Semantic ML for Manufacturing Monitoring at Bosch

Baifan Zhou[1,2], Yulia Svetashova[1,2], Tim Pychynski[1], and Evgeny Kharlamov[3,4]

[1] Bosch CR, [2] Karlsruhe Institute of Technology, [3] Bosch Center for AI, [4] University of Oslo

Motivation. Technological advances that come with Industry 4.0, in e.g. sensoring and communication, unlock unprecedented large volumes of manufacturing data. This opens new horizons for data-driven methods like Machine Learning (ML) in analysis of manufacturing processes for a wide range of industries. An important scenario here is *monitoring* of manufacturing processes, including e.g. analysing the quality of the manufactured products and predicting the health state of machines and equipment. Consider an example of *welding quality monitoring* at Bosch, where welding is performed with automated machines that connect pieces of metal together by pressing them and passing high current electricity through them. Development of ML approaches for welding quality monitoring used in Bosch follows an iterative workflow that includes data collection (Step 1), task negotiation (Step 2), data preparation (Step 3), ML model development (Step 4), result interpretation and model selection (Step 5), model deployment (Step 6).

Challenges. Development of such ML approaches is complex and costly and at Bosch it is mostly affected by the following 3 challenges:

– *Transparency*: Steps 2 and 5 of welding quality monitoring require collaborative work of experts from different areas. The asymmetric knowledge backgrounds, including complexity of engineering practices in manufacturing and sophistication of ML algorithms that constrain the transparency of ML results and models, make the communication time consuming and error-prone.
– *Data preparation*: Step 3 requires to integrate data from dozens of sources and this is a labor-intensive effort that requires necessary understanding of multi-faceted domain knowledge and plentiful data complications.
– *Generalisability* of ML quality models: each ML model developed in Step 4 is typically tailored to a specific dataset and one process and thus its reuse for other data or processes, which is often needed, requires a significant effort.

Semantic Enhancement of ML Development. In order to address these challenges we propose to rely on semantic technologies to enhance ML pipelines that are based on feature engineering and developed a system SemML [3] that implements our ideas (see Fig. 1). The core of our approach is to incorporate domain, e.g. welding, and machine learning knowledge in the ML development in such a way that it allows us to automate data integration, ML modelling, and improve model explainability and generalisability.

In particular, we capture the domain knowledge as *ontologies* (Fig. 1) and rely on two high level ontologies: *Core* that captures high level manufacturing knowledge, e.g., of discrete manufacturing processes, and *ML* that captures ML aspects like feature groups, feature processing and ML algorithms. We also rely on a set of specific ontologies of two categories: *Domain* and *ML Pipeline* that focus on particularities of specific manufacturing processes, e.g., welding, and specific ML-pipelines for such processes. Moreover, we developed *Manufacturing* and *ML templates* that allow to encode domain knowledge (see the *Knowledge encoding* box in Fig. 1) by constructing and extending Domain and ML-pipeline ontologies in accordance to the Core and ML ontologies.

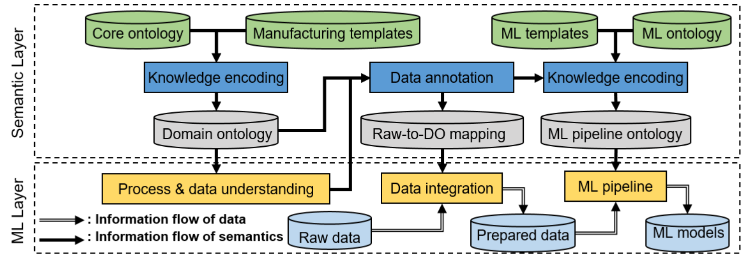

Fig. 1: Overview architecture of our approach and the SemML system.

By reasoning over ontologies we automate ML modelling. In particular, when welding experts annotate raw welding data with features from the *Domain* (see *Data annotation* box in Fig. 1), the reasoner derives from these features the relevant feature groups, and then with the help of the *ML Pipeline* it derives relevant feature processing algorithms, and relevant ML algorithms. Moreover, we use domain ontologies to annotate resulting ML models and their selected features, thus enhancing their explainablility.

Evaluation at Bosch. We implemented and evaluated our approach at Bosch for welding quality monitoring and prediction. The purpose was to enable predictive measures such as equipment (re-)configuration or maintenance and thus to prevent welding quality failures. The evaluation was done in the offline mode on two car manufacturing lines. These lines generated large volumes of heterogeneous data such as historic data with sensor measurements, welding machine configurations, manufacturing specifications, and the quality estimates of finished welding operations. The raw data came in SQL, Excel, plain text files, RUI-files and resulted in 263 fields with 53.2 million records, and 1.4 billion items after data integration. The accuracy of welding quality prediction by ML models constructed with the help of SemML is good: the mean absolute percentage error is within 1.61%–2.50% comparing to 3.19%–7.74% for simple baselines [2]. Moreover, we evaluated the usability of our system [1] with very promising results: Bosch engineers were able to construct domain and ML-pipeline ontologies relatively fast with minor training and gave high overall usability scores to our system.

Outlook. Our current results on semanticfication and thus simplification and automation of ML model development for monitoring of manufacturing processes are promising. At the same time there is still a number of important steps to be done to bring them the necessary degree of maturity and deploy them in production at Bosch by integrating them into a new welding control system and automation platform. We plan to further develop our ML pipelines with more methods, e.g. Feature Learning, ARIMA, and conduct more extensive study specifically focusing on analysis of input feature importance. We also plan to evaluate all our ML pipelines on scenarios different than welding.

References

1. Svetashova, Y., Zhou, B., Pychynski, T., Schmidt, S., Sure-Vetter, Y., Mikut, R., Kharlamov, E.: Ontology-Enhanced Machine Learning Pipeline: A Bosch Use Case of Welding Quality Monitoring. In: ISWC (2020)
2. Zhou, B., Svetashova, Y., Byeon, S., Pychynski, T., Mikut, R., Kharlamov, E.: Predicting Quality of Automated Welding with Machine Learning and Semantics: a Bosch Case Study. In: CIKM (2020)
3. Zhou, B., Svetashova, Y., Pychynski, T., Kharlamov, E.: SemFE: Facilitating ML Pipeline Development with Semantics. In: CIKM (2020)