Data Analytics  
Master Digitale Prozesse und Technologien  
Hochschule für Technik Stuttgart

Projektbericht

**[Titel]**

[ggf. Untertitel]

Vorgelegt von:

Antonino Forte   
Matrikelnummer: XXXX, Fachsemester: 2

Bruno Vidal dos Santos  
Matrikelnummer: XXXX, Fachsemester: 1

Vinko Jelić

Matrikelnummer: XXXX, Fachsemester: 2

Abgabedatum: 09.01.2024

Modul: Data Analytics

Semester: Wintersemester 2023/24

Dozenten: Prof. Dr. Sebastian Speiser, Prof. Dr. Ulrike Pado

Inhaltsverzeichnis

[Abbildungsverzeichnis I](#_Toc155450665)

[Tabellenverzeichnis II](#_Toc155450666)

[1. Motivation 1](#_Toc155450667)

[2. Problemstellung 2](#_Toc155450668)

[3. Lösungsansatz 3](#_Toc155450669)

[4. Haupterkenntnisse aus den Daten 4](#_Toc155450670)

[5. Haupterkenntnisse aus der Umsetzung inklusive Schwierigkeiten 5](#_Toc155450671)

[6. Fazit und Ausblick 6](#_Toc155450672)

[Literaturverzeichnis IV](#_Toc155450673)

[Anhang V](#_Toc155450674)

[Eidesstattliche Erklärung VI](#_Toc155450675)

# Abbildungsverzeichnis

# Tabellenverzeichnis

# 1. Motivation

Für unser Projekt liegt die Motivation in der immer weiter fortschreitenden Entwicklung der Datenanalyse und des maschinellen Lernens im Kontext von „Big Data“. Dabei gerät der Trend zur Entwicklung einer künstlichen Intelligenz durch diese Ansätze in den Hintergrund.

Statt das menschliche Denken zu simulieren, sollen spezifische Lösungen für klar definierte Probleme geschaffen werden. Setzt man Algorithmen des maschinellen Lernens ein, können entsprechende Muster aus historischen Daten erkannt und verarbeitet werden.

In der Regel greifen die Methoden des maschinellen Lernens auf objektive und umfassende Daten („Big Data“). Dadurch werden genauere und effizientere Vorhersagen ermöglicht, da die Algorithmen nach Mustern suchen, die am ehesten zur jeweiligen Situation passen.

Dementsprechend lassen sich drei Schlüsselvoraussetzungen aufstellen:

* Datenverfügbarkeit (z.B. durch IoT-Sensoren)
* Rechen- und Speicherkapazitäten
* Verfügbare Algorithmen und Tools (Python inkl. Bibliotheken)

Unser Projekt verfolgt die Zielsetzung, das Potenzial der Datenanalyse und des maschinellen Lernens auszunutzen und tiefe Einblicke in das komplexe Problemfeld zu wahren. Zum Einsatz kommen das Tool Jupyter Notebook sowie die Programmiersprache Python, unterstützt durch leistungsstarke Bibliotheken, wie „pandas“, „numpy“, „scikit-learn“ und weitere.

# 2. Problemstellung

Unser Fokus im Projekt liegt auf der prädiktiven Analyse der Schwere („Severity“) von Verkehrsunfällen in den USA. Als Grundlage nutzen wir den Datensatz „US Accidents“ (vgl. US Accidents (2016 - 2023) 2024) von Kaggle. Somit konzentrieren wir uns auf die Entwicklung von Modellen zur Vorhersage der Unfallschwere und dem zeitgleichen Einblick in die Faktoren, die zu den schwerwiegenderen Unfällen führen.

Die Hauptaspekte unserer Problemstellung werden durch folgende Punkte untermauert:

1. Zielvariable – Severity: Die Vorhersage der Unfallschwere dient als zentrales Ziel der Analyse. Die verschiedenen Kategorien der Schwere (eins bis vier) werden betrachtet, um einordnen zu können, welche Faktoren oder Zusammenhänge mit schweren Unfällen in Verbindung stehen.
2. Analyse: Die Analyse der Faktoren auf die Unfallschwere anhand der verfügbaren Daten, wie Witterungsbedingungen, Verkehrsbedingungen, Straßenzustand, geografische Koordinaten und weiterer relevanter Daten.
3. Modell: Mit Hilfe der ausgewählten Tools sollen prädiktive Modelle erstellt werden. Dabei kommen z.B. Methoden zur Klassifikation zum Einsatz, um die Unfallschwere mit Hilfe der ermittelten und relevanten Faktoren vorherzusagen.
4. Interpretation: Aus den gewonnenen Erkenntnissen aus den Modellen sollen Interpretationen abgeleitet werden. Dabei spielen z.B. mögliche Maßnahmen zur Verbesserung der Verkehrssicherheit oder Infrastruktur eine Rolle. Dazu gehören mögliche Erweiterungen oder Umverteilungen von Ressourcen in Prozessen rund um das Verkehrs- und Rettungsmanagement.

# 3. Lösungsansatz

# 4. Haupterkenntnisse aus den Daten

# 5. Erkenntnisse aus der Umsetzung inklusive Schwierigkeiten

# 6. Fazit und Ausblick

# Literaturverzeichnis

# Anhang

# Eidesstattliche Erklärung

Der Wortlaut der eidesstattlichen Erklärung wird dir meist von deinem Fachbereich vorgegeben. Wenn das nicht der Fall ist, findest du hier ein Beispiel: [Eidesstattliche Erklärung | Muster zum Download](https://www.scribbr.de/hausarbeit/eidesstattliche-erklaerung-hausarbeit/)

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe, dass ich sie zuvor an keiner anderen Hochschule und in keinem anderen Studiengang als Prüfungsleistung eingereicht habe und dass ich keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus Veröffentlichungen oder aus anderweitigen fremden Äußerungen entnommen wurden, sind als solche kenntlich gemacht.

\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_ \_

Ort, Datum Unterschrift