

author1un=0,uniquepart=base,hash=ba648cebf10e32fb9ca84ffd4c9dca3bfamily=Bock,  
familyi=B., given=Peter, giveni=P., givenun=0 sortinitT aut-  
hor1un=0,uniquepart=base,hash=de6539ea1dd3e30d69e6ae978858b2e3family=Vaughn  
Koen, familyi=V. K., given=Peter, giveni=P., givenun=0, suffix=Billy & Bock,  
suffixi=B.  
bibinitperiod B., suffixun=0



**Hochschule Harz**

Hochschule für angewandte Wissenschaften

# METHODIK EINES INGENIEURS

Vorgelegt von:

**Philipp Thüring**

m26665

Friedrichstraße 12

38855 Wernigerode

Erstprüfer	Johr	Alexander
Zweitprüfer	Singer	Jürgen
Abgabedatum	28. Februar	2021

# Inhaltsverzeichnis

<b>Hinweise zur Abgabe</b>	<b>4</b>
<b>1 Was ist ein Problem?</b>	<b>6</b>
1.1 Veränderung . . . . .	6
1.1.1 Am Besten . . . . .	6
1.1.2 Ungewissheit . . . . .	7
1.2 Der Kern des Ingenieursproblems . . . . .	7
<b>2 Analyse</b>	<b>8</b>
2.1 Beschreibung des Problems . . . . .	8
2.2 Performance Kriterien festlegen . . . . .	9
2.3 Themenverwandte Arbeiten untersuchen . . . . .	10
2.4 Ziel formulieren . . . . .	11
<b>3 Synthese</b>	<b>12</b>
3.1 Lösungsansätze implementieren . . . . .	12
3.1.1 Experimente Designen . . . . .	13
3.1.2 Experimente durchführen . . . . .	15
3.1.3 Ergebnisse extrahieren . . . . .	15
<b>Anhang</b>	<b>19</b>
A Nachbereitung . . . . .	19
<b>Eidesstattliche Erklärung</b>	<b>22</b>

