Enterprise Modeling for Machine Learning: Case-based Analysis and Initial Framework Proposal

Dominik Bork, Panagiotis Papapetrou, Jelena Zdravkovic

To appear in:

17th International Conference on Research Challenges in Information Science (RCIS 2023)

Final version available :

https://doi.org/10.1007/978-3-031-33080-3_33

Enterprise Modeling for Machine Learning: Case-based Analysis and Initial Framework Proposal

Dominik Bork[1 [0000−0001−8259−2297]], Panagiotis Papapetrou[2 [0000−0002−4632−4815]], and Jelena Zdravkovic[2 [0000−0002−0870−0330]]

[1] Business Informatics Group, TU Wien, Vienna, Austria
dominik.bork@tuwien.ac.at
[2] Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden {jelenaz, panagiotis}@dsv.su.se

Abstract. Artificial Intelligence (AI) continuously paves its way into even the most traditional business domains. This particularly applies to data-driven AI, like machine learning (ML). Several data-driven approaches like CRISP-DM and KKD exist that help develop and engineer new ML-enhanced solutions. A new breed of approaches, often called canvas-driven or visual ideation approaches, extend the scope by a perspective on the business value an ML-enhanced solution shall enable. In this paper, we reflect on two recent ML projects. We show that the data-driven and canvas-driven approaches cover only some necessary information for developing and operating ML-enhanced solutions. Consequently, we propose to put ML into an enterprise context for which we sketch a first framework and spark the role enterprise modeling can play.

Keywords: Enterprise modeling · Conceptual modeling · Artificial intelligence · Machine learning · Model-driven engineering.

1 Introduction

Artificial Intelligence (AI) and Machine Learning (ML) are continuously paving their way into even the most traditional business domains. Many products and services nowadays entail data-driven components, often realized by an ML model. This new breed of products and services require adjustments to the development and operations practices in enterprises as existing methods are either purely (software) product-focused (like Scrum, Waterfall, and Business Model Canvas) – i.e., not incorporating the ML part – or from the opposite direction, ML focused (like CRISP-DM, KDD, and MLOps), thereby lacking focus on the enterprise context within the ML solution needs to be integrated into, e.g., the business domain, business processes, resources, etc. What is essential to consider is that such ML-based solutions do most often not run in isolation [4]. In contrast, our experience - which we will report on later- shows that ML-based solutions must be integrated into an enterprise context. Similarly to, e.g., the business and IT alignment, enterprises need to ensure that an ML solution is aligned with its enterprise context [2].

The paper at hand stresses the need to account for such an extended context when designing, implementing, monitoring, and deploying ML-enhanced solutions in an enterprise. We report on two recent cases where the authors were involved and use these cases to identify relevant context dimensions that sketch a vision of a comprehensive framework. This framework aims to put ML into an enterprise context. We equip this framework with exemplary questions that aim to engage business people in the AI/ML discussion, which is, based on the cases we present in the following, dominated by data scientists and focused on data aspects of an ML solution, ignoring to a great extent the enterprise context.

In the remainder of this paper, we first introduce state-of-the-art methods for developing and operating ML-enhanced solutions. Based on two exemplary ML project cases, we derive a multi-dimensional conceptual framework that we envision as capable of addressing challenges adhering to such ML projects from the enterprise context perspective.

2 Background

In the following, we provide an overview of existing approaches that support enterprises in realizing ML projects.

Data-driven approaches. The *CRoss Industry Standard Process for Data Mining (CRISP-DM)*, describes how the data science research process is currently implemented. The model comprises six sequential steps. The first, referred to as business understanding, aims to understand and exemplify the objectives and requirements of the project. The second step focuses on understanding the data that is available for the problem at hand, verifying data quality, exploring and visualizing the data variables, and eventually determining whether the data is appropriate for addressing the objectives and requirements. The third step concerns data processing and preparation (e.g., missing values, normalization, integration and transformation). Step four determines which algorithms should be selected and applied for solving the defined problem, deciding how to split the data into training, validation, and testing, and evaluating whether the model solves the particular task(s) at hand effectively. The fifth step focuses on the evaluation of the model with regard to meeting the defined objectives and requirements. Finally, during step six a deployment plan is defined, including monitoring and maintenance of the deployed model while in operation, final reporting of the project and the results with directions for future improvements.

