Executable Knowledge Graph for Transparent Machine Learning in Welding Monitoring at Bosch

Zhuoxun Zheng
zhuoxun.zheng@de.bosch.com
Bosch Center for AI, Germany
Oslo Metropolitan University, Norway

Baifan Zhou
baifanz@ifi.uio.no
SIRIUS Centre
University of Oslo, Norway

Dongzhuoran Zhou
dongzhuoran.zhou@de.bosch.com
Bosch Center for AI, Germany
University of Oslo, Norway

Ahmet Soylu
ahmet.soylu@oslomet.no
Department of Computer Science
Oslo Metropolitan University, Norway

Evgeny Kharlamov
evgeny.kharlamov@de.bosch.com
Bosch Center for AI, Germany
University of Oslo, Norway

ABSTRACT

With the development of Industry 4.0 technology, modern industries such as Bosch's welding monitoring witnessed the rapid widespread of machine learning (ML) based data analytical applications, which in the case of welding monitoring has led to more efficient and accurate welding monitoring quality. However, industrial ML is affected by the low transparency of ML towards non-ML experts needs. The lack of understanding by domain experts of ML methods hampers the application of ML methods in industry and the reuse of developed ML pipelines, as ML methods are often developed in an ad hoc manner for specific problems. To address these challenges, we propose the concept and a system of executable Knowledge Graph (KG), which formally encode ML knowledge and solutions in KGs, which serve as common language between ML experts and non-ML experts, thus facilitate their communication and increase the transparency of ML methods. We evaluated our system extensively with an industrial use case at Bosch, showing promising results.

Speaker's Bio: Zhuoxun Zheng is a Data Scientist and Doctoral Researcher at Bosch Center for Artificial Intelligence. He obtained his Master's degree in Mechanical Engineering at Karlsruhe Institute of Technology. Currently, he is working on knowledge enhanced machine learning and neuro-symbolic reasoning to increase transparency, trust, and performance of ML.

Relevancy to CIKM: Knowledge enhanced ML in the industry.

ACM Reference Format:
Zhuoxun Zheng, Baifan Zhou, Dongzhuoran Zhou, Ahmet Soylu, and Evgeny Kharlamov. 2022. Executable Knowledge Graph for Transparent Machine Learning in Welding Monitoring at Bosch. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22), October 17–21, 2022, Atlanta, GA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3511808.3557512

CIKM '22, October 17–21, 2022, Atlanta, GA, USA

ACM ISBN 978-1-4503-9236-5/22/10.
https://doi.org/10.1145/3511808.3557512

TALK DESCRIPTION

Motivation. In modern manufacturing industries, machine learning (ML) technologies attracts substantial yet increasing attention thanks to its strong modelling capability without the need of explicit programming . Take the Resistance Spot Welding at Bosch as an example, which is a type of fully automated and impactful manufacturing process widely applied in automotive industry, accounting for the production of millions of cars globally every year [1]. To ensure the welding quality, traditional quality monitoring approaches often require tearing the welded car bodies apart , which is extremely costly and produces much waste . In contrast, ML-based methods will help reduce the waste and contribute to more economical and sustainable manufacturing industry [7].

Interdisciplinary ML practice at Bosch. Three important activities of ML practice at Bosch include visual, statistical analytics of the data (these two are often known as exploratory data analysis and seen as important preceding steps for ML analytics), and ML analytics, which involve experts of distinct backgrounds, such as welding experts, managers etc. They work together for ML development yet speak different language. Their communication requires the transparency of ML practice (knowledge, solution, options, etc.), so that the non-ML experts can understand ML and trust that ML applied in heavy robots that operate with high electricity can ensure product quality and personnel safety. In addition, Bosch has strict regulations on documenting and reporting ML projects for later review or audit. Thus, the process of ML development, and the developed ML solutions, knowledge, and insights need to be documented properly by the experts.

Challenge. However, there exist still challenges of ML practice in industry [2]. ML projects in modern industries often involve an interdisciplinary team of experts with distinct background. The transparency of ML (C1) to non-ML experts is usually challenging, since the latter didn't receive excessive training of ML that is often required to understand the sophisticated ML methods and interpret the ML results [5]. The non-ML experts need to understand ML and trust that ML applied in manufacturing robots operating with high electricity can ensure product quality and personnel safety. In addition, in traditional ML projects, the ML procedures, methods, scripts, and decisions are described in the technical language of ML,

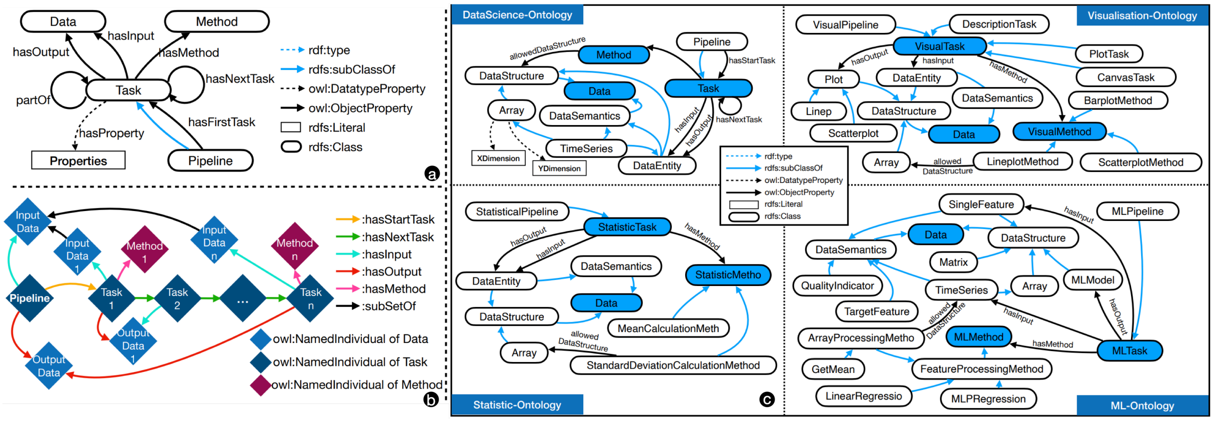

Figure 1: (a) Executable KG framework in the schema level, and (b) in the individual level, (c) KG schemata (ontologies) for the executable KG

which is highly dependent on the person who writes the document. ML knowledge and solutions are hardly described or documented in a standardised way (C2), causing difficulties for later review and retrospective comprehension of the projects in big companies like Bosch, which have strict regulations in reporting the details for later audit and analysis.

Our Contribution. To address these challenges, we propose to combine semantic technologies and ML, to encode ML solutions in knowledge graphs (KG), which is named as executable KG and helps in describing ML knowledge and solutions in a standardised way and increase the transparency via graphic user interface (GUI)-based system and visualisation of KGs [3]. In addition, executable KGs can be translated to modularised executable ML scripts that can be modified and reused for new data and new questions [4].

Our Approach. We first define the framework of the executable KGs as the left part of Fig 1, such executable KG should take the form as the right part of Fig. 1. There are three main concepts in the framework: *Data*, *Method* and *Task*. *Method* denotes a function in form of language-dependent script (such as in C++ or Python). A method takes some data which fulfils certain constraints $C_{\mathcal{F}}$ as input and can output specific data. *Task* is the process of invoking a method by feeding it with some data that meets certain constraints, and by doing so to obtain some other data. Except those *Tasks* with their *Methods* already been integrated in script, all other *Tasks* can be modularised in a *Pipeline* and be unfolded into a sequence of *Tasks*. The objectProperty *:hasFirstTask* connects the *Pipeline* with the first task in its unfolded sequence, while *:hasNextTask* connects the task in the sequence with its following task.

We then introduce briefly the semantic artefacts (ontologies) used in our approach (Fig. 1.c), which determine the conceptions (classes) of terms in executable KGs representing concrete ML data pipelines. The ontologies are created by Bosch data scientists, and are expressed in OWL 2 EL language. The data science ontology (O^{ds}) as the upper task ontology formalises the executable KG framework mentioned before, which is also the general knowledge of data science activities. Based on O^{ds}, three task ontologies, namely visual ontology, statistic ontology and ML ontology are created in such a way that all classes/properties in the task ontologies are sub-classes/sub-properties of that in O^{ds}.

Based on the ontologies, users can generated executable KGs semi-automatically through a GUI-based software system by adding information in the instance level into the ontologies [6]. In particular, the system constructs a KG that represents an executable data analytical pipeline with a series of tasks. After that, the system verifies whether the generated executable KGs fulfil the constraints. The system then executes the KGs by invoking the scripts represented by the tasks in the order defined by the pipeline.

Evaluation. We organised workshops at Bosch and collected 28 reports from experts of different background, such as welding engineers and data scientists. We have also collected some typical ML data analysis tasks in Bosch's welding manufacturing . The users were divided into two groups; each group first perform an analytical task without our method, and then answer several single selection questions (SSQ); after that, the two groups exchange their tasks and do the tasks with our method by generating an executable KG to represent the tasks, and then answer the SSQs.

The correctness of SSQ for the participants with and without the help of our approach reaches 96.7% and 93.3%, respectively, which show that the executable KG indeed helps users to better understand the ML analysis pipeline, thus increase the transparency (C1). The evaluation also demonstrates the ability of executable KGs to represent KG data pipelines in a standardised manner, thus addressing the challenges of missing standardised documentation of the ML approaches in the industry (C2).

Acknowledgements. The work was partially supported by the H2020 projects Dome 4.0 (Grant Agreement No. 953163), OntoCommons (Grant Agreement No. 958371), and DataCloud (Grant Agreement No. 101016835) and the SIRIUS Centre, Norwegian Research Council project number 237898.

REFERENCES

[1] M. Yahya, B. Zhou, et al. 2022. Towards generalized welding ontology in line with ISO and knowledge graph construction. In *ESWC*. Springer.
[2] Z. Zheng et al. 2022. Executable Knowledge Graph for Machine Learning: A Bosch Case for Welding Monitoring. In *ISWC*.
[3] Zhuoxun Zheng et al. 2022. ExeKG: Executable Knowledge Graph System for User-friendly Data Analytics. In *CIKM*.
[4] Z. Zheng, B. Zhou, and Others. 2022. Query-based industrial analytics over knowledge graphs with ontology reshaping. In *ESWC*. Springer.
[5] B. Zhou, Z. Zheng, et al. 2022. The data value quest: a holistic semantic approach at Bosch. In *ESWC*. Springer, 287–290.
[6] D. Zhou et al. 2022. Ontology Reshaping for Knowledge Graph Construction: Applied on Bosch Welding Case. In *ISWC*. Springer.
[7] D. Zhou, B. Zhou, et al. 2022. Enhancing knowledge graph generation with ontology reshaping–Bosch case. *ESWC, Springer.*