

Hochschule für angewandte Wissenschaften Coburg  
Fakultät Elektrotechnik und Informatik

Studiengang: Informatik-Bachelor

Bachelorarbeit

Klassifikation von Verkehrsteilnehmern auf Basis realer Positionszeitreihen mit Verfahren des maschinellen Lernens

Lennart Köpper

Abgabe der Arbeit: 10. 09. 2023

Betreut durch:  
Prof. Dr. Thomas Wieland, Hochschule Coburg

# Abstract

# Inhaltsverzeichnis

Abstract 2

Inhaltsverzeichnis 3

Abbildungsverzeichnis 5

Tabellenverzeichnis 6

Programmcodeverzeichnis 7

Abkürzungsverzeichnis 8

1 Einleitung 9

1.1 Motivation und Projekthintergrund 9

1.2 Zielsetzung und Forschungsfragen 9

1.3 Aufbau der Arbeit 9

2 Theoretischer Hintergrund 10

2.1 Global Navigation Satellite System 10

2.2 Map-Matching 10

2.3 Maschinelles Lernen und Klassifikation 10

2.4 Klassische Klassifikationsverfahren 10

2.4.1 Klassifikator 1, z.B. Support-Vector-Machine 10

2.4.2 Klassifikator 2, z.B. Entscheidungsbäume/Random-Forest 10

2.5 Klassifikation mit Künstlichen Neuronalen Netzen 10

2.5.1 Grundlagen 11

2.5.2 Feed-Forward-Netze 11

2.5.3 Rekurrente Neuronale Netze 11

2.6 Bewertungsmaße für die Klassifikation 11

2.6.1 Accuracy 11

2.6.2 Precision 11

2.6.3 Recall 11

2.6.4 F1-Score 11

2.7 Eingesetzte Technologien und Frameworks 11

2.7.1 Python 11

2.7.2 ScikitLearn 11

2.7.3 Keras 11

3 Verwandte Arbeiten 12

3.1 Vorangegangene Abschlussarbeiten 12

3.2 Vehicle Classification from Low-Frequency GPS Data with Recurrent Neural Networks 12

3.3 Vehicle Classification using GPS Data 12

4 Datengrundlage 13

4.1 Gewinnung der Ausgangsdaten 13

4.2 Beschreibung der Ausgangsdaten 13

4.3 Weiterverarbeitung zu Bewegungsdaten 13

5 Anforderungen und Gesamtkonzept der Klassifikation 14

6 Umsetzung des Map-Matchings 15

6.1 Einrichtung der Valhalla-Engine 15

6.2 Umsetzung und Evaluierung des Road-Snappings 15

6.3 Vorklassifikation zur Bestimmung des korrekten Matching-Modus 15

6.3.1 Erzeugung des Trainingsdatensatzes 15

6.3.2 Auswahl des Klassifikators 15

6.3.3 Evaluierung 15

7 Umsetzung der Klassifikation 16

7.1.1 Erzeugung der Trainingsdatensätze (inkl. Vorverarbeitungen) 16

7.1.2 Training und Optimierung der Modelle 16

7.1.3 Vorstellung der besten Klassifikatoren 16

8 Evaluierung 17

9 Zusammenfassung und Ausblick 18

Quellenverzeichnis 19

Anhang A 1. Test 20

Anhang A 2. Test 2 21

Ehrenwörtliche Erklärung 22

# Abbildungsverzeichnis

# Tabellenverzeichnis

# Programmcodeverzeichnis

# Abkürzungsverzeichnis

# Einleitung

## Motivation und Projekthintergrund

Brücke über das autonome Fahren (erweiterte Umfeldwahrnehmung) als mögliches Einsatzgebiet. 🡪 Forschung in diesem Gebiet auch Teil des Projektes 5GKC an welchem die Hochschule beteiligt ist.

## Zielsetzung und Forschungsfragen

Ziel: Klassifikation von Verkehrsteilnehmern auf Basis ihrer sequenziell bereitgestellten Positionsdaten durch den Einsatz maschineller Lernverfahren.

Frage 1: Wie können reale Positionsdaten, die Ungenauigkeiten und Rauschen aufweisen, so vorverarbeitet werden, dass sie sich gut für den Einsatz maschineller Lernverfahren eignen? Frage 2: Welche Verfahren des maschinellen Lernens sind für die Klassifizierung von Verkehrsteilnehmern auf Basis von sequenziellen Positionsdaten geeignet?

## Aufbau der Arbeit

Beschreibung der Gliederung.

# Theoretischer Hintergrund

## Global Navigation Satellite System

Bildet technische Grundlage für das gesamte Vorhaben (keine Positionsdaten ohne GNSS). Gewisse Umstände, wie die nach wie vor existierenden Ungenauigkeiten in GNSS-Daten bilden Grundlage für den Einsatz von Vorverarbeitungen wie Map-Matching.

## Map-Matching

Vielversprechender Ansatz in der Datenvorverarbeitung. Damit die Idee klar wird, sollte der Ansatz theoretisch erläutert werden.

## Maschinelles Lernen und Klassifikation

Was ist maschinelles Lernen? Wo ist mein Problem darin einzuordnen? -> Klassifikationsproblem; überwachtes Lernen

## Klassische Klassifikationsverfahren

Für den Fall, dass „einfachere“ Klassifikationsverfahren im Zuge der Umsetzung und damit auch Evaluierung zum Einsatz kommen.

### Klassifikator 1, z.B. Support-Vector-Machine

### Klassifikator 2, z.B. Entscheidungsbäume/Random-Forest

## Klassifikation mit Künstlichen Neuronalen Netzen

Angestrebte Art der Klassifikation (insbesondere Hoffnungsträger bei Klassifikation motorisierter Fahrzeuge) in dieser Arbeit. Funktionsweise von KNNs ist alles andere als trivial und sollte deshalb erläutert werden.

### Grundlagen

### Feed-Forward-Netze

### Rekurrente Neuronale Netze

## Bewertungsmaße für die Klassifikation

### Accuracy

### Precision

### Recall

### F1-Score

## Eingesetzte Technologien und Frameworks

### Python

Kurze Anführung der verwendeten Python-Version und der genutzten Standardbibliotheken (insb. Pandas und numpy).

### ScikitLearn

### Keras

# Verwandte Arbeiten

## Vorangegangene Abschlussarbeiten

Kurze Beschreibung der Methodiken und Erkenntnisse der wichtigsten vorangegangenen Arbeiten, die sich ebenfalls mit der Klassifikation von Verkehrsteilnehmern auseinadergesetzt haben: Sohl, Torlak und Fischer. Aufzeigen der Grenzen und Schwächen dieser Arbeiten.

[Torlak 2022]

[Sohl 2022]

[Fischer 2023]

## Vehicle Classification from Low-Frequency GPS Data with Recurrent Neural Networks

Bildet Grundlage für meinen angestrebten Ansatz (Sequenzielle Klassifikation mit RNNs). Datenbasis ist jedoch nicht vergleichbar (GPS-Punkte mit stark variierenden und großen zeitlichen Abständen).

[Matteo Simoncini et al. 2018]

## Vehicle Classification using GPS Data

Ähnlich zur Arbeit von Sohl. Liefert gute Erkenntnisse darüber, welche Merkmale zur Klassifikation von Verkehrsteilnehmern genutzt werden können. Auch hier ist die Datenbasis (bildet lediglich Fahrten von PKWs und LKWs auf Hauptverkehrsadern ab) und bei genauerem Blick auch die Aufgabenstellung (nicht-sequenzielle Klassifikation) eine völlig verschiedene zu meiner Arbeit.

[Sun et al. 2013]

# Datengrundlage

## Gewinnung der Ausgangsdaten

Wie wurden die Daten gewonnen? -> MotionTrace

Unter welchen Umständen und unter welchen Einschränkungen wurden die Daten erzeugt?

## Beschreibung der Ausgangsdaten

Wie umfangreich ist die Datenbasis?

Wie sehen die Rohdaten aus?

Sind die Daten ausgeglichen?

Probleme in den Daten, welche beachtet werden müssen?

## Weiterverarbeitung zu Bewegungsdaten

# Anforderungen und Gesamtkonzept der Klassifikation

Teil 1: Wie genau sieht das zugrundeliegende Problem aus und welche konkreten Aufgaben/Anforderungen ergeben sich daraus?

Teil 2: Wie könnte man dieser Aufgaben Herr werden? Welche Ansätze und Ideen werden gewählt?

# Umsetzung des Map-Matchings

## Einrichtung der Valhalla-Engine

## Umsetzung und Evaluierung des Road-Snappings

## Vorklassifikation zur Bestimmung des korrekten Matching-Modus

### Erzeugung des Trainingsdatensatzes

### Auswahl des Klassifikators

### Evaluierung

# Umsetzung der Klassifikation

### Erzeugung der Trainingsdatensätze (inkl. Vorverarbeitungen)

### Training und Optimierung der Modelle

### Vorstellung der besten Klassifikatoren

Weitere Unterkapitel ggf. abhängig von schlussendlich verwendeten Verfahren.

# Evaluierung

Weitere Unterkapitel abhängig von schlussendlich verwendeten Verfahren. Dieses Kapitel wird aber auf jeden Fall einen Vergleich zwischen gematchten und ungematchten Daten bei der finalen Klassifikation enthalten.

# Zusammenfassung und Ausblick

…

# Quellenverzeichnis

Fischer, D. 2023. *Verwendung von Positionsdaten zur automatisierten Klassifizierung von Verkehrsteilnehmern mittels maschinellen Lernverfahren.* Masterarbeit.

Matteo Simoncini et al. 2018. Vehicle Classification from Low-Frequency GPS Data with Recurrent Neural Networks [online]. *Transportation Research Part C: Emerging Technologies*. doi: 10.1016/j.trc.2018.03.024.

Sohl, M. 2022. *Klassifizierung der Bewegungsmuster von Mobilfunkteilnehmern zur erweiterten Umfeldwahrnehmung autonomer Fahrzeuge.* Bachelorarbeit.

Sun, Z. und Ban, X. 2013. Vehicle classification using GPS data [online]. *Transportation Research Part C: Emerging Technologies*. doi: 10.1016/j.trc.2013.09.015.

Torlak, R. 2022. *Detektion der Bewegung von Verkehrsteilnehmern aus Positionsdaten.* Bachelorarbeit.

1. Test
2. Test 2

# Ehrenwörtliche Erklärung

Ich versichere hiermit, dass ich meine *Bachelorarbeit* mit dem Titel

*Klassifikation von Verkehrsteilnehmern auf Basis realer Positionszeitreihen mit Verfahren des maschinellen Lernens*

selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie nicht an anderer Stelle als Prüfungsarbeit vorgelegt habe.

|  |  |  |
| --- | --- | --- |
| Ort |  |  |
| Datum |  | Unterschrift |