RAS Working Group meeting 20.11.2020

Participants: S. Albright, R. Alemany-Fernandez, A. Apollonio, G. Azzopardi, A. Camacho, T. Cartier-Michaud, Lajos Cseppentö, D. Dlugosz, L. Felsberger, B. Fernandez Adiego, E. Fol, B. Goddard, Julia H., V. Kain, B. Li, A. Müller, D. Nisbet, C. Obermair, K. Samaras-Tsakiris, R. Schmidt, V. Schramm, B. Todd, J. Uythoven, S. Uznanski, E. Vergara Fernandez, F. M. Velotti, W. Vigano, J. Wenninger

This RASWG has been held in combination with the weekly ML coffee.

The slides of the presentation can be found on the ML coffee Indico event page:

https://indico.cern.ch/event/974740/

Putting ML into Practice: Data-Driven Reliability Optimization for Particle Accelerator Systems - Speaker: L. Felsberger

L. Felsberger presented an overview of his PhD thesis, which is based on the work carried out at TE-EPC-CCE in the recent three years. He gave a motivation for the work by pointing out that future particle accelerators, such as the FCC, will require even more reliable equipment at lower cost than the LHC to reach the availability and cost targets. The goal of the thesis is to study how data driven methods can be used to optimize the reliability of equipment at low costs. The speaker tested data driven methods for reliability optimization on three use cases at CERN and provided some lessons learned.

Slide 25: B. Goddard asked whether recurrent neural networks were used. L. Felsberger responded that in benchmark tests they do not perform better than convolutional neural networks but tend to be harder to train. Hence, they were not used but the focus was on convolutional neural networks.

After the presentation, W. Vigano pointed out that a proper use of traditional reliability prediction methods leads to very precise reliability predictions. J. Uythoven added that the experience at CERN confirms this. A. Apollonio commented that for new and more complex machines, the traditional approach might prove less cost effective than data driven approaches.

V. Kain mentioned that in IT methods have been developed to facilitate the labelling of data by system experts (presented during the previous MLcoffe event: https://indico.cern.ch/event/974740/). This could solve some problems of data availability.

B. Todd commented that the necessary data on failure mechanisms might not be collected when the failing systems are cheap and simply get replaced without further analysis.

R. Schmidt asked whether the conclusion from the last use case is that converters should always be treated as if they are developed and produced in large quantities. L. Felsberger responded that one has to be cautious about the interpretation of the results and that it would need further investigation.

V. Schramm asked whether the third use case was also used with data indicating if failures occur early or late in the life cycle. L. Felsberger answered that the data was not available to do so but that it could reveal interesting findings.

B. Goddard asked if unsupervised learning was considered as the failure data is very sparse. L. Felsberger responded that supervised learning was preferred, as for unsupervised learning and the popular anomaly detection methods it cannot be said if an anomaly is of operational origin without consequences or leading to failure.

R. Alemany-Fernandez asked whether root cause detection was addressed by Machine Learning (ML) methods. L. Felsberger explained that ML methods were used to facilitate root cause detection by human experts but not to fully automatize this process as he believes that it requires a set of skills which AI is far from demonstrating.

K. Samaras-Tsakiris recommended to use scientific ML methods to infer the parameters on the modelling approach in the second use case.