METeOR-SPM

Metamorphic Testing of Reliability and Robustness in Sequential Prediction Models

Programm / Ausschreibung	KS 24/26, Bridge 2025, BRIDGE 2025/01	**Status**	laufend
Projektstart	01.01.2026	**Projektende**	31.12.2028
Zeitraum	2026 - 2028	**Projektlaufzeit**	36 Monate
Barwert der Projektförderung	€ 359.982		
Keywords	Metamorphic Testing; Combinatorial Testing; Differential Testing; Sequential Prediction Models; Trustworthy AI		

Projektbeschreibung

Künstliche Intelligenz (KI) findet zunehmend Einsatz in kritischen Bereichen wie Logistik, Fertigung und Energieprognosen. Dennoch bleibt ihre Nutzung in sicherheitskritischen Szenarien eingeschränkt, da Fehlvorhersagen schwerwiegende wirtschaftliche Schäden oder Risiken für Menschenleben verursachen können. Bestehende Evaluationsverfahren, die vor allem auf aggregierten Kennzahlen wie Accuracy oder F1-Scores beruhen, erfassen systematische Fehlermodi nicht zuverlässig. Damit entsteht eine zentrale Lücke, die eine breitere und vertrauenswürdige Nutzung von KI verhindert. Dies schafft einen dringenden Bedarf an rigorosen Softwaretestmethoden, um die Zuverlässigkeit von KI-Systemen sicherzustellen.

Das Projekt adressiert diese Herausforderung durch die Entwicklung neuer Testmethoden für Sequential Prediction Models (SPMs), wie LSTMs oder Transformer-Architekturen. Unser Ansatz kombiniert und erweitert modernste Testverfahren – metamorphes Testen, kombinatorisches Testen und differentielles Testen. Um das KI-spezifische Oracle-Problem zu überwinden, integrieren wir diese Methoden mit fortgeschrittenen statistischen Analysen, die es ermöglichen, verborgene und emergente Fehlermodi zu erkennen, die durch traditionelle Evaluationsverfahren nicht erfasst werden können. Dieser integrierte Ansatz erlaubt es, nicht nur isolierte Fehler, sondern auch systematische und wiederkehrende multifaktorielle Ausfälle aufzudecken und so eine umfassendere und robustere Bewertung der Modellzuverlässigkeit über die Zeit zu unterstützen.

Durch die Bereitstellung neuer Testmethoden wird das Projekt die Zuverlässigkeit und Interpretierbarkeit von KI-Systemen stärken und so die Grundlage für eine breitere und sicherere Einführung in der Industrie schaffen. Auf diese Weise leistet es einen Beitrag sowohl für die Softwaretest-Forschungscommunity als auch zu Europas strategischem Ziel, vertrauenswürdige und qualitativ hochwertige KI-Systeme zu entwickeln.

Abstract

Artificial intelligence (AI) is increasingly applied in critical domains such as logistics, manufacturing, and energy forecasting. Yet, in many application areas AI systems remain inadmissible for safety reasons, since wrong predictions may entail severe economic damage or even risk human lives. Current evaluation practices, which rely largely on aggregate performance metrics such as accuracy or F1-scores, fail to uncover subtle but systematic failure modes. This gap prevents wider and trustworthy adoption of AI in real-world, safety-sensitive contexts and creates a pressing need for rigorous software testing methods to ensure the reliability of AI systems.

The proposed project addresses this challenge by advancing the testing of Sequential Prediction Models (SPMs), such as LSTMs and Transformer-based architectures. We will combine and extend state-of-the-art software testing methods—metamorphic testing, combinatorial testing, and differential (cross-version) testing. To overcome the AI-specific oracle problem, we integrate these methods with advanced statistical analyses, enabling the detection of hidden and emergent failure modes that traditional evaluation cannot capture. This integrated approach allows us to uncover not only isolated faults but also systematic and recurring multi-factor failures, supporting a more comprehensive and robust assessment of model reliability over time.

By delivering new testing methodologies, the project will strengthen the reliability and interpretability of AI systems, creating the foundation for broader and safer adoption in industry. In doing so, it contributes both to the software testing research community and to Europe’s strategic objective of developing trustworthy, high-quality AI systems.

Projektkoordinator

- Software Competence Center Hagenberg GmbH

Projektpartner

- NXAI GmbH