The *Knowledge Discovery in Databases (KDD)* process overlaps with CRISP-DM. The first step is to develop an understanding of the underlying application domain and the existing domain knowledge, identify the available data, and set the requirements from a customer's perspective. The next step is to select a data subset that is more relevant and suitable to the problem solution. The third step is about data preprocessing, removing missing values, and dealing with data complexity. Next, the appropriate features and their representations are chosen and defined (e.g., using feature selection and dimensionality reduction). The fifth step is modeling, i.e., selecting and applying a set of appropriate ML/data mining techniques. Next, model evaluation and analysis of the findings is performed, coupled with the previous step. The appropriate parameters are tuned and the model outperforming the other candidate models in terms of chosen performance metrics is selected. Moreover, the results, extracted patterns and rules are visualized. The final step concerns the exploitation of the extracted knowledge (and model) by either integrating it into the current knowledge base or domain knowledge in general or deploying the model to a software system.

Machine learning operations (MLOps) is an approach for streamlining ML software application life cycle management with the main principles inherited from DevOps in software engineering. MLOps aims for a higher software quality, release frequency, and user customization, by integrating and automating the tasks of development and operations, and by moving between them continuously. MLOps supports continuous integration and testing of ML models. MLOps is a collaborative approach, comprising different roles such as data scientists and architects, DevOps engineers, and traditional software engineers. By following MLOps practices, these roles increase the pace and synergy of development and production, monitoring, validation, and governance of ML models. The approach also enables high scalability where a number of ML models can be managed and monitored for continuous delivery and deployment. It also provides reproducibility of ML pipelines, thereby enabling more tightly-coupled collaboration across data teams, efficiency in model regulatory scrutiny and drift-check through transparency, and compliance with organizational policies.

Canvas-driven approaches. The previously introduced data-driven approaches primarily target data scientists which is also reflected in the focus on the data-related aspects of a ML project. To account for the needs of business people (i.e., domain experts) and the characteristics of the business that aims to develop and use a new ML solution, several canvas-based approaches have been proposed recently. All of these approaches follow the general paradigm of design canvases as pioneered by the Business Movel Canvas. These canvas-driven approaches all entail a view on the data (e.g., [3]) of a ML project but also accommodate other aspects like the value that is aimed to be delivered by a ML solution [7, 9], the affected business processes [6], the heterogeneous stakeholders [4, 5], and regulative aspects [9]. These approaches are not yet matured and industry-proven compared to the data-driven ones.

3 ML in an Enterprise Context: Two Cases

Case 1: Explainable machine learning for healthcare.
This case concerns a 5-year ongoing collaborative research project between two universities, one research institute, and medical practitioners as stakeholders from two hospitals in Sweden. The first objective of the project is to develop explainable ML models and workflows that exploit the complexity of medical data sources and produce useful and insightful predictions and treatment rec-

ommendations for patients suffering from cardiovascular conditions as well as patients suffering from adverse drug effects. Secondly, the project aims to leverage explainable and responsible AI principles when used for decision making in healthcare, and demonstrate the benefits and pitfalls of AI-assisted diagnostics. Moreover, the third goal is to develop a software prototype that integrates several ML methods and provides a diagnosis and a treatment recommendation for an ongoing patient visit alongside a medically valid and trustworthy rationale behind these recommendations.

In terms of its implementation, the project follows a standard ML workflow, starting with the exploration of the available data sources and their relevance to the project objectives. More concretely a database of 1.3 million patients has been used, with one of the two hospitals being the data owner. The dataset comprises over 180 million timestamped hospital events, including diagnoses, medications, lab tests, medical procedures, and clinical text. The dataset can be readily used to train ML models, while having been de-anonymized in order to preserve the privacy and integrity of the individual patients. At a first stage, an exploratory analysis of the data variables was performed to assess data quality in terms of missingness, uncertainty, and any other potential errors. Next, the elicitation of the requirements was done by consulting with some of the stakeholders, including a small number of medical practitioners (doctors and specialists in cardiology and clinical pharmacology), patients, and lawyers.

