EDITORIAL

Neural networks and deep learning: a brief introduction

Adrian Iustin Georgevici[1] and Marius Terblanche[2]*

Artificial intelligence (AI) refers to the theory and development of computer systems able to perform tasks normally requiring human intelligence. AI is at the top of the hype curve and an increasing number of studies based on machine learning are starting to appear in the intensive care literature [1–9]. One subset of machine learning techniques, deep learning (DL), has the potential to revolutionize the way healthcare is delivered. To make the most of the opportunity presented by this technological paradigm shift clinicians will need to acquire the knowledge and skills to correctly interpret the information generated by these complicated algorithms. Here we provide a brief introduction to deep learning and the mechanics of what happens inside the so-called black box.

Deep learning use representation learning, also known as feature learning, to map input features (analogues to prediction variables in traditional statistics) to an output [10, 11]. This mapping process occurs inside multiple connected layers which each contain multiple neurons. Each neuron is a mathematical processing unit which, combined with all other neurons, is designed to learn the relationship between the input features and the output (Fig. 1).

The first step in developing a deep neural network (DNN) is to determine the type of problem that needs to be solved. Examples of problem types include clustering, regression, classification, prediction, optimization, usage of sensors and motor controls in robotics, and vision. If, for example, the purpose is to predict mortality one is dealing with a classification problem, while aiming to predict a future event represents a prediction problem. Reading a chest radiograph would be a vision problem. Each of these types of problems requires a DNN designed specifically to solve it.

The nature of the data available to train a DNN is another important consideration. Supervised learning can occur if the data used to train the DNN is labeled, meaning the output (e.g. mortality in the case of a binary classification problem) is known. If the output label is unknown, unsupervised approaches are required. An example is the use of clustering to identify groups in the data, or dimensionality reduction to generate a less complex representation of the data. A third category known as semi-supervised learning uses labeled data sets to improve the results of unsupervised learning [12].
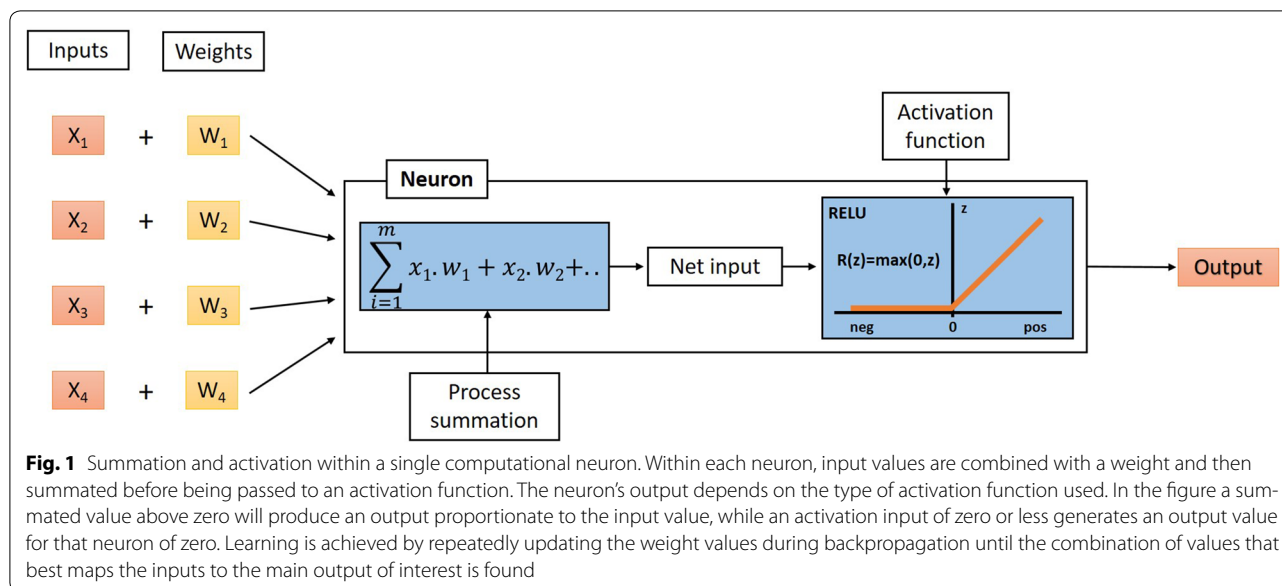
DL algorithms are complex mathematical structures with several processing layers that can separate the features of data, or representations, into various abstraction layers. In supervised learning, a DNN sequentially passes the input feature data from the neurons in one layer to the neurons in the next layer during a process that is repeated many, often thousands, of times (one cycle or iteration is known as an epoch). At each step information is extracted and passed to the next layer. Each neuron accepts weighted inputs from multiple other neurons (see Fig. 1). These inputs are summated and passed to an internal activation function, a hyperparameter chosen to optimize the model's performance. If the activation threshold is exceeded the neuron will generate an output which is combined with a weight value before being passed to multiple neurons in the next layer. If the threshold is not exceeded, the output will equal zero.

