

# Machine Learning und tiefe neuronale Netze mit TensorFlow

DAVID BAUMGARTNER



BACHELORARBEIT

Nr. XXXXXXXXXXXX-A

eingereicht am  
Fachhochschul-Bachelorstudiengang

Software Engineering

in Hagenberg

im Januar 2017

Diese Arbeit entstand im Rahmen des Gegenstands

.....

im

Wintersemester 2016/17

Betreuer:

Stephan Dreiseitl, FH-Prof. PD DI Dr.

# Erklärung

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen entnommenen Stellen als solche gekennzeichnet habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt.

Hagenberg, am 14. Januar 2017

David Baumgartner

# Inhaltsverzeichnis

<b>Erklärung</b>	<b>iii</b>
<b>Vorwort</b>	<b>vi</b>
<b>Kurzfassung</b>	<b>vii</b>
<b>Abstract</b>	<b>viii</b>
<b>1 Begriffe im Maschinellen Lernen</b>	<b>1</b>
1.1 Data Science . . . . .	1
1.2 Machine Intelligence . . . . .	2
1.3 Machine Learning . . . . .	2
1.4 Neuronale Netzwerke . . . . .	2
1.5 Neuron . . . . .	3
1.6 Bias Neuron . . . . .	3
1.7 Ebenen/Layer . . . . .	4
1.8 Informationen Merken und wieder Erkennung . . . . .	4
1.9 Backpropagation . . . . .	4
1.10 Allgemeine Probleme . . . . .	5
1.10.1 Overfitting . . . . .	5
1.10.2 Datensätze . . . . .	5
1.11 Trainieren . . . . .	5
1.11.1 Supervised Trainieren . . . . .	5
1.11.2 Unsupervised Trainieren . . . . .	5
1.12 Domänenklassen . . . . .	5
1.12.1 Clustering . . . . .	6
1.12.2 Regression . . . . .	6
1.12.3 Classification . . . . .	6
1.12.4 Predict . . . . .	6
1.12.5 Robotics . . . . .	6
1.12.6 Computer Vision . . . . .	7
1.12.7 Optimization . . . . .	7
1.13 Neuronale Netzwerktypen . . . . .	7

1.13.1	FeedForward . . . . .	7
1.13.2	Self-Organizing Map . . . . .	8
1.13.3	Hopfield . . . . .	9
1.13.4	Boltzmann Machine . . . . .	9
1.13.5	Deep Belief Network . . . . .	9
1.13.6	Deep Feedforward . . . . .	9
1.13.7	NEAT . . . . .	9
1.13.8	CPPN . . . . .	9
1.13.9	HyperNEAT . . . . .	9
1.13.10	Convolutional neural network . . . . .	9
1.13.11	Elman Network . . . . .	9
1.13.12	Jordan Network . . . . .	9
1.13.13	Recurrent Network . . . . .	9
1.14	Domänen und Typen Matrix . . . . .	9
1.15	Tensorflow Typen Unterstützung . . . . .	9
<b>2</b>	<b>TensorFlow Bibliothek</b>	<b>10</b>
2.1	Python API . . . . .	11
2.2	C++ API . . . . .	11
2.3	Go API . . . . .	11
2.4	TensorFlow Python . . . . .	11
2.4.1	Graphs / Dataflowgraph . . . . .	11
2.4.2	Operation . . . . .	11
2.4.3	Tensor . . . . .	11
2.4.4	Operationen . . . . .	11
2.4.5	Probleme . . . . .	11
<b>3</b>	<b>Facial Keypoints Detection</b>	<b>12</b>
3.1	Ausgangssituation . . . . .	12
3.2	Vorbereitung . . . . .	12
3.2.1	Daten vorbereiten und normalisieren . . . . .	12
3.2.2	Evaluation- und Errorfunktion . . . . .	12
3.3	Neuronale Ebenen vorbereiten . . . . .	12
3.4	Neuronale Ebenen verknüpfen . . . . .	12
3.5	Trainieren . . . . .	12
3.6	Validierungsergebnisse . . . . .	12
	<b>Quellenverzeichnis</b>	<b>13</b>

# Vorwort

# Kurzfassung

# Abstract



# Kapitel 1

## Begriffe im Maschinellen Lernen

Diese Erklärung der Begriffe und Elemente verfolgt zwei Ziele. Zum einen stellt dies Grundlage des gesamten Themas dar und soll für Interessierte die nicht so vertraut sind, eine Einführung in die Thematik bieten. Und zum anderen werden viele dieser Begriffe erläutert, welche noch häufig zum Einsatz kommen (u.A. Neuron, Aktivierungsfunktion, ...).

