A checklist for papers applying machine learning to healthcare

These questions are intended to provoke thought and aid the systematic examination of papers applying machine learning to healthcare. They are not comprehensive nor particularly detailed. For those interested in a more thorough examination of such research papers, we refer you to the references at the bottom of this document.

Overview

- What did they *do*? (in simple English)
- How does this fit into the wider context?
    - What other studies have been done in that specialty?
    - What other studies have been done looking specifically at this problem?
    - What does this study add?
- Who did the study? (what authors and institutions)

Data

- What type of data is used? (images? blood test results? text? scoring systems?)
- How much data did they use? (is it enough?)
- How was the 'ground truth' defined? (histology vs radiological vs expert opinion? how many experts labelled?)
- Is the data skewed towards particular classes? (how was this handled?)
- Does the data represent the population the model will be deployed on? (appropriate mix of demographics, locations and diseases?)

Methodology

- What techniques did they use? (e.g. classifiers? CNNs?)
- What were the output measures? (e.g. diagnosis? prediction? decision? detection rate?)
- What type of study was it? (Was it retrospective or prospective?)
- What was the rationale for their approach? Did they explain in enough detail to replicate the study?

Performance

- How did the model perform? (what was the accuracy? AUC? precision and recall?)
- What features were important in the model? Do these make sense?
- If compared with clinicians, was the control group appropriate? (how many experts? what level of expertise?)
- How would this be used in a clinical setting?

Conclusions

After having read all this, consider:

- What did the authors conclude? Is it justified?
- What are the next steps? (Is this ready to be applied clinically? What further evidence would you want? If you were to design the next study, what would it look like?)

Further Reading

1. *"How to Read Articles That Use Machine Learning - Users' Guides to the Medical Literature" – Faes et al., Translational Vision Science & Technology, Feb 2020*
2. *"A Clinician's Guide to Artificial Intelligence: How to Critically Appraise Machine Learning Studies" – Liu et al., JAMA, Nov 2019*

Dr Chris Lovejoy
www.chrislovejoy.me