

Neural Networks for Medical Applications

Dr Ana C. Calderon* and Dr Simon Thorne⁺

{*acalderon, +sthorne}@cardiffmet.ac.uk

Department of Computing, Cardiff Metropolitan University, Cardiff, U.K.

With increases in data size and richness of available data, machine learning (henceforth ML) has seen a resurrection of interest in recent years. In a nutshell ML is a field that studies mechanisms in which computers carry out tasks without being explicitly programmed. It essentially uses large data sets to answer questions, by producing reliable results, as well as providing researchers with mechanisms for uncovering hidden results. As more research is accomplished, and its underlying algorithms become more powerful, we will increasingly witness machine learning's ubiquity in automated processes of benefit to all areas of society. ML's usages can be as varied as sorting through produce, self-driving cars, or speech recognition. In this article we will focus on a particular usage, how machine learning can benefit medical diagnostics and data analysis. The basic components of machine learning consists of:

- model creation,
- data collection and data preparation,
- training,
- evaluation.

Because of the complexity of the domain this article focuses on, we will firstly explain machine learning via a simple example, before moving on to explicating its benefits for medical purposes. Typically researchers will have a model they wish to train, and will therefore need to collect some data with which to train it; an important aspect is ensuring biases are, as much as possible, eliminated from the data. Importantly the stage at which the model is considered 'good enough' during training is pre-determined by researchers, depending on what question the model was created to answer. By 'good enough' we mean when the model is considered to have learnt enough. Machine learning, by its very nature, cannot guarantee the performance of its algorithms, and it is thus necessary for researchers to stipulate probabilistic bounds on the performance.

For instance, for a very simple example suppose you wish to create a model that can distinguish between wine, juice, and beer. Your data might consist of say sugar content, colour (maybe measured through a spectrometer), and alcohol content. The quality and quantity of data is directly correlated to how well your model will perform after training, so if more data was gathered about juice than the other categories, then there is clearly an imbalance here, and the model would typically guess 'juice' more often as it would have a greater chance of being right. Once the model has been trained, the next step is evaluating it, which must be done on different data, as otherwise it would just guess the answer.

Within the field ML, lies Neural Networks (NN), the first neural network ever made was in 1943 by neurophysiologist Warren McCulloch and mathematician Walter Pitts, based on a paper that sought to describe how neurons in the brain work. They created an approximate model using electrical circuits to explain how neurons might work in the brain. The first computer based neural network came in 1959, it was a system called ADALANE from Stanford university that detected and cancelled echo on telecommunications lines using adaptive filters which is still in use today. Whilst

neural networks provided a new and interesting approach to information processing and computation, computing power was at a premium at the time. Marvin Minsky, who laid the foundations of Neural Network computation, also discouraged research in the discipline in his book *Perceptrons* which some say led to the “AI winter” of reduced research funding, publishing and progress in Neural Networks that lasted between the 1970’s and the early 2000’s. However, since then, as computing power became more readily available, interest in Neural Networks has increased once again. This additional computing power has allowed more computationally demanding techniques such as “Deep Learning” to be deployed and generate compelling results. Despite their unpopularity in computer science, some researchers still pursued Neural Network research through the “AI winter”, recognising it’s great potential.

Neural networks are particularly useful when the problem being analysed has a degree of uncertainty, they tend to work best when our

conventional computation approaches have failed to turn up robust models. There are numerous examples of Neural Networks being used in medicine to this end. An example of some importance in the area of medical application of neural networks is in the diagnosis and surgical planning for horizontal strabismus. Strabismus is an anomaly of the eyes in which the eyes lose alignment with one another. One eye may fixate on the frontal point and the other may turn inwards, outwards, upwards or downwards. The treatment for this condition is usually surgery to the eye muscles to return binocular vision. The surgery involves either weakening one set of muscles or increasing muscular torque in another depending on the patients condition. There are four muscles that can be operated on and surgical intervention often involves surgery to two or three of these. The factors that influence the surgeons choice are complex, demanding both theoretical knowledge and practical experience from the surgeon. When planning the surgery, the patient’s condition is analysed in a number of ways. Measurements of visual acuity, refraction exam, binocular fixation, fundoscopic exam, primary and secondary gaze positions in the leftwards, rightwards, upwards and downwards positions. The critical unit of measure is how many millimetres of torque should be applied through either tightening or loosening the muscles to return the best binocular vision to the patient. Reserachers have analysed data from 114 surgical interventions and collated along with an expert judgement of each patient’s best course of action.

Example of neural network used for improving diagnostics in breast cancer

Diagnosing breast cancer is fraught with difficulties, the primary means of doing so is via mammogram scans and visual identification of tumours. In decades past, fine needle aspirations were also used up until they proved to be of questionable value and are an invasive technique. During the early 90s researchers attempted to minimise false positives in mammogram scans in order to decrease the occurrence of diagnosticians misidentifying malignant tumours as benign and visa versa due to the varieties of visual data and properties of both malignant and benign tumours. There has since been a strand of research demonstrating that neural networks can be used to improve the accuracy of breast cancer diagnosis, and features of tumours in mammogram scans can be measured in terms of: tumour radius, texture, perimeter, area, smoothness, compactness, concavity, symmetry and fractal dimension and the neural networks trained accordingly.

Researchers from the University of Wisconsin took measurements applied to 589 known cases of benign and malignant tumours and a dataset of benign and malignant tumour features was created. This data was then used to train a neural network to tell the difference, based on the measurement data, between cases of malignant and benign tumours. The resulting model was able to distinguish between malignant and benign tumours with an accuracy of 97%. The system was then deployed in Wisconsin University Hospitals as a decision support aid to the timely and correct diagnosis of many breast cancer cases potentially helping save many lives.

