Industrial Machine Learning for Enterprises (IML4E)

Jürgen Großmann[1], Jukka K. Nurminen[2]

[1] Fraunhofer FOKUS, Berlin, Germany
juergen.grossmann@fokus.fraunhofer.de
[2] University of Helsinki, Finland
jukka.k.nurminen@helsinki.fi

Abstract. Smart software solutions, i.e., software that includes artificial intelligence (AI) and machine learning (ML), have shown a great potential to automate processes that were previously not accessible to automation. These areas include predictive maintenance, the creation of clinical diagnoses, recommendations systems, speech, image and scenario recognition, automated driving etc. Since AI and ML differ from classical software development regarding fundamental activities and processes, it is currently unclear how AI and ML can be integrated into existing industrial-grade software development processes. Addressing the industrialization of ML development and operations, the IML4E project will directly address the specifics of AI and ML by providing interoperability, automation and reusability in the data and the training pipeline. Moreover, IML4E enables continuous quality assurance and supervision for different types of machine learning (supervised learning, unsupervised learning, etc.) throughout the whole life cycle of a smart software solution. In this project presentation, we will focus in particular on the quality assurance and testing research planned in the project.

Keywords: ML, ML-Testing, MLOps

1 IML4E background

Estimates show that the use of AI-based solutions for business applications will also experience significant growth in Europe over the next five years, with projected revenues worldwide rising from $14.69 billion in 2019 to an expected $126 billion in 2025. However, the high growth rates for AI-based software and services can only be achieved if AI- and ML-based software can be produced, operated and maintained with similar efficiency and quality as classic software. In analogy to classical software, AI-based software must be implemented and validated according to the requirements of the end user and fulfil the established quality characteristics of classical software as well as a number of new quality characteristics (e.g., interpretability, intelligent behavior, non-discrimination, etc.). Their use must be technologically, socially and ethically acceptable and safe. All this must be carefully planned, realized, validated and maintained throughout the software life cycle.

2 IML4E objectives

Against this background, the IML4E project[1][2] develops a European framework for the development, operation and maintenance of AI-based software, thereby ensuring the development and quality assurance of intelligent services and intelligent software on an industrial scale. The main IML4E objectives are:

- Improving the modularity and reuse of development and data artefacts throughout the development process by (a) providing datasets and metadata that may serve the training of models in different application contexts, (b) pretrained models that are reused as a basis for further training in different application contexts, and (c) test patterns and test procedures that allow for standardized test suites to ensure dedicated ML specific quality attributes like security, robustness, transparency etc.
- Boosting the automation, interoperability and tool support throughout the whole ML lifecycle. In particular, there is currently a lack of tools that allow for (a) automated processing with integrated quality assurance of data in the data preparation pipeline, (b) continuous testing and verification of ML artefacts during development, reuse and deployment, (c) versioning and traceability of development and data artefacts (data sets, models, parameters, test results) in the course of data preparation, training and (d) operations, and systematic surveillance and monitoring of models in the field (monitoring corner cases, model evolution, functional fitness, security etc.) including the ability to intervene in case severe deviations are reported.

It has become clear that a paradigm shift is imminent, especially in the area of quality assurance for AI-supported software. Classical verification approaches are not applicable due to the complexity of deep neural networks. Basic assumptions about the stability of the software and its non-determinism no longer hold, as the well-known problems with adversarial attacks and concept drift show. Quality assurance must become a continuous process that accompanies the entire life cycle of the software. Appropriate methods and tools with the necessary degree of automation as well as a focus on known ML vulnerabilities and stochastic applications are missing.

3 Expected IML4E results

The IML4E project will develop methods and tools for risk-based quality assurance over the next three years that are specifically adapted to the characteristics of deep neural networks. This includes a catalogue of formalized quality attributes dedicated to data, ML and ML-based software, tools for debugging, testing and safeguarding ML especially in safety critical areas as well as an MLOps methodology that seamlessly integrate ML and data science activities with processes and best practices from software engineering, quality assurance and safety engineering

[1] https://itea4.org/project/iml4e.html

[2] IML4E is partly funded by the Federal Ministry for Research and Education of Germany