

Submitted by **Viktor Maximilian Loreth** k12006268

Submitted at Institute of Computational Perception

Supervisor **Katharina Hoedt, PHD**

August 14, 2023

INTPRETABILITY METHODS AND THEIR EVALUATION: AN IN-DEPTH LOOK



Bachelor Thesis to obtain the academic degree of Bachelor of Science

in the Bachelor's Program

Artificial Intelligence

JOHANNES KEPLER UNIVERSITY LINZ

Altenbergerstraße 69 4040 Linz, Österreich www.jku.at DVR 0093696



Sworn Declaration

I hereby declare under oath that the submitted Bachelor Thesis has been written solely by me without any third-party assistance, information other than provided sources or aids have not been used and those used have been fully documented. Sources for literal, paraphrased and cited quotes have been accurately credited.

The submitted document here present is identical to the electronically submitted text document.

Linz, August 14, 2023

August 14, 2023 i/3



Abstract

- Motivation zum Thema [Wieso brauchen wir das Thema? Welche Rolle spielt es in der Welt?]
- Welche Rolle spielt AI in der Welt? Welche Rolle spielen Interpretability Methoden in der Welt? Wieso wird es erforscht?
- Worüber geht es in der Thesis [Interpreatbility of AI models,]
- Die Thesis fasst zusammen, was sind Interpretability Methoden, was für Arten gibt es, wie werden Sie angewandt?

August 14, 2023 ii/3



Contents

1	Introduction	1
2	Fundamental theory 2.0.1 abc	2
3	Project work	3

August 14, 2023 iii/3



List of Figures

August 14, 2023 iv/3



1 Introduction

- Worüber geht es in der Thesis [Interpreatbility of AI models,]
- Die Thesis fasst zusammen, was sind Interpretability Methoden, was für Arten gibt es, wie werden Sie angewandt?

Struktur erklären [Was wird in dieser Bachelorarbeit besprochen?]

August 14, 2023 1/3



2 Fundamental theory

2.1 abc

2.1.1 abc

August 14, 2023 2/3



3 Project work

August 14, 2023 3/3