Thesis for The Degree of Doctor of Philosophy
in Computer Science and Engineering

# Non-Functional Requirements for Machine Learning Systems


Khan Mohammad Habibullah

**Department of Computer Science and Engineering**


The thesis will be defended in public on Thusrday, 18 September, 2025, at 13:00 in Room Omega, Jupiter Building, Campus Lindholmen, Hörselgången 5, Göteborg

Faculty opponent: Research Scientist, Dr. Sallam Abualhaija, Department of Software Verification and Validation, University of Luxembourg

UNIVERSITY OF GOTHENBURG


University of Gothenburg
SE-405 30 Gothenburg, Sweden
Phone: +46 31 786 0000

## Abstract

**Background:** Machine learning (ML) systems are increasingly being deployed in complex and safety-critical domains such as autonomous driving, healthcare, and finance. ML systems learn using big data and solve a wide range of prediction and decision-making problems that would be difficult to solve with traditional systems. However, increasing use of ML in different systems has raised concerns about quality requirements, which are defined as non-functional requirements (NFRs). Many NFRs, such as fairness, transparency, explainability, and safety, are critical in ensuring the success and acceptance of ML systems. However, many NFRs for ML systems are not well understood (e.g., maintainability), some known NFRs may become more important (e.g., fairness), while some may become irrelevant in the ML context (e.g., modularity), some new NFRs may come into play (e.g., retrainability), and the scope of defining and measuring NFRs in ML systems is also a challenging task.

**Objective:** The research project focuses on addressing and managing issues related to NFRs for ML systems. The objective of the research is to identify current practices and challenges related to NFRs in an ML context, and to develop solutions to manage NFRs for ML systems.

**Method:** This research follows a design science methodology and consists of a series of empirical and design-oriented studies. First, we conducted an interview study to explore practitioners' perceptions of NFRs and the challenges associated with defining and measuring them in ML systems. Then we conducted a subsequent survey study to validate and expand these findings with broader practitioner input. To complement these studies, we conducted a partial systematic mapping study to assess the coverage of NFRs in the academic literature, revealing discrepancies between research focus and industrial needs. Additionally, we conducted group interviews with domain experts in the automotive industry to uncover requirements engineering (RE) practices and challenges specific to ML-enabled perception systems. Based on these insights, we proposed a structured, five-step quality framework and evaluated it through practitioner interviews. Finally, we proposed revised maintainability metrics adapted to the unique structure of ML systems, and we evaluated them using ten real-world open-source ML projects.

**Findings:** We found that NFRs are crucial and play an important role in the success of the ML systems. However, there is a research gap in this area, and managing NFRs for ML systems is challenging. To address the research objectives, we have identified important NFRs for ML systems, such as accuracy, reliability, fairness, transparency, retrainability, and explainability. We also identified challenges in defining, scoping, and measuring NFRs, including domain dependence, lack of standardized metrics, and difficulty in tracing NFRs across ML system components. Furthermore, we found that practitioners face significant challenges in applying RE to ML systems—particularly in autonomous perception—due to uncertainty, evolving components, and lack of systematic approaches for managing quality trade-offs, data quality, and cross-organizational collaboration. To address these gaps, we proposed a five-step NFR management framework, covering NFR selection, scoping, trade-off analysis, measurement planning, and structured specification using templates. Finally, given that maintainability is an important NFR for ML systems, we proposed scope-aware definitions and measurement strategies for maintainability in ML systems and demonstrated their usefulness through empirical evaluation.

**Conclusion:** NFRs are critical for ML systems, but they are difficult to define, allocate, specify, and measure due to challenges like unintended bias, non-deterministic behavior, and the high cost of thorough testing. Industry and research lack well-structured solutions to manage NFRs for ML systems effectively. This research addresses this critical gap by providing a comprehensive understanding of NFRs and the unique challenges they pose in the ML context. Through a combination of empirical studies and the development of a structured NFR management framework, this research offers a solution for identifying, prioritizing, scoping, measuring, and specifying NFRs across granular-level components of ML systems. Contributions also include scope-aware definitions and measurement metrics of maintainability for ML systems. These findings enrich the theoretical understanding of NFRs for ML systems, provide empirically grounded insights into their challenges, and introduce artifacts and metrics to support future research. These outcomes also provide valuable guidance for practitioners to build trustworthy, maintainable, and high-quality ML systems. This research will help practitioners make better engineering decisions, improve quality assurance processes, and provide a foundation for more systematic and accountable ML system development.

**Keywords**
Non-functional Requirements, NFRs, Machine Learning, Quality Requirements, Requirements Engineering

ISBN: 978-91-8115-395-8 (PRINT)
ISBN: 978-91-8115-396-5 (PDF)