This data was then used to train a Neural Network to determine which muscles should be operated on and how much torque or loosening should be applied to each muscle measured in millimetres. The efficiency of the trained NN is measured by comparing the networks predictions and the expert judgement. The models error rate was 0.5mm for tightening the muscle and 0.7mm for loosening. This model therefore shows great potential over previous formula based techniques which relied on averaging patient values which doesn't capture the variability in each case. The trained model can then be used as a decision support aid for surgeons which should improve the outcome of these surgical interventions.

The beneficial potential of neural networks for the basis of clinical decision support is clear. They are able to represent complex relationships found in data that are not immediately obvious to human inspection. Depending on the algorithm used in the learning process, neural networks are also tolerant to noise and error in the training set data. Depending on the size and quality of the data set, neural networks can tolerate up to 15% error and still produce a robust result. In some cases, introducing noise to the dataset can actually improve the performance of the network.

However, the major problem with using these networks is the availability of data in sufficient quality and quantity. If the data is somehow unrepresentative or in insufficient quantity to cover all eventualities in the area of study, the resulting models can be brittle or unrepresentative of the real world situation.

Medical Decision Making

Clinical decision support systems are designed to facilitate clinical decision-making by health care providers. The landscape of CDSS providers is rather mixed. Our own experience is with groups providing large scale resources with 20,000 medical algorithms in multiple medical disciplines either online via XML forms or as a downloadable Android or IOS application. Other providers such as MDCalc and QXMD offer around 200 different medical calculators covering a range of medical specialties either as a web tool or a downloadable app for IOS, Android and Windows.

Surgical Example

A notable example of neural networks used in surgical interventions is on cranial hematoma, namely in making the decision of whether to open the skull and remove to remove the hematoma or wait for it to be reabsorbed into the body. Many variables are important in this judgement, the type of trauma and position of the hematoma have to be evaluated before reaching a decision. Taiwanese doctors and computer scientists have collated data relating to traumatic brain injuries since 1991. The dataset contains 12,640 cases cover 132 different parameters including gender, age, type of collision, type of fracture and type cognitive deficit. With the help of an expert surgeon, 32 of the 132 parameters were chosen as important features for the neural network training data. In this case, two types of neural network were used. The first is a Multi-Layer Perceptron which is standard, highly adaptable multi layered network that can solve almost any type of problem. The second neural network used was a Radial Basis Function network which is a 'feed forward' 'data clustering' network. Feedforward networks are simpler in that they have only one processing layer and input data is directly fed forward to the output layer. The results of the study show that both neural networks performed well, the radial basis function network correctly predicted 88% of cases and the multilayer perceptron model accuracy predicted 89% of cases. These results improve on other clinical decision support systems which are based on conventional statistics such as logistic regression and demonstrate the power of neural network based clinical decision support.

There are numerous examples of one off designed CDSS that covers one or two specific illnesses, generally these have better design, build and test ethics than mass produced CDSS systems.

There is no agreed standardised way of designing CDSS and each application tends to be unique to the specific condition it is intended to assist with. Therefore, there is very little continuity in the user interface or the functionality of each CDSS which can cause usability issues with the end user. Further, there are very few agreed standards to govern the use of such technology and as such many medical organisations do not have policies in place to guide users in the technologies safe use. However, the potential benefits of CDSS for clinicians are significant, in theory CDSS offer clinicians more reliable decision making, greater performance and fewer medical errors. Combined with machine learning, and in particular with neural networks, these benefits are greatly amplified, these are known as non-knowledge based CDSS (as oppose to their knowledge based counterparts, which will consist of a knowledge base and several if-then-rules). Non-knowledge based CDSS rely on some form of artificial intelligence, they are typically used post-diagnostic to suggest patterns for clinicians and researchers to investigate more deeply. Some of our findings on working with CDSS systems is highlighted in the text box.¹

Our experience with CDSS

We have found, in our investigations of impact of knowledge-based CDSS have on clinical practice, that human trust increases with usage, interactions between frequency of usage and intended usage, as well as that practitioner's validation of results changed with changes in usage frequency and type.

We also conducted a study in which one particular system was deployed with clinicians, the system is capable of diagnostic algorithms, blogging, (presenting and recording) clinical case studies, unit converters and social media feed. We found that the preferred focus was on algorithms for aiding in diagnosis (approximately 48% of the time, followed by case studies approximately 24% of the time), showing the potential for non-knowledge based CDSS, in particular for incorporating neural networks to aid in better diagnosis of diseases.

A note on combining ML and distributed ledger technology

Distributed ledger (or blockchain) technology has been called the 'next big thing', its powers lies both in how machines in a blockchain can maintain their anonymity, as well as alleviating security concerns, given how difficult a blockchain system is to penetrate. Block chain is the underlying mechanism in crypto currencies, which might be more familiar to the reader, the mechanism is such that the privacy of data transfer and authentication are guaranteed. Security and privacy in the medical discipline are of the utmost importance. A blockchain is a database composed of digital (unchangeable) records, called blocks, stored in a chain, each block contains cryptographically encrypted data which contains information about the "previous" block. The database is shared between several parties, such that everyone has a consist view of it, and the integrity of the records is formed by a consensus amongst all authenticating parties. Distributed ledger technologies can make the services they are applied to more transparent and trustworthy, without compromises of security or privacy. If block-chain was fully implemented in a country's public health service, it could, in theory, allow researchers to investigate and gather patient data without compromise to patient anonymity. The possibilities would then be virtually endless, for instance researchers could pattern spot for disease commonalities, to investigate unknown factors linked to a particular illness;

¹ If you would like more information on this, or any other research mentioned in this article, please contact the authors.

combined neural networks this would be a fruitful area to explore pattern-recognition for medical uses as well as enhance the NN applications mentioned in this article.