

Bachelorarbeit

# **Automatische Analyse von Stimmerkmalen zur Vorhersage von Persönlichkeitsprofilen mittels Künstlicher Intelligenz**

*Sebastian Schmidt*

Matrikelnummer: 10053783

E-Mail: schmidt.sebastian2@fh-swf.de

Erstprüfer: Prof. Dr. Michael Rübsam

Zweitprüfer: Prof. Dr. Christian Gawron

25. Oktober 2019

## **Eigenständigkeitserklärung**

Ich erkläre, dass ich die Arbeit selbständig angefertigt und nur die angegebenen Hilfsmittel benutzt habe. Alle Stellen, die dem Wortlaut oder dem Sinn nach anderen Werken, gegebenenfalls auch elektronischen Medien, entnommen sind, sind von mir durch Angabe der Quelle als Entlehnung kenntlich gemacht. Entlehnungen aus dem Internet sind durch Angabe der Quelle und des Zugriffsdatums belegt. Weiterhin habe ich die vorliegende Arbeit an keiner anderen Stelle zur Erlangung eines Abschlusses vorgelegt.

---

Datum, Ort

---

Unterschrift

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Eingrenzung</b>	<b>1</b>
<b>3. Theoretische Grundlagen</b>	<b>1</b>
3.1. Allgemeine Grundlagen des Maschinellen Lernens	1
3.1.1. Arten von Verfahren des Maschinellen Lernens	1
3.1.2. Typische Probleme beim Maschinellen Lernen	3
3.1.3. Datenvorverarbeitung für Bilddaten	5
3.1.4. Qualitätsmaße	7
3.1.5. Testen und Validieren	7
3.2. Neuronale Netze	7
3.3. Konvolutionelle Neuronale Netze	7
<b>4. Datensätze</b>	<b>7</b>
4.1. Sign Language MNIST Dataset	7
4.2. ASL Alphabet Dataset	7
4.3. Videodatensätze	7
<b>5. Entwicklungsprozess</b>	<b>7</b>
5.1. Entwickelte Modelle	7
5.1.1. Neuronale Netze	7
5.1.2. Konvolutionelle Neuronale Netze	7
5.1.3. Transferlearning mittels VGG19	7
5.2. Backend der Webanwendung	7
5.3. Frontend der Webanwendung	7
5.4. Ergebnisse der Entwicklung	7
<b>6. Fazit</b>	<b>7</b>
<b>7. Ausblick</b>	<b>7</b>
<b>8. Quellen</b>	<b>7</b>

## Abbildungsverzeichnis

1	Overfitting in einem mehrschichtigen Neuronalen Netz . . . . .	4
---	--	---

## Listingverzeichnis

# Formelverzeichnis

## 1. Einleitung

## 2. Eingrenzung

## 3. Theoretische Grundlagen

In diesem Abschnitt sollen theoretische Grundlagen des Maschinellen Lernens vorgestellt werden. Es wird zunächst auf allgemeine Grundlagen eingegangen, um im folgenden auf die spezielleren Neuronalen Netze und Konvolutionellen Neuronale Netze einzugehen.

### 3.1. Allgemeine Grundlagen des Maschinellen Lernens

Maschinelles Lernen gibt einem Computer die Möglichkeit zu lernen ohne explizit von einem Entwickler für eine Aufgabe programmiert zu werden. Tom Mitchell beschreibt Maschinelles Lernen 1997 mit einem Computerprogramm das ohne Veränderungen des Programmcodes aus Erfahrungen  $E$  im Bezug auf eine Aufgabe  $T$  und ein Maß für die Leistung  $P$  lernt, indem seine Leistung  $P$  mit der Erfahrung  $E$  anwächst [1, S. 4]. Erfahrungen sammelt ein System, indem Daten mit speziellen Algorithmen betrachtet werden. Auf Basis dieser Daten erstellt der Algorithmus dann eine Strukturbeschreibung, welche auch als Modell bezeichnet wird [2, S. 2].

In diesem Projekt wäre die Aufgabe das Klassifizieren von Gebärdensprache. Die Leistung des Modells würde durch die korrekte Klassifizierung beschrieben werden, welche idealerweise während des Trainingsvorgangs beim Anlernen von Datensätzen zu Gebärdensprache ansteigen sollte.

#### 3.1.1. Arten von Verfahren des Maschinellen Lernens

Verfahren des Maschinellen Lernens werden in der Literatur häufig mittels verschiedener Kategorien in verschiedene Arten aufgeteilt. Auch wenn viele Arten von

Verfahren in dieser Arbeit nicht betrachtet werden, ist es wichtig diese zu benennen, um ein passendes Verfahren auszuwählen. Es werden typischerweise drei Kategorien betrachtet, die Verfahren einteilen [1, S.8-14][3, S.2].