# **1 Einleitung**

## **1.1 Motivation und Problemstellung**

## **1.2 Zielsetzung und Forschungsfragen**

## **1.3 Wissenschaftlicher Beitrag und Forschungslücke**

## **1.4 Aufbau der Arbeit**

## **2 Theoretische Grundlagen und Stand der Forschung**

### **2.1 Natural Language Processing für Verkehrsinformationen**

#### **2.1.1 Grundlagen der Sprachverarbeitung**

#### **2.1.2 NLP-Anwendungen im Verkehrssektor**

#### **2.1.3 Domänenspezifische Sprachmodelle**

### **2.2 Sprachmodelle und Fine-Tuning**

#### **2.2.1 Fine-Tuning-Methoden**

#### **2.2.2 Prompt Engineering**

### **2.3 Ressourceneffizienz und Modelloptimierung**

#### **2.3.1 Problematik großer Sprachmodelle**

#### **2.3.2 Quantisierung**

#### **2.3.3 Weitere Optimierungsansätze**

### **2.4 Qualitätssicherung bei KI-generierten Texten**

#### **2.4.1 Evaluationsmetriken für NLP**

#### **2.4.2 Halluzination Detection und Validierung**

## **3 Anwendungskontext und Anforderungen**

### **3.1 Leipziger Verkehrsbetriebe**

### **3.2 Analyse der Verkehrsanweisungen**

### **3.3 Anforderungen an das System**

#### **3.3.1 Funktionale Anforderungen**

#### **3.3.2 Nicht-funktionale Anforderungen**

## 4 Modellauswahl: leo-mistral-hessianai-7b

### 4.1 Begründung der Wahl

### 4.2 Modellvarianten

## 5 Datensatzerstellung und -aufbereitung

### 5.1 Datenerhebung

### 5.2 Datensatzentwicklung für Fine-Tuning

#### 5.2.1 Annotationsstrategie

#### 5.2.2 Datenaugmentierung

#### 5.2.3 Datensatzstruktur

#### 5.2.4 Herausforderungen bei kleinem Datensatz



## **6 Implementierung**

### **6.1 Technische Umsetzung**

### **6.2 Fine-Tuning des Modells**

#### **6.2.1 LoRA-basiertes Fine-Tuning**

#### **6.2.2 Prompt Engineering und Automatisierung**

#### **6.2.3 Quantisierung**

### **6.3 Qualitätssicherung und Validierung**

## **7 Evaluation**

### **7.1 Evaluationsmethodik**

### **7.2 Qualitätsevaluation**

#### **7.2.1 Vollmodell vs. Quantisierte Versionen**

#### **7.2.2 Automatisierung und Konsistenz**

### **7.3 Gesamtbewertung**

## 8 Diskussion

### 8.1 Interpretation der Ergebnisse

### 8.2 Kritische Bewertung

#### 8.2.1 Limitationen der Arbeit

#### 8.2.2 Validität der Ergebnisse

## 9 Zusammenfassung und Ausblick

### 9.1 Zusammenfassung der Arbeit

### 9.2 Beantwortung der Forschungsfragen

### 9.3 Wissenschaftlicher und praktischer Beitrag

### 9.4 Ausblick und zukünftige Forschung

#### 9.4.1 Praktische Implikationen

#### 9.4.2 Erweiterungsmöglichkeiten

#### 9.4.3 Offene Forschungsfragen

## Abbildungsverzeichnis

1	Problemerkennung . . . . .	8
2	Performance Kriterien . . . . .	9
3	Arten von Medien . . . . .	11
A.1	Hochschullogo . . . . .	21

## Listingverzeichnis

1	Loesungsimplementation Codebeispiel Java . . . . .	13
2	Anzeigemöglichkeit Experimente via HTML . . . . .	15
3	SQL Befehl, zur Extraktion der Ergebnisse . . . . .	16
A.1	Java Hello World Beispiel . . . . .	20

# Anhang

## A Nachbereitung

Als allererstes, hier noch die letzte Referenz zum folgenden Listing im Anhang: A.1 Die hier benutzte Arbeit entstand im Rahmen des Moduls *Wissenschaftliches Arbeiten* aus dem 4. Semester. Da keine Bilder oder Grafiken vorgesehen waren, habe ich ein paar eher weniger relevante eingefügt. Trotzdem hoffe ich, dass der Leseflow der Arbeit einigermaßen angenehm ist. Alle für diese Abgabe relevanten Inhalte sind gleichmäßig über die Arbeit verteilt und können über die Seite 'Hinweise' angewählt werden. Überflüssige Stellen der Arbeit wurden gelöscht, um den Inhalt bündig beisammen zu halten und nicht zu dünn zu verstreuen. Deshalb könnte der Text bei genauerem lesen allerdings auch teilweise nur wenig Sinn ergeben.

---

```

5
6      /*
7      Java Hello World example.
8      */
9
10     public class HelloWorldExample{
11
12         public static void main(String args[]){
13
14             /*
15             Use System.out.println() to print on console.
16             */
17             System.out.println("Hello World !");
18
19         }
20
21     }
22
23     /*
24     OUTPUT of the above given Java Hello World Example would be :
25     Hello World !
26     */
27

```

---

Listing A.1: Every java dev says 'Hello world!', but noone ever says 'How are you doing, world?'...

Hier noch die letzte Anforderung an das Dokument in Form eines Bildes von unserem Hochschullogo im Anhang.

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam



et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

# ▲ Hochschule Harz

## Hochschule für angewandte Wissenschaften

Abbildung A.1: Das Logo unserer Hochschule

## Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst und dabei keine anderen als die angegebenen Hilfsmittel benutzt habe. Sämtliche Stellen der Arbeit, die im Wortlaut oder dem Sinn nach Publikationen oder Vorträgen anderer Autoren entnommen sind, habe ich als solche kenntlich gemacht. Die Arbeit wurde bisher weder gesamt noch in Teilen einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Wernigerode, den 18.08.1999



---

Philipp Thüring