### 1.1 Data Science

Data Science wird generell als die Extraktion von Wissen aus Daten bezeichnet. Dabei werden die Fachbereiche Statistik und Mathematik, Computer Science und Machine Learning sowie einige weitere, mit diesem Begriff zusammengefasst. Das Gebiet für sich wird auch als Berufstätigkeit bezeichnet, wobei meist spezialisierte Formen für die Berufsbezeichnung verwendet werden.

Damit Wissen aus Daten überhaupt extrahiert werden kann, muss ein ganzer Prozess durchlaufen werden. Dieser beginnt mit dem zusammentragen von Rohdaten aus der Realität, welche zu diesem Zeitpunkt noch keinen Zusammenhang offenbaren. Im zweiten Prozessschritt werden diese Daten meist umgebaut und neu sortiert, wobei dieser Schritt nicht immer erforderlich ist. Auf die zurecht gelegten Daten besteht nun die Möglichkeit Modelle, Algorithmen sowie weitere Extraktionen durchzuführen. Die erneut extrahierten Daten werden weiterer Folge als Ausgangsdaten verwendet. Auf diese Daten ausgeführte Modelle und Algorithmen liefern Ergebnisse die visuell dargestellt für eine größer Gruppe an Personen geeignet sind. — Aus diesem gelernten Wissen besteht zusätzlich die Möglichkeit dieses zum Generieren von neuen Daten zu verwenden und neue Modelle zu entwickeln, die zum Beispiel Vorgänge in der Natur noch akkurater widerspiegeln.

## 1.2 Machine Intelligence

Maschine Intelligence ist ein Begriff der in dieser Form noch nicht definiert worden ist. Einige namhafte Unternehmen wie Google Inc. und Microsoft Corporation bieten jeweils unterschiedliche Definitionen oder Beschreibungen. Beschreibungen, weicht aber unwesentlich von einander ab. Er wird als Überbegriff über das gesamte Gebiet mit Machine Learning, Künstlicher Intelligenz, Konversationsintelligenz und alle Themen die in näherer Beziehung dazu stehen genannt.

## 1.3 Machine Learning

Machine Learning definiert eine große Anzahl an Theorien und Umsetzungen von nicht explizit programmierten Abläufen. Diese wurden aus Studien in den Bereichen der Mustererkennung und der rechnerischen Lerntheorie mit Künstlicher Intelligenz teilweise entwickelt. Dieses Gebiet umfasst im Jahr 2016 aber sehr viel mehr. So existieren zusätzliche Ansätze aus dem Bereich der Biologie wie zum Beispiel Neuronale Netzwerke, die dem Gehirn nachempfunden sind und genetische Algorithmen die der Weiterentwicklung eines Lebewesen ähneln. Ein ganz anderer Zugang wurde in der Sowjetunion verfolgt, mit sogenannten 'Support Vektor Machines', bei welchen man einen rein mathematischen Ansatz anstrebt.

## 1.4 Neuronale Netzwerke

Neuronale Netzwerke sind im Jahr 2016 auch bekannt unter dem Begriff 'Deep Learning'.

Die Theorie und die ersten Grundlagen wurden im Jahre 1943 von Warrn McCulloch und Walter Pitts geschaffen, die ein Modell entwickelten, jedoch nicht die technischen Möglichkeiten hatten dieses umzusetzen. Dieses führte zur 'Threshold Logik'. Durch den 'Backpropagation'-Algorithmus im Jahre 1975 war es möglich, Netzwerke mit mehr als drei Ebenen zu trainieren.

Neuronale Netzwerke bestehen aus Neuronen die miteinander verbunden sind und gemeinsam ein Netzwerk ergeben, welches dem Gehirn gleicht. Die Verbindungen sind nicht fest vorgegeben, sondern können auch zum Beispiel eine Schleife bilden.

## 1.5 Neuron

Abbildung ( Aufbau eines Neurons ) - TikZ Grafik

Ein Neuron wurde einer Nervenzelle in einem Gehirn nachempfunden mit den folgenden Bestandteilen:

**Informationseingangsstrom** ist der Dateneingang, wobei ein Neuron ein bis theoretisch unendlich viele solcher Eingänge haben könnte. Dies hängt von der jeweiligen Architektur des Netzwerks ab.