Zunächst wird ein Verfahren danach eingeteilt, welche Informationen zu den Trainingsdaten anhand von Label nötig sind. Muss jeder Datenpunkt mit einem Label versehen sein, so handelt es sich um Überwachtes Lernen. Dem gegenüber steht das Unüberwachte Lernen, bei dem ein Algorithmus versucht Aussagen über mögliche Label aufgrund der Struktur der Daten zu treffen. Das Halbüberwachte Lernen kombiniert beide Eigenschaften und der Algorithmus versucht zunächst Label anzulernen, um im Folgenden autonom fortzufahren. Abschließend sei noch das Reinforcement Lernen zu nennen. Dieses bestraft und belohnt Aktionen auf eine vorher definierte Weise. Der Algorithmus versucht dann möglichst viele Belohnungen in möglichst wenig Zeit zu erhalten und gleichzeitig Bestrafungen zu verhindern [1, S.8-14][3, S.2].

Eine weitere Kategorie ist die Art wie ein Lernverfahren mit Daten versorgt werden muss. Beim Batch-Learning muss das Verfahren auf Basis eines kompletten Datensatzes trainiert und kann danach nicht mehr mit weiteren Daten verbessert werden. Es muss von Grund auf neu trainieren. Dem gegenüber steht das Online-Learning, welches nach und nach trainiert wird, und noch eine spätere Verfeinerung ermöglicht [1, S.14-17][3, S.2].

Mit einer weiteren Kategorie wird festgelegt wie ein Verfahren Daten verallgemeinert, um neue Aussagen treffen zu können. Mit modellbasierten Lernen wird anhand der Daten ein mathematisches Modell erstellt, welches versucht die Struktur der Daten korrekt zu beschreiben, um neue Vorhersagen treffen zu können. Instanzbasiertes Lernen vergleicht bei vorherzusagenden Daten die Merkmale mit bereits angelernten Datensätzen. Werden hier Ähnlichkeiten in den bereits erlernten Instanzen entdeckt, werden diese für eine neue Aussage genutzt [1, S.14-17][3, S.2].

Ebenfalls lassen sich Verfahren nach der Art des zu lösenden Problems einteilen. Bei der Klassifikation werden Daten anhand ihrer Label einer gewissen Klasse zugeordnet. Das Ziel ist es nun die Klasse von neuen Datensätzen korrekt vorherzusagen. Demgegenüber stehen Regressionsprobleme bei denen ein Label einen

konkreten Wert angibt, welcher bei einem neuen Datensatz vom Modell korrekt vorhergesagt werden muss [1, S.8-9][3, S.2].

Bei dem in dieser Arbeit zu lösenden Problem handelt es sich um ein typisches Klassifikationsproblem des Überwachten Lernens. Ein Modell wird auf Basis von vor-klassifizierten Daten, also typischerweise Datensätze mit Bildern von Händen die ein Zeichen in Gebärdensprache repräsentieren, angelernt. Dieses soll dann im Folgenden weitere Bilder korrekt einem Zeichen in Gebärdensprache zuweisen können. Instanzbasiertes Lernen könnte bei Bilddaten zu Problemen führen. Da Bilder in der Regel relativ viele Daten beinhalten, könnte ein Merken der Beispiele zu sehr großen Modellen führen [1, S.18]. Aus diesem Grund sollte modellbasiertes Lernen vorgezogen werden. Es ist allerdings sowohl Batch-Learning als auch Online-Learning denkbar.

### 3.1.2. Typische Probleme beim Maschinellen Lernen

Beim Maschinellen Lernen können viele Probleme auftreten, die den Lernerfolg eines Modells behindern können. Um diesen entgegenwirken zu können, sollten typische Probleme vor dem Entwickeln eines Modells betrachtet werden.

Schon bei einfachen Problemen ist eine ausreichende Datenmengen nötig, um Modelle antrainieren zu können. Abhängig von der Komplexität eines Problems und des genutzten Algorithmus, können Tausende bis Millionen von Datensätzen für einen Lernvorgang nötig sein [1, S.23][3, S.4].

Wurden Daten fehlerhaft gesammelt und repräsentieren nicht das Spektrum der möglichen Eingabe im Einsatz, so kann ein Modell keine korrekten Vorhersagen für den geplanten Einsatz treffen, da hier das Wertespektrum nicht das Gelernte repräsentiert. Auch eine ungünstige Verteilung von Klassen in den Daten, kann zu Probleme führen, besonders wenn dieser Fall mit unpassenden Qualitätsmaßen ausgewertet wird [1, S.24-25][3, S.4].

Minderwertige Daten erschweren ebenfalls den Lernvorgang. Dies kann sich durch verrauschte, fehlerhafte und fehlende Daten sowie Ausreißer äußern. Eine manuelle oder automatische Vorverarbeitung könnte dies verbessern. Sogar ein Auslassen bestimmter Merkmale kann sich als sinnvoll herausstellen. [1, S.26][3, S.3]. Hilft

dies nicht ist ein Sammeln von neuen Datensätzen sinnvoll.

Irrelevante oder schlecht gewählte Merkmale für das Antrainieren eines Datensatzes führen ebenfalls zu einem schlechteren oder gar keinem Lernerfolg. Für das Vorhersagen von Gebärdensprache, sollte zum Beispiel das Datum des vorherzugesagten Bildes keinen Einfluss haben. Wird diese Information im Trainingsvorgang trotzdem einbezogen, so ist mit einer schlechteren Performanz des Lernvorgangs zu rechnen [1, S.26][3, S.4].

Overfitting tritt auf, wenn ein Modell für die Trainingsdaten zu komplex gewählt ist. Es führt dazu, dass ein Modell nicht korrekt verallgemeinert, sondern zufällige Variationen in den Trainingsdaten lernt. Es äußert sich in einer guten Performanz im Training, während diese in Tests nachlässt. Verhindern lässt sich dies im Modell durch eine Verringerung der Zahl der Parameter oder bestimmter Techniken, die dem Modell Restriktionen auferlegen, der sogenannten Regularisierung. Außerdem lässt es sich durch Sammeln neuer, sowie der Verringerung von Rauschen in vorhandenen, Daten reduzieren [1, S.27][3, S.13].

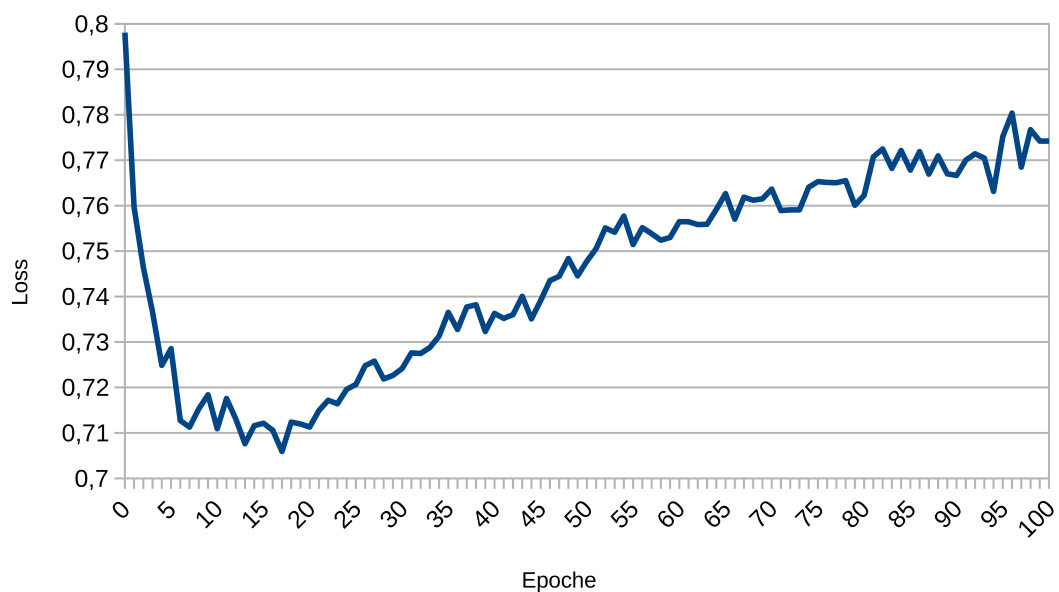


Abbildung 1: Nach der 20. Epoche, tritt in diesem mehrschichtigen Neuronalen Netz ein Overfitting auf [3, S.14].



Zuletzt sei das Underfitting genannt, welches dem Overfitting gegenübersteht. Ein Modell ist dabei zu einfach, um die Strukturen in den Trainingsdaten anzulernen. Modelle mit mehr Parametern, bessere Merkmale oder eine Verringerung der Regularisierung, reduzieren in der Regel ein Overfitting [1, S.29][3, S.14].

### 3.1.3. Datenvorverarbeitung für Bilddaten

Wie im vorherigen Kapitel dargestellt, würden unzureichende Daten zu einer schlechteren Performanz beim Anlernen eines Modells führen. Dies kann auch bei Bilddaten der Fall sein und sollte durch Vorverarbeitungsschritte verhindert werden. In dieser Sektion sollen einige typische Schritte der Bildvorverarbeitung dargestellt werden.



**3.1.4. Qualitätsmaße**

**3.1.5. Testen und Validieren**

**3.2. Neuronale Netze**

**3.3. Konvolutionelle Neuronale Netze**

**4. Datensätze**

**4.1. Sign Language MNIST Dataset**

**4.2. ASL Alphabet Dataset**

**4.3. Videodatensätze**

**5. Entwicklungsprozess**

**5.1. Entwickelte Modelle**

**5.1.1. Neuronale Netze**

**5.1.2. Konvolutionelle Neuronale Netze**

**5.1.3. Transferlearning mittels VGG19**

**5.2. Backend der Webanwendung**

**5.3. Frontend der Webanwendung**

**5.4. Ergebnisse der Entwicklung**

**6. Fazit**

**7. Ausblick**

**8. Quellen**

- [2] Josh Patterson und Adam Gibson. *Deep Learning: A Practitioner's Approach*. Sebastopol: O'Reilly, 2017. isbn: 978-1-491-91425-0.
- [3] Sebastian Schmidt. "Automatische Analyse von Stimmmerkmalen zur Vorhersage von Persönlichkeitsprofilen mittels Künstlicher Intelligenz". Iserlohn: Fachhochschule Südwestfalen, 25. Okt. 2019.