Several ML models have been explored, such as random forests and logistic regressors, and new models have been designed, including deep learning-based architectures, such as RNNs and CNNs. The models were trained and validated using common ML model evaluation procedures, such as cross validation and repeated holdout. Standard predictive performance metrics were used, such as precision, recall, and AUC. At the same time, model explainability was assessed both qualitatively and quantitatively. Finally, and in accordance to the project's objectives, the compliance of the ML models to the existing legislative rules of Sweden was assessed and enforced during design and evaluation.

Case 2: Simulation of pandemics using machine learning. This was a 1-year collaborative project between university researchers and a public agency in Sweden. The project had two main objectives: To use reinforcement learning (RL) for policy recommendation during an ongoing pandemic (with COVID-19 as a case), and to update and improve the existing simulator used by the agency so as to include policy recommendations for contact reduction of the pandemic spread using ML. Following a similar workflow, the data sources available for this project were explored and assessed in terms of quality, validity, and relevance. Real data from the COVID-19 pandemic spread was available both at a national as well as international level. The data variables included various epidemic spread indicators, mitigation measures taken, as well as people's sentiment as quantified by Twitter posts. The epidemic data was partly public and partly owned by the agency as it concerned the epidemic spread in Sweden in particular.

Several contrasts to the first case emerged. First, policy and decision makers were strongly present in the development of both the RL-based ML models as well as during their integration to the existing software of the health agency. Hence, all constraints related to the feasibility and potential societal implications of the designed model and the contact reduction measures proposed by the model were taken into consideration. Second, methods and software had to be updated several times due to *context drift* caused by the mutation of the virus and the availability of vaccinations against the virus.

The ML methods used in this project were mostly restricted to RL techniques as they were highly suitable for the particular problem formulation. The learned policies and proposed mitigation measures were thoroughly assessed both quantitatively and qualitatively. On one hand, quantitative metrics such as model convergence, stability, and accuracy were used. On the other hand, the policies were assessed by a team of policy makers in terms of feasibility, before being integrated into the existing epidemic simulator software tool. Finally, the agency was continuously in the loop, protecting the RL agent from taking infeasible and potentially unlawful decisions. The RL model was never employed and used in practice, but it was only used for retrospective analysis while the pandemic was ongoing. In that respect, any unlawful or unreasonable recommendation would not effectively impact the population.

Lessons learned.
While the workflow followed by both cases is consistent with CRISP-DM and KDD, the two cases differ in terms of implementation. Firstly, the omission of some hospital stakeholders (e.g., nurses, specialists in other pathologies except for cardiology and clinical pharmacology) and the hospital leadership from the process resulted in several deficiencies during the development of the software prototype. On one hand, omitting nurses may neglect inaccuracies related to the content of the electronic health records, such as delayed registry of the patients' blood tests and erroneous or incomplete diagnosis codes. On the other hand, missing the hospital's leadership team resulted in inadequate information concerning software adoption and integration at the hospital.

Furthermore, the extracted rules and reasoning used by the ML models had to be aligned with the medical guidelines and the hospital decision-making processes. However, they have not been taken into full consideration, since the focus was, as is often the case, mostly on the quantitative side of model performance. These processes would have been detrimental for the design of the software prototype, as well as on the underlying mechanisms and rules that the ML models employ during training. For example, in the case of treating heart failure, the national guidelines in Sweden recommend a particular line of medication unless some other underlying condition is present. Such rules are easily integrated into the ML models, e.g., by means of constraints during model training and validation. Nonetheless, patient prioritization may differ between hospital units as they are primarily based on demand, underlying costs, or availability of specialists and personnel. Such constraints are harder to be integrated and require thorough consultation with the hospital leadership and decision-making team.