**Informationsgewichtung** bezeichnet die Gewichtung mit der der Eingangsstrom gewertet wird. So wird ein Informationseingangsstrom mehr oder weniger berücksichtigt. Diese Gewichtung wird durch den Backpropagation-Algorithmus angepasst und nachjustiert.

**Kernfunktion** bewirkt das Verarbeiten der gewichteten Informationseingänge. Im einfachsten Fall werden alle Werte aufsummiert. Es wäre aber möglich, jegliche Berechnung hier einfließen zu lassen, welche mehrere Werte verwenden und daraus einen neuen Wert berechnet.

**Aktivierungsfunktion** stellt den Ausgang eines Neurons dar. Dabei wird eine weitere Funktion auf das im Kern berechnete Ergebnis ausgeführt und führt dazu, dass ein Ergebnis noch stärker ausgeprägt weitergegeben wird oder minimiert wird. Diese Aktivierungsfunktion ist meist die Sigmoid-Funktion oder eine Lineare-Funktion.

Die einfachste Repräsentation eines Neurons lässt sich mathematisch so darstellen. Im Kern wird eine Summenberechnung durchgeführt. Dabei werden die Eingangswerte und deren Gewichtung miteinander multipliziert, sowie diese Ergebnisse aufsummiert. Der griechische Buchstabe  $\phi$  (phi) steht für die Aktivierungsfunktion des Neurons und stellt damit die Ausgabe des Neurons dar.

$$f(x_i, w_i) := \phi\left(\sum_i (w_i * x_i)\right) \quad (1.1)$$

## 1.6 Bias Neuron

Ein Bias definiert einen Spezialfall eines Neurons, welches keine Dateneingänge somit auch keine Gewichtung hat und keine Berechnung im Kern durchführt. Dieses liefert nur einen konstanten Wert, wie zum Beispiel eine Eins. Durch die konstante Auslieferung wird auch die Aktivierungsfunktion überflüssig. Das Bias Neuron stellt somit einen stetig Wert für das Netzwerk dar, beziehungsweise für die darauffolgende Ebene.

## 1.7 Ebenen/Layer

Ebenen sind Zusammenschlüsse von Neuronen, welche sich auf der selben Stufe befinden. Diese Neurone sind aber nicht miteinander verbunden, sondern bekommen Daten aus der Ebene davor und geben diese an die darauffolgende Ebene weiter. Dieser Typ wird **Hiddenlayer** bezeichnet. Jedes Netzwerk benötigt zusätzlich zwei weitere Ausprägungen an Ebenen. Diese sind:

**Inputlayer** sind Ebenen die keine vorgeschaltete Ebene haben. Diese Ebene nimmt die Daten ohne Gewichtung auf und gibt sie verarbeitet an die darauffolgende Ebene weiter. In gewisser Weise stellt sie einen Übergang zwischen der Außenwelt und des Neuronalen Netzwerkes dar.

**Outputlayer** befindet sich am Ende eines Netzwerkes. Dieser Layer hat die Aufgabe, die Daten nach außen weiter zu geben anstatt an das darauffolgende Netzwerk. Hierbei werden die Informationen meist nur mehr für die Ausgabe aufbereitet. So wird zum Beispiel die Wahrscheinlichkeit für jeden Ausgang berechnet.

Abbildung ( Aufbau eines Einfachen FeedForward NN) - TikZ Grafik

## 1.8 Informationen Merken und wieder Erkennung

Durch das Anpassen der Gewichtungen bei jedem Dateneingangsstrom mit Hilfe des Backpropagation-Algorithmus ist es möglich, Zustände zu speichern und diese auch zu merken. Dies führt dazu, sollte ein ähnlicher Dateneingang stattfinden, wo zuvor schon ein ähnlicher in einem Trainingszyklus vorhanden war, dieser ähnliche erkannt wird. Dieser kann möglicherweise zu derselben Kategorie gehören, wie der zuvor schon bekannte gemachte und gelernte Dateneingang.

## 1.9 Backpropagation

TODO: relative genauer Algorithmus mit Aufgabe

## 1.10 Allgemeine Probleme

Neuronale Netzwerke sind nicht Fehler- und Problemlos und können nicht jegliche Fragen in dem Themengebiet lösen. Es existieren bekannte Probleme für die es Lösungen oder Lösungsansätze gibt. Es gibt aber auch noch

genügend andere Probleme, welche noch nicht festgestellt wurden oder als solche noch nicht erkannt worden sind.

