

## Smoke testing for machine learning: simple tests to discover severe bugs

Steffen Herbold<sup>1</sup> Tobias Haar<sup>2</sup>

**Abstract:** We summarize the article *Smoke testing for machine learning: simple tests to discover severe bugs* [HH22], which was published in *Empirical Software Engineering* in 2022.

**Keywords:** Machine learning; Classification; Software testing; Smoke testing; Combinatorial testing; Equivalence classes; Boundary-value analysis

### 1 Overview

The article “Smoke testing for machine learning: simple tests to discover severe bugs” was published in *Empirical Software Engineering* in 2022. Machine learning is nowadays a standard technique for data analysis within software applications. Software engineers need quality assurance techniques that are suitable for these new kinds of systems. Within this article, we discuss the question whether standard software testing techniques that have been part of textbooks since decades are also useful for the testing of machine learning software. Concretely, we try to determine generic and simple smoke tests that can be used to assert that basic functions can be executed without crashing.

### 2 Results

We found that we can derive such tests using techniques similar to equivalence classes and boundary value analysis. Moreover, we found that these concepts can also be applied to hyperparameters, to further improve the quality of the smoke tests. Even though our approach is almost trivial, we were able to find bugs in all three machine learning libraries that we tested and severe bugs in two of the three libraries. This demonstrates that common software testing techniques are still valid in the age of machine learning and that considerations how they can be adapted to this new context can help to find and prevent severe bugs, even in mature machine learning libraries.

---

<sup>1</sup> Universität Passau, Fakultät für Informatik und Mathematik, Dr.-Hans-Kapfner-Straße 30, 94032 Passau, Deutschland [steffen.herbold@uni-passau.de](mailto:steffen.herbold@uni-passau.de)

<sup>2</sup> Universität Göttingen, Deutschland

### **3 Data Availability**

The data and all analysis scripts are available online [He21].

#### **Literatur**

- [He21] Herbold, S.: sherbold/replication-kit-2020-smoke-testing: v1.0.0, Version v1.0.0, Dez. 2021, URL: <https://doi.org/10.5281/zenodo.5752045>.
- [HH22] Herbold, S.; Haar, T.: Smoke testing for machine learning: simple tests to discover severe bugs. *Empirical Software Engineering* 27/2, Jan. 2022, URL: <https://doi.org/10.1007/s10664-021-10073-7>.