A Conducted Data Tour in the Health Domain

Anjan Kr Dasgupta

31/07/2020

"Water, water, every where,

Nor any drop to drink."

The Rime of the Ancient Mariner

Samuel Coleridge.

Synopsis

Background

E-health which may have a number of synonyms, essentially is a balanced superset of Data Science and ICT. These two backed up by suitable models are the oft-used "invisible hands" that set the trends. In health and drug research domain, AI applications for example, is considered as potential but as dangerous as the "self-driving cars".

Unlike bioinformatics which is based on ordered and annotated data, and has mainly indirect predictive abilities, E-health has direct implications in the wet clinical laboratories. Unlike bioinformatics, a small error in prediction may cause loss of human life which is more fatal than the crash of the "self-driving-cars". An example that is often showcased is the follwing. A programming error in a smart pacemaker by a reputed manufacturer caused numerous loss in life, and FDA approval was withdrawn for that model.

Data versus the model - "Content and Form"

The tremendous challenge a data scientists face is often not an absence of data. Following Coleridge, we may say "data data every where", but not a drop left for "analytics". Data is the new "gold/oil" reserve of the advanced world

Several laboratories from the West took initiatives to keep the digital data records, perhaps anticipating the emergence of some future data based commodities. Many pharma-companies indeed used data as a commodity, and data on health and clinical records became subjects of hot extreme political and rights-issue based debate.

In contrast, we have a poor record of book-keeping data. Unless health data is aggregated and organized, the dream of E-health will be far fetched. Even West also has the problem

in inappropriate data recordings, the Lancet gate involving HCQ, being the most recent evidence.

The model and data are often as closely related as the structure and function of a molecule, or the form and content of any language. Quoting Kalidasa "Vagarthaviva sampriktau vagarthah pratipattaye" where he highlighted the inseparablity of speech and its meaning,we have similar entanglement between data and model. The presentation raises some insights into this interplay particularly in context of health issues and the management of a pandemic like scenario.

Analytics ethics and design in E-health

Design and ethics of data are as important as the analytics. Every day we are wasting several terabytes of data, by not collecting it. We need to motivate the health workers and researchers so that they realize the importance of data accountancy auditing and ethical recording. We also need to develop the appropriate ICT framework that enables affordable acquisition of data and its aggregation and integration to a fog or cloud.

We cite an example of Physionet (MIT package) and show the successful combination of smart data acquisition and generation of a futuristic E-health platform.

A glimse of E-health Future - Block chains in E-health

Finally, we shall show how some additional baby steps in the E-health domain which is promising, but at this stage as risky as self-driving cars. The possible application of block chain technology is thus mentioned at the end.