Diese Arbeit wird auf die bekanntesten Probleme eingehen und auch Lösungen oder mögliche Lösungsansätze beinhalten.

### 1.10.1 Overfitting

Overfitting bezeichnet ein Problem welches nicht nur Machine Learning betrifft, sondern auch Menschen und andere Lebewesen. Zum Beispiel ein Student lernt auf eine Prüfung und ist im Besitz einer Klausur aus einem Vorjahr. Nach öfterem Durchspielen der Fragen und sich selbst testen, befindet er sich in der Lage diese Klausur mit einer sehr hohen Wahrscheinlichkeit zu bestehen. Dabei hat sich die Klausur in seinem Gehirn eingeprägt aber nicht das Stoffgebiet zu welchem er eine Klausur schreiben muss. Das Problem wird als Overfitting bezeichnet und beschreibt das etwas gemerkt wurde aber nicht gelernt worden ist und somit eine Abwandlung von Informationen nicht wiedererkannt werden.

### 1.10.2 Datensätze

TODO: jeder datensatz ist unterschiedlich -> neue Daten = neues Netzwerk

## 1.11 Trainieren

TODO

### 1.11.1 Supervised Trainieren

TODO

### 1.11.2 Unsupervised Trainieren

TODO

## 1.12 Domänenklassen

Neuronale Netzwerke können sehr vielseitig eingesetzt werden. Grundsätzlich lässt sich jedes Problem, welches als Funktion repräsentiert werden kann, durch ein Neuronales Netzwerk approximieren.

In dieser Arbeit werden sieben Hauptdomänen erklärt und beschrieben. Im speziellen wird auf die Neuronalen Netzwerke eingegangen, welche öfter zum Einsatz kommen und eingesetzt werden.

### 1.12.1 Clustering

Clustering Problem bezeichnet das Einordnen von Daten in Klassen oder Gruppierungen. Diese Gruppierungen können von einem Netzwerk selbst definiert werden oder manuell festgelegt werden. Es kann die Anzahl der Gruppen in welche die Daten eingeordnet werden festgelegt werden. Diese Netzwerke agieren unsupervised und besitzen meist die Möglichkeit sich nach dem Trainieren, noch selbst weiter anzupassen. Dies erlaubt dem Netzwerk auf sich ändernde Daten anzupassen und das Clustern und Gruppieren weiterzuführen ohne ein neues Netzwerk erstellen zu müssen.

### 1.12.2 Regression

TODO

### 1.12.3 Classification

Classification Problem ist in gewisser Weise das Gegenstück zum Regression Problem. Hier werden Daten dem Netzwerk übergeben, und dieses muss vorhersagen zu welcher Klasse sie gehören. Dies wird in einer supervised Umgebung durchgeführt. Die Klassen für die Vorhersage sind vorab schon bekannt und können im den Daten aus dem Outputlayer des Netzwerks verglichen werden und infolge Justierungen durchgeführt werden.

### 1.12.4 Predict

Predict Problemstellungen kommen im Kontext von Business beziehungsweise von E-Business zur Anwendung. Hier muss anhand von meist zeitgesteuerten Ereignisse eine Vorhersage getroffen werden. Zum Beispiel an der Börse ändern sich täglich die Kurse relativ rasch, sodass es für Menschen praktisch nicht mehr möglich ist, diesen zu folgen. Im Falle der Börse sind nun aber zeitlich gestützte Kurse aus der Vergangenheit verfügbar. Diese werden als Trainingsset für ein Netzwerk verwendet, um den nächsten Tag möglicherweise vorherzusagen.

### 1.12.5 Robotics

Im Bereich von Robotics entstehen häufig zu große Datenmengen. Diese Datenmengen müssten von einem Menschen analysiert werden. Daraus würde ein Algorithmus entstehen, wobei die Analyse sowie Entwicklungsphase zu viel Zeit in Anspruch nimmt. Netzwerke verwenden im Grunde die Sensorinformationen und steuern damit Motoren. Diese kann von Anfang an durchgeführt werden, sowie automatisiert.