	Enterprise context dimension						
	Business	Data	Process	Stakeholders	Technology	Regulation	Legislation
Case: Issues faced in which dimensions							
Case#1:	◐	●	○	◐	◐	●	●
Case#2:	◐	●	○	◐	◐	◐	●
Data-driven approaches: Coverage of which dimensions							
CRISP-DM	●	●	○	○	◐	○	○
KDD	●	●	○	○	◐	○	○
MLOps	◐	●	○	◐	●	◐	◐
Canvas-driven approaches: Coverage of which dimensions							
Data Collection Map [3]	○	●	○	○	○	○	○
Enterprise AI Canvas [4]	◐	●	○	◐	○	○	○
Data Innovation Board [5]	◐	●	○	●	○	○	○
Machine Learning Canvas [6]	◐	●	◐	○	○	○	○
Data Science Canvas [7]	●	●	○	○	●	○	○
Prescriptive Modeling Canvas [9]	●	●	◐	◐	○	◐	○
○ = Not applicable; ◐ = Partially applicable; ● = Fully applicable							

Table 1: Mapping ML cases and related approaches to an enterprise context

With regard to the second case, the methods and software used experience context drift, which had to be taken into consideration during the development and implementation. Finally, while in the first case law experts were part of the project both as researchers and stakeholders, in this case law experts were not included in the training and development of the RL agent. This implies that some recommended pandemic mitigation policies were not thoroughly assessed in terms of their actual feasibility.

4 A Framework for ML Projects in Enterprise Context

We now elicit the relevant dimensions when developing enterprise-wide ML solutions and combine them toward a vision for a framework. By analyzing the two presented cases and the primary existing ML development methods, we collected the following dimensions: *Business*, describing business value(s) for organizations and business goals to be achieved; *Data*, that is relevant for the ML models; *Process*, the workflows of business activities related to a specific concern; *Stakeholder*, people (internal or external to the enterprise) with a particular interest or role to the development or use of the ML solution; *Technology*, encompasses all technological frameworks, algorithms, and tools used in the ML solution development and operation; *Regulation*, refers to enterprise policies; and *Legislation*, concerns laws, directives, or decisions of a relevant governing body (e.g., state or municipality).

Upon a detailed analysis of the support of each case and the existing works, we have concluded their outcomes as presented in Table 1. Regarding the two cases, the results show that: the Business dimension was supported in terms of goals (objectives) for the project, while business values were not considered; both

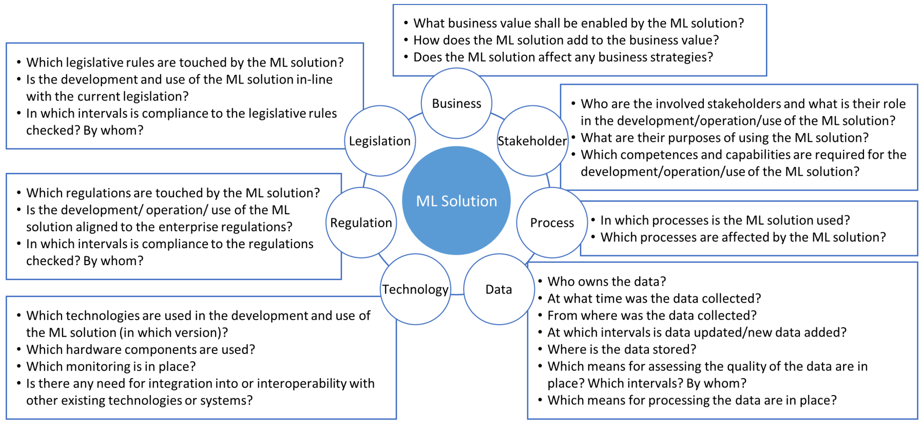

Fig. 1: Enterprise context of AI/ML solutions

cases covered the Data dimension, as well as the algorithmic part of Technology, while deployment environments where not exercised; understanding of the existing Processes was neglected which negatively influenced the quality of the ML models; the Regulation and Legislation dimensions were reasonably considered, confirming thus a maturity of these aspects about the use of ML in enterprises.

Regarding the data-driven approaches, the analysis has shown full support for the Data and Technology dimensions; Stakeholders are considered only in MLOps, yet with the focus on the development team roles; the Business is well-recognized in KDD and CRISP-DM, i.e., both values and objectives, while in MLOps only the latter are considered. The Process dimension needs to be addressed, thus showing low support for aligning the software systems and its enterprise. In contrast, the Regulation and Legislation dimensions are not considered in the first two and, to an extent, are guided for addressing in the MLOps specification. Regarding the canvas-driven approaches, a recent systematic review [11] also confirms our observations of a need for more integration of the enterprise context. The authors analyzed a total of 25 ML canvases. They concluded that many canvases focus on the Data and Technology dimensions, and a few on the Business and (partly) the Process dimensions. At the same time, more consideration of the Regulation and Legislation dimensions must be considered.

