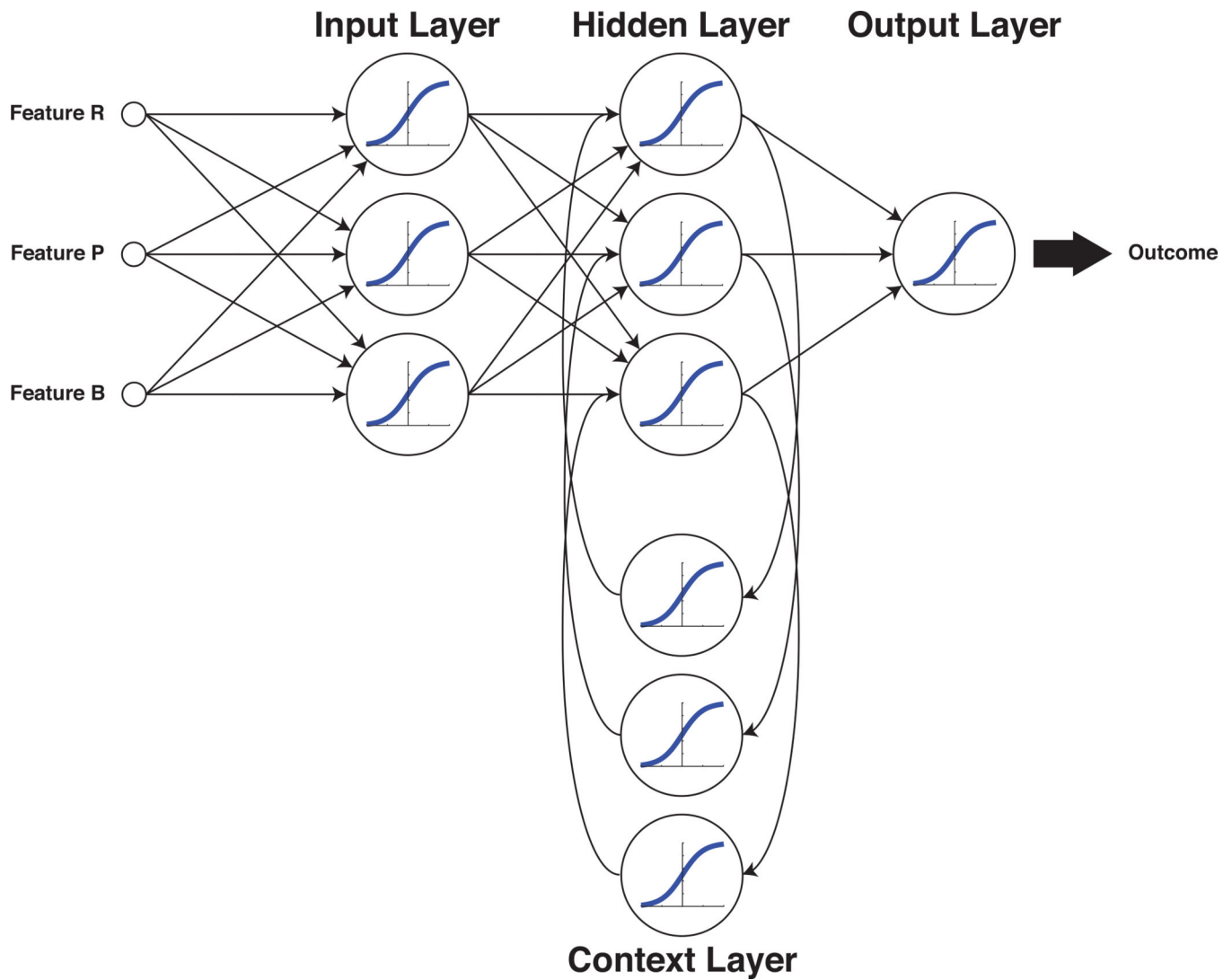


Figure 5: Recurrent Neural Networks (RNNs) employ feedback such that the output of the system is dependent on both the current input state and also the preceding inputs, enabling the network to respond to trends that evolve over time. In the Elman network arrangement shown here, the Context Layer feeds from and to the Hidden Layer.

Table 1

Results of a survey of the primary literature for articles on the application of artificial intelligence and machine learning to clinical decision-making processes in anesthesiology.

Topic	References
ASA Score and Preoperative Assessment	35,59,60
Depth of Anesthesia and EEG processing	61–65
PK/PD and Control Theory	41,66–73
Blood Pressure Homeostasis and Euvolemia	6,31,39,74–76
Surgical Complications and Trauma	77–81
Post-operative Care	82–84
Acute Pain Management and Regional Anesthesia	85–91
ICU Sedation, Ventilation and Morbidity	17,92–97

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript