International Journal of Population Data Science

Journal Website: www.ijpds.org

Approaches to Capacity Building for Machine Learning and Artificial Intelligence Applications in Health

Paprica, PA¹, Sullivan, F², Aphinyanaphongs, Y³, and Gibson, G¹

¹Vector Institute

²University of St. Andrews

³NYU Langone Health

Many health systems and research institutes are interested in supplementing their traditional analyses of linked data with machine learning (ML) and other artificial intelligence (AI) methods and tools. However, the availability of individuals who have the required skills to develop and/or implement ML/AI is a constraint, as there is high demand for ML/AI talent in many sectors. The three organizations presenting are all actively involved in training and capacity building for ML/AI broadly, and each has a focus on, and/or discrete initiatives for, particular trainees.

P. Alison Paprica, Vector Institute for artificial intelligence, Institute for Clinical Evaluative Sciences, University of Toronto, Canada. Alison is VP, Health Strategy and Partnerships at Vector, responsible for health strategy and also playing a lead role in “1000AIMs” – a Vector-led initiative in support of the Province of Ontario’s \$30 million investment to increase the number of AI-related master’s program graduates to 1,000 per year within five years.

Frank Sullivan, University of St Andrews Scotland. Frank is a family physician and an associate director of HDRUK@Scotland. Health Data Research UK <https://hdruk.ac.uk/> has recently provided funding to six sites across the UK to address challenging healthcare issues through use of data science. A 50 PhD student Doctoral Training Scheme in AI has also been announced. Each site works in close partnership with National Health Service bodies and the public to translate research findings into benefits for patients and populations.

Yin Aphinyanaphongs – INTREPID NYU clinical training program for incoming clinical fellows. Yin is the Director of the Clinical Informatics Training Program at NYU Langone Health. He is deeply interested in the intersection of computer science and health care and as a physician and a scientist, he has a unique perspective on how to train medical professionals for a data drive world. One version of this teaching process is demonstrated in the INTREPID clinical training program. Yin teaches clinicians to work with large scale data within the R environment and generate hypothesis and insights.

The session will begin with three brief presentations followed by a facilitated session where all participants share their insights about the essential skills and competencies required

for different kinds of ML/AI application and contributions. Live polling and voting will be used at the end of the session to capture participants’ view on the key learnings and take away points.

The intended outputs and outcomes of the session are:

- Participants will have a better understanding of the skills and competencies required for individuals to contribute to AI applications in health in various ways
- Participants will gain knowledge about different options for capacity building from targeted enhancement of the skills of clinical fellows, to producing large number of applied master’s graduates, to doctoral-level training
- After the session, the co-leads will work together to create a resource that summarizes the learnings from the session and make them public (though publication in a peer-reviewed journal and/or through the IPDLN website)