Contribution ID: **305**   Type: **Oral Presentation**

ANALYSIS OF TYPE 1 DIABETES WITH MACHINE LEARNING METHODS

Wednesday, 27 June 2018 10:40 (20 minutes)

ANALYSIS OF TYPE 1 DIABETES WITH MACHINE LEARNING METHODS

ABSTRACT

Big data is a term that we can use for data sets with large, diverse and complex data structures that are difficult to analyse or visualize using traditional computing methods and approaches. Health care sectors rely on interpreting data gathered from patients, which ranges from graphic scans to medical images of which there is a limitation on its analysis by medical practitioners. It is this same limitation for datasets and advances in big-data methodology that sparked interest an interest in this sector as well. We proposed to use the Type 1 Diabetes (T1DM) data from INDEPTH VA as a test case for applying modern machine learning (ML) methods to analyse the data. Due to a high prevalence of Type 1 Diabetes (T1DM) it has huge amounts of data available from on-going researches, and there are possible misclassifications of such data in past researches. In collaboration with Dr Alisha Wade(DG,Johannesburg) and a study by Prof Justine Davies (KCL,London,UK) which aimed to estimate the numbers of deaths in people with T1DM and the causes of those deaths by using VA data in the under 40 age-group across all INDEPTH sites. The same data will be processed by classification algorithms based on machine learning. The basic idea is that the algorithm automatically learns to make accurate predictions based on past observations. For this particular case, it will learn the features present in diabetic patients, thereby able to identify patients with high chances of unidentified T1DM. One of the advantages of this approach is that it can avoid human bias, but to not fall prey to any training bias, and understand the correlation between different properties, inputs from medical experts will also be necessary. The feasibility of this approach will be studied to see if accuracy obtained with current methods can be reached or bettered. If this preliminary study is successful, the scope can be expanded. For this study, we will use Tensorflow [2] or Scikit Learning [3], two of the most popular deep learning libraries to classify the dataset. They are both open source software.

**Please confirm that you
have carefully read the
abstract submission instructions
under the menu item
"Call for Abstracts"
(Yes / No)**

Yes

**Consideration for
student awards
Choose one option
from those below.
N/A
Hons
MSc
PhD**

MSc

**Supervisor details
If not a student, type N/A.
Student abstract submision
requires supervisor permission:
please give their name,
 institution and email address.**

Dr Deepak Kar- The University of the Witwatersrand

deepak.kar@cern.ch

Dr Jong Soo Kim- The University of the Witwatersrand

jongsoo.kim@tu-dortmund.de

Primary author: Ms MANAKA, Thokozile (University of the Witwatersrand)

Co-authors: KAR, Deepak (University of Witwatersrand);   Dr SOO KIM, Jong (University of the Witwatersrand/NITheP)

Presenter: Ms MANAKA, Thokozile (University of the Witwatersrand)

Session Classification: Applied Physics

Track Classification: Track F - Applied Physics