The 'knowledge' of the DNN is captured in its weight values. The weights are analogues to coefficients or odds ratios in a traditional statistical model. Compared to a large multivariable statistical model which might contain fewer than 50 coefficients, even small DNN can have

*Correspondence: marius.terblanche@kcl.ac.uk
[2] Department of Critical Care Medicine, St. Thomas' Hospital, Guy's & St. Thomas' NHS Foundation Trust, Westminster Bridge Road, London SE1 7EH, UK
Full author information is available at the end of the article

🌱 Springer

Fig. 1 Summation and activation within a single computational neuron. Within each neuron, input values are combined with a weight and then summated before being passed to an activation function. The neuron's output depends on the type of activation function used. In the figure a summated value above zero will produce an output proportionate to the input value, while an activation input of zero or less generates an output value for that neuron of zero. Learning is achieved by repeatedly updating the weight values during backpropagation until the combination of values that best maps the inputs to the main output of interest is found

many thousands of weights, while large recurrent or convolutional networks often have many millions.

Once information has passed through all layers, as described above, the model generates an output which is compared with the truth, based on the label value. An error is calculated and then used to update the weights in a process known a backpropagation. For this, a special mathematical function, known as the loss function, needs to be specified. The exact type of loss function depends on the nature of the model, but it is essentially a tool for evaluating the performance of a model on some given data, with a low value indicating better performance than a higher value. During the multiple learning epochs the model aims to minimize the error and find the combination of weight values that generates the lowest error value. This repeated updating of weights based on the size of the error is what is referred to as 'learning'.

Artificial intelligence, particularly deep learning, represents a considerable and, quite possibly, disruptive leap forward in improving the technologies supporting healthcare. Such computer algorithms are tools that can be trained to improve patient care quality by increasing diagnosis accuracy and decreasing workload stress on human care providers. They offer opportunities for automation and prediction not previously seen in healthcare. Very large data sets can be used to find correlations and patterns in clinical data in a way that is probably impossible using the unaided human brain or traditional approaches. Finally, deep learning systems can help care organizations to engage more effectively with staff and patients.

In future, deep learning systems may evolve to become sine qua non in that they perform the functions computers are good at while supporting caregivers in doing the things humans are good at.

Author details
[1] Department of Anesthesiology and Intensive Care Medicine, Katholisches Klinikum, St. Josef Hospital, Ruhr University, Gudrunstr. 56, 44791 Bochum, Germany. [2] Department of Critical Care Medicine, St. Thomas' Hospital, Guy's & St. Thomas' NHS Foundation Trust, Westminster Bridge Road, London SE1 7EH, UK.

Compliance with ethical standards

Conflicts of interest
The corresponding author has a majority shareholding in a technology startup company developing artificial intelligence solutions in healthcare.

Ethical approval
An approval by an ethics committee was not applicable.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Taylor K (2017) 12 medical technology innovations likely to transform healthcare in 2017. https://blogs.deloitte.co.uk/health/2017/01/12-medical-technology-innovations-likely-to-transform-healthcare-in-2017.html. Accessed 8 Nov 2018
2. Aczon M, Ledbetter D, Ho L, Gunny A, Flynn A, Williams J, Wetzel R (2017) Dynamic mortality risk predictions in pediatric critical care using recurrent neural networks. arXiv 1701.06675
3. Arguello Casteleiro M, Maseda Fernandez D, Demetriou G, Read W, Fernandez-Prieto M, Des Diz J, Nenadic G, Keane J, Stevens R (2017) A case study on sepsis using PubMed and deep learning for ontology learning. Informat Health 235:516–520

4. Raghu A, Komorowski M, Ahmed I, Celi L, Szolovits P, Ghassemi M (2017) Deep reinforcement learning for sepsis treatment. arXiv 1711.09602

5. Raghu A, Komorowski M, Celi LA, Szolovits P, Ghassemi M (2017) Continuous state-space models for optimal sepsis treatment – a deep reinforcement learning approach. arXiv 1705.08422

6. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA (2018) The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 24:1716–1720

7. Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH (2017) Improving palliative care with deep learning. In: 2017 IEEE international conference on bioinformatics and biomedicine (BIBM), Kansas City, 2017, pp. 311–316. https://doi.org/10.1109/BIBM.2017.8217669

8. Carneiro G, Oakden-Rayner L, Bradley AP, Nascimento J, Palmer L (2017) Automated 5-year mortality prediction using deep learning and radiomics features from chest computed tomography. In: 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017), Melbourne, 2017, pp. 130–134. https://doi.org/10.1109/ISBI.2017.7950485

9. Beaulieu-Jones BK, Orzechowski P, Moore JH (2017) Mapping patient trajectories using longitudinal extraction and deep learning in the MIMIC-III critical care database. bioRxiv 5:177428

10. Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning. MIT Press, Cambridge

11. Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2017) Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 19(6):1236–1246

12. Mamoshina P, Vieira A, Putin E, Zhavoronkov A (2016) Applications of deep learning in biomedicine. Mol Pharm 13:1445–1454