Mining condensed set of patterns from time series data for explainable prediction

Planed funding: ANR XQUALITY project
Postdoc supervisors: Ahmed SAMET
Contact: ahmed.samet@insa-strasbourg.fr
Workplace: INSA Strasbourg
Laboratory: ICUBE laboratory
Team: SDC
Startup scheduled: Date of the start could be planned with the selected candidate after January 1st 2023

Project in a few words:

French and German manufacturing companies are known for their high-quality production and their orientation towards smart factories. Quality assurance and control of complex production systems is a major challenge and is further exacerbated by the shortage of skilled labour in this domain. Quality problems must be detected and eliminated quickly. When a quality problem is detected, it is necessary to quickly understand the multiple possible causes for it (which can sometimes be contradictory to each other) in order to propose the most appropriate corrective actions to return the manufacturing process to its normal operating mode.

The **XQuality** project is researching hybrid and explainable AI approaches to help manufacturing companies implement intelligent and automated quality assurance. The project combines data-based machine learning, semantic technologies and expert knowledge to monitor and explain product and process quality targets in a company. The goal is to develop an AI-based system that will assist the staff in identifying the main causes of quality issues as early as possible, to achieve reliability engineering in the domain of manufacturing, thanks to the new quality assurance models.

PostDoc Subject:

XQuality will develop new mining algorithms that analyze time series data and generate a condensed set of rich patterns called chronicles (Sellami et al. 2020). Unfortunately, as highlighted in literature, the number of chronicles is still high and their mining is time expensive. Therefore, we aim to reduce the set of generated patterns by optimising several quality set metrics such as discriminance, confidence to cite a few (Sahoo et al 2015) aiming toward more explainable patterns.

From a methodological point of view, we aim to propose a new mining algorithm to mine rich and sequential patterns called chronicles. A new condensed set based on the closure definition of patterns is proposed. Additionally, machine troubleshooting documents will be mined using a CRF based technique to find corrective actions to each quality loss situation. The idea is to associate patterns to situations and corrective actions. This quality of this set is evaluated on a predictive task to assess the accuracy of quality loss prediction. we extend the work of (Sellami et al 2020) by

integrating several constraints in the chronicle mining algorithm. These quality constraints are related to either the domain knowledge or to quality rule metrics such as discriminance, confidence.

Two directions could be undertaken: either by declarative programing or optimization process (Guyet et al 2017). The found condensed set of chronicles is applied on a prediction task of quality loss prediction to ensure quality loss monitoring.

Keywords: Pattern mining, condensed sets, Constraint programming, Explainable AI, Uncertainty modelling, Industry 4.0, Quality control.

Candidate's skills: The job requires a strong pattern mining and mathematical background. A previous experience on data mining, constraint programming/uncertainty modelling are highly recommended. Having a PhD contribution pattern mining and explainable AI on time series is highly appreciated. Additionally, previous experience on Industry 4.0 and quality control is highly appreciated.

Specific knowledge: Knowledge in time series pattern mining is expected from any application. Applicants should show the adequacy of their profile to the project description. Having published in top ranked data mining conference is highly appreciated.

Desired education: Junior PhD profile in computer science is desired.

Salary: Monthly gross salary (before tax) : ~ 2500 euros. Postdoc fellowship according to the INSA salary scale

References:

Sellami, C., Miranda, C., **Samet**, A., Tobji, M. A. B., & de Beuvron, F. (2020). On mining frequent chronicles for machine failure prediction. Journal of Intelligent Manufacturing, 31(4), 1019-1035.

Guyet, T., Happe, A., & Dauxais, Y. (2017, June). Declarative sequential pattern mining of care pathways. In Conference on Artificial Intelligence in Medicine in Europe (pp. 261-266). Springer, Cham.

Sahoo, J., Das, A. K., & Goswami, A. (2015). An effective association rule mining scheme using a new generic basis. Knowledge and Information Systems, 43(1), 127-156.