Using Machine Learning and Natural Language Processing to Automatically Diagnose Medical Conditions

Amartya Vadlamani

With an growing and aging population, getting fast access to medical advice is vital to ensuring a good quality of life but ever growing pressures on the NHS mean that it can take between 3 and 18 weeks to see a specialist[1], and up to four hours to go into a walk-in clinic to see a nurse if you ever get seen at all[2]. Online symptom checkers aren't that useful either as according to a BMJ study there is only a 57% chance that correct diagnosis will be in the top 20 options[4]. The bottleneck in all of these cases is the lack of doctors who can interview patients and produce a diagnosis quickly and reliably.

Now if you break up the tasks doctors do while in a diagnostic appointment; process human input; analysis this input to find relations and patterns; and produce a given treatment plan; the task seems ripe for automation by computers. There are two fields in computer science, Machine Learning and Natural Language Processing, two fields that started in the 50s but only recently had enough computing power to become applicable to many different tasks. Machine Learning lets machines learn from structured datasets, i.e. datasets in a machine readable format, and perform tasks that they could never be manually programmed to do.[5] While this is all well and good, where would you find a structured dataset large enough to teach a computer the whole of medicine? This is where Natural Language Processing comes in. This subfield of computational linguistics is based around finding methods that computers can turn the unstructured datasets that people produce when they write and turn them into structured datasets that a computer can understand.[5] Put these two fields together and feed it from a medical database of thousands of research papers and you get a digital doctor that scales with the number of computers that you can attach it to rather than the number of people that have graduated from medical school in the last few years. That's the theory at least.

In practice, there are a few teething issues but the technology is now mostly here, in the form of machine learning solutions like IBM's Watson[5]. A question answering computer program that is able to answer questions in plain English. A point to note is that AI systems like Watson do not have to be perfect just better than people and in some instances they already are. Watson, for example, already diagnoses lung cancer with 90% accuracy as opposed to 50% accuracy by human doctors and has made a leukemia diagnosis that had stumped a suite of human doctors for months. [6][7].

Now while these figures show quite a lot of promise there still a long way before a future of fast and easy access to digital doctors. But the technology is currently expensive as the algorithms need a lot of computation power to run and this means that data centers need to either be set up or rented and while strides are being made to make these systems more efficient they still need large computer arrays to serve the ever growing population that was the catalyst for their creation. As well as this more data needs to be gathered about where the inherent biases are in this system as well as reducing the number of times that it makes false predictions. But this could mean the earlier diagnosis and treatment of a whole range of time sensitive diseases like cancer.

References

[1] NHS UK. Guide to NHS waiting times in England . http://www.nhs.uk/NHSEngland/appointment-booking/Pages/nhs-waiting-times.aspx (accessed 05 December 2016).

[2] Mark Tran. GP appointment waiting times crisis revealed in Labour research. https://www.theguardian.com/society/2016/apr/17/gp-appointment-waiting-times-crisis-revealed-in-labour-research (accessed 05 December 2016).

[3] Central London Community Healthcare. NHS Walk-in Centres. http://www.clch.nhs.uk/services/walk-in-and-urgent-care-centres.aspx (accessed 05 December 2016).

[4] Hannah L Semigran, Jeffrey A Linder, Courtney Gidengil, Ateev Mehrotra. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ (British Medical Journal) 2015; http://www.bmj.com/content/351/bmj.h3480 (accessed 05 December 2016).

[5] David Ferrucci , ferrucci@us.ibm.com, Anthony Levas, Sugato Bagchi, David Gondek, Erik T. Mueller. Watson: Beyond Jeopardy!. Science Direct 2012; http://www.sciencedirect.com/science/article/pii/S0004370212000872 (accessed 05 December 2016).

[6] Alfred Ng. IBMs Watson gives proper diagnosis for Japanese leukemia patient after doctors were stumped for months . http://www.nydailynews.com/news/world/ibm-watson-proper-diagnosis-doctors-stumped-article-1.2741857 (accessed 05 December 2016).

[7] Ian Steadman. IBM's Watson is better at diagnosing cancer than human doctors. http://www.wired.co.uk/article/ibm-watson-medical-doctor (accessed 05 December 2016).