Machine Learning for Healthcare: Course Summary

What is machine learning (ML) and why healthcare?

Artificial intelligence can be defined as "the science of making machines do things that would require intelligence if done by men". Machine learning can be used to obtain this goal, by computers learning without being explicitly programmed.

Machine learning can be defined technically as:

"A computer program learns from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E." – Tom Mitchell (1998)

The attributes of machine learning put it in a good place to have a significant impact on several aspects medicine: it is excellent at analysing large volumes of data to identify complex relationships, and enables analysis of previously-hard-to-analyse data types, such as images and text.

It is one of the most active areas of medical research and innovation, with many exciting studies being published in the last few years and many companies working on exciting projects.
Recommended resources: ***https://explainthispaper.com/***, ***https://chrislovejoy.me/research*** ***and*** ***https://chrislovejoy.me/companies/***

It will affect different specialties differently, in part determined by the amount of 'pattern recognition' the specialty involves, as well as the volumes and type of data the specialty uses for diagnosis, monitoring and treatment decisions.

Machine learning in medicine is a cross-disciplinary area. Many big contributions are being made by people from computer science and data science backgrounds. We believe that people from medical backgrounds are key contributors to on-going research and innovation.
Recommended reading: ***https://chrislovejoy.me/why***

How does the machine *learn*?

There are broadly four types of machine learning; supervised, unsupervised, semi-supervised and reinforcement learning. The majority of medical studies to-date have involved supervised learning.

The primary supervised learning tasks are regression, the prediction of a continuous variable from known inputs, and classification, the prediction of a categorical variable from known inputs. The aim of the machine learning algorithm is to accurately predict the value or category of a new object, when given a known variable. An example of regression is to predict someone's weight if we know their height. An example of classification is to predict the presence or absence of cancer given information about a skin mole.

The machine learning algorithm is 'trained' to spot the patterns by analysing data where the value or category is already known. This enables it to predict the value or category when presented with new examples, where the outcome is unknown.

There are several methods by which an algorithm learns. Gradient descent is a common technique used in linear regression, logistic regression, neural networks and deep learning. It involves iteratively updating the values in an algorithm to minimise the amount of error.
Recommended resource: First five weeks of 'Machine Learning by Andrew Ng' on Coursera.

Other machine learning algorithms, such as nearest neighbour classification and random forest classification, have their own methods of learning.
Recommend resource: Khan Academy

Unsupervised learning is where the machine learns from data without a labelled meaning. For example, in supervised machine learning the data may be 1,000 patients' blood test results and each is labelled as either having an infection or not. In unsupervised learning, the blood test results wouldn't be labelled. A common unsupervised learning task is to cluster data into groups. In the above example, if asked to find two clusters, the algorithm may identify the infected and not infected groups, but not be able to say what they are. Alternatively, if asked to find three groups, it may find uninfected, moderate infection and severe infection. One application of research to date is identifying subgroups and phenotypes of particular patients.
Recommended resource: Informatics and Machine Learning to Define the Phenotype by Basile and Ritchie (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6080627/)

Reinforcement learning is where a machine is given rewards or punishments for how well it performs on a task. If a reward is given, for example if a patient survives, it will try to learn the actions giving the best chance of that outcome. There are challenges of attempting such studies in medicine. Existing research most often looks at trying to find optimal treatment strategies.
Recommended reading: Review of Recent Reinforcement learning applications to healthcare – Towards Data Science (https://towardsdatascience.com/a-review-of-recent-reinforcment-learning-applications-to-healthcare-1f8357600407)

Deep Learning in Medicine

Neural networks are biologically-inspired machine learning algorithms which can perform regression and classification tasks well on large volumes of data with many variables.

Their calculations involve multiplying 'nodes' with 'weights', summing this up and then multiplying it by an 'activation function'.
Recommended resource: First five weeks of 'Machine Learning' by Andrew Ng on Coursera.

'Deep learning' involves neural networks with multiple inner layers. They can take different structures, including convolutional neural networks (CNNs) which excel at medical imaging and recurrent neural networks (RNNs) which excel at sequential data. The majority of published healthcare AI research involves CNNs.
Recommended resource: 'Deep Learning Specialisation' by Andrew Ng on Coursera.

One use of RNNs is in natural language processing (NLP). In healthcare, this may enable analysis of medical notes, automatic dictation and the development of chatbots.
Recommended reading: "Healthcare chatbots are here to help" (https://pulse.embs.org/may-2019/health-care-chatbots-are-here-to-help/)

Assessing performance and clinical utility

For a model to be useful it must fit the data appropriately and it must represent the wider population. This is demonstrated by testing models in new datasets and quantifying performance with performance metrics.

Some of the most popular performance metrics in machine learning are the area under the ROC curve (AUC), precision, recall and F1 score. These metrics enable us to interpret different elements of an algorithm's performance.
Recommended reading: ***https://lukeoakdenrayner.wordpress.com/2018/01/07/the-philosophical-argument-for-using-roc-curves/***

When training a model, we should divide up some data to train the model and other data to test it.
Recommended resource: Week 6 of 'Machine Learning' by Andrew Ng on Coursera.

To assess clinical utility, we should ask:

1. Does it work? (Assessed by performance metrics)
2. Does it improve clinical outcomes?
3. Is it safe?

There are many cases where promising new ideas may not be able to tick all three when translated to reality.

There are several levels of clinical evidence for machine learning algorithms being used in healthcare:

- Retrospective studies, where you have test a model on data that has already been collected
- Prospective (non-randomized) studies, where you test a model on new data as it comes in
- Randomized control trial (RCT) studies, where you divide up people into two groups, then use the AI model on one and not the other, and compare the outcomes

To-date, the majority of studies are retrospective, although we are starting to see more prospective studies and RCTs.
Recommended resources: "Artificial intelligence versus clinicians" – Nagendran et al., BMJ (https://www.bmj.com/content/368/bmj.m689)

Next steps, careers and resources

The most common roles for those from a medical background looking to apply machine learning to healthcare are healthcare AI research and input to health technology start-ups and companies, although there are other roles.
Recommended resources: ***https://chrislovejoy.me/career***

The areas to learn can be broadly divided into (i) mathematical foundations (ii) machine learning theory, (iii) computer programming and (iv) machine learning applied to healthcare. The level of understanding required in each area depends on the role desired.
Recommended resources: ***https://chrislovejoy.me/foundation***