Fig. 1 shows our vision for a framework comprising seven dimensions, each equipped with exemplary questions supporting collecting relevant information about an ML project. These questions are based on the lessons learned from the two presented cases and shall aim to operationalize our framework. The aim is twofold: first and foremost, these questions shall help mitigate many of the issues faced in the two presented cases; second, they shall enable business people and domain experts to engage in the ML discussion already during the development stage and not only at the time they face issues when using the ML solution. Enterprise models can play an influential role here by providing richer specifications of individual dimensions and representing an integrated view of the different dimensions (cf. [10]). This endeavor fits nicely within the prospective future of combining enterprise modeling and AI [1, 8].

5 Conclusive Discussion

In this paper, we reflected on the lessons learned from two recent ML projects to derive a set of dimensions forming a vision for a framework that puts ML-enhanced solutions into an enterprise context. Based on the limitations of existing approaches, we spark the role enterprise modeling can play in providing a more holistic description of ML-enhanced solutions. In our future work, we aim to validate and revise our framework with more cases and to formalize the presented dimensions using well-defined requirements of a modeling method. Ultimately, we aim to propose a model-driven method for guiding the development of ML-enhanced solutions encompassing the enterprise context. Such a method might be informed by the data-driven and canvas-driven approaches and follow a structured process model.

References

1. Bork, D., Ali, S.J., Roelens, B.: Conceptual modeling and artificial intelligence: A systematic mapping study. CoRR **abs/2303.06758** (2023). https://doi.org/10.48550/arXiv.2303.06758
2. Haller, K.: Managing AI in the Enterprise. Springer (2022)
3. Kayser, L., Mueller, R.M., Kronsbein, T.: Data collection map: A canvas for shared data awareness in data-driven innovation projects. In: Pre-ICIS Symposium on Inspiring mindset for Innovation with Business Analytics and Data Science (2019)
4. Kerzel, U.: Enterprise AI canvas integrating artificial intelligence into business. Appl. Artif. Intell. **35**(1), 1–12 (2021)
5. Kronsbein, T., Müller, R.M.: Data thinking: A canvas for data-driven ideation workshops. In: Bui, T. (ed.) 52nd Hawaii International Conference on System Sciences, HICSS 2019. pp. 1–10. ScholarSpace (2019)
6. Marin, I.: Data science and development team remote communication: the use of the machine learning canvas. In: Calefato, F., Tell, P., Dubey, A. (eds.) 14th Int. Conf. on Global Software Engineering. pp. 18–21 (2019)
7. Neifer, T., Lawo, D., Esau, M.: Data science canvas: Evaluation of a tool to manage data science projects. In: 54th Hawaii International Conference on System Sciences, HICSS 2021. pp. 1–10. ScholarSpace (2021)
8. Rittelmeyer, J.D., Sandkuhl, K.: Features of AI solutions and their use in AI context modeling. In: Modellierung 2022 - Workshop Proceedings. pp. 18–29. GI (2022)
9. Shteingart, H., Oostra, G., Levinkron, O., Parush, N., Shabat, G., Aronovich, D.: Machine learning prescriptive canvas for optimizing business outcomes. CoRR **abs/2206.10333** (2022). https://doi.org/10.48550/arXiv.2206.10333
10. Takeuchi, H., Ito, Y., Yamamoto, S.: Method for constructing machine learning project canvas based on enterprise architecture modeling. In: Int. Conf. on Knowledge-Based and Intelligent Information & Engineering Systems (2022)
11. Thiée, L.W.: A systematic literature review of machine learning canvases. In: für Informatik, G. (ed.) 51. Jahrestagung der Gesellschaft für Informatik. LNI, vol. P-314, pp. 1221–1235. Gesellschaft für Informatik, Bonn (2021)