### 1.12.6 Computer Vision

Computer Vision zielt darauf ab, einem Computer das Sehen und Verstehen von Bildern zu ermöglichen. Diese Technik findet im Jahr 2016 schon häufig Einsatz. So werden automatisiert Bilder analysiert, beschrieben sowie auch klassifiziert. Das Dienste werden auch kommerziell eingesetzt und auch angeboten. In autonom gesteuerten Fahrzeugen findet diese Anwendung auch Verwendung um Objekte zu erkennen und zu verstehen. So muss zum Beispiel ein Verkehrszeichen von einem Passanten unterschieden werden können.

### 1.12.7 Optimization

Optimization bezieht sich auf eines der Grundprobleme der Informationstechnologie. So werden immer bessere schnellere Algorithmen entwickelt, welcher konkrete Probleme noch effizienter lösen können. Durch das Thema BigData entstand ein Problem, sodass selbst sehr effiziente Algorithmen einige Problemstellungen nicht mehr in konstanter oder adäquater Zeit lösen können. Das 'Salesman' Problem gehört zu diesen Problemen. Durch Optimization wird in konstanter Zeit eine Lösung ermittelt, welche nicht die beste Lösung repräsentiert. Diese Lösung liegt aber im Rahmen von einer bestimmten definierten Toleranz und kann als Lösung verwendet werden.

## 1.13 Neuronale Netzwerktypen

In den letzten Jahren heraus haben sich diverse gut funktionierende Neuronale Netzwerktypen gebildet, beziehungsweise sind entwickelt und erforscht worden. Diese Netzwerktypen definieren Richtlinien oder Ansätze zu möglichen Netzwerken, welche aber nicht komplett übernommen werden müssen, sondern einen kreativen Spielraum ermöglichen.

In dieser Arbeit werde die wichtigsten Typen und ihre Eigenschaften erklärt. Für die folgenden Punkte ist es erforderlich ein Grundverständnis für Neuronale Netzwerk sowie über ihre Funktionsweise zu besitzen, sollte dies nicht der Fall sein - Punkt 1.3 bis 1.12.

### 1.13.1 FeedForward

FeedForward Netzwerke (FFN) waren bis für einigen Jahren noch der Stand der Forschung. Auf ihnen basieren einige andere Typen von Netzwerken, die bekanntesten werden in dieser Arbeit noch behandelt. Ein FFN passiert auf den Grundlagen eines Neurons 1.5, sowie dem ausbauen dieses zu Ebenen mit mehreren Neuronen 1.7. So ein Netzwerk besitzt einen Inputlayer, einen Hiddenlayer sowie einen Outputlayer. Sobald das Netzwerk mehrere Hiddenlayer aufweist wird es als Deep FeedForward Netzwerk 1.13.6 bezeichnet. Das FFN weist dabei eine Charakteristik auf, indem dass der Datenfluss

eindeutig definiert ist. Der Datenfluss beginnt bei dem Inputlayer und endet bei dem Outputlayer ohne dass ein Datenrückfluss zum Beispiel von dem Hiddenlayer in den Hiddelayer vorhanden ist. Dies würde eine Rekursion oder einem Kurzzeitgedächtnis entsprechen. Die einzelnen Ebenen müssen dabei aber nicht voll verbunden sein, sondern die Vernetzung kann selbst bestimmt werden.

### 1.13.2 Self-Organizing Map

Self-Organizing Map (SOM) findet vor allem im Bereich der Classification 1.12.3 Verwendung und wurden von Kohonen (1988) erfunden. Es ist nicht erforderlich einer SOM die Information zu geben, in wie viele Gruppen oder Klassen die Daten unterteilt werden sollen. Dadurch gehören es zu den Systemen, welche unsupervised trainiert werden. Außerdem besitzen sie die Möglichkeit, sich auch nach der Trainingsphase auf sich ändernde Eingangsdaten anzupassen. Kohonen entwarf die SOM mit zwei Ebenen, ein Inputlayer und ein Outputlayer ohne Hiddenlayer. Der Inputlayer propagiert Muster an den Outputlayer, wo der Dateneingang gewichtet wird. In dem Outputlayer gewinnt das Neuron, welches den geringsten Abstand zu den Eingangsdaten hat. Dies geschieht durch das Berechnen der euklidischen Distanz. Diese Art von Netzwerk kommt ohne Bias Werte 1.6 aus und es kommen ausschließlich Lineare Aktivierungsfunktionen zur Verwendung.



1.13.3 Hopfield

1.13.4 Boltzmann Machine

1.13.5 Deep Belief Network

1.13.6 Deep Feedforward

1.13.7 NEAT

1.13.8 CPPN

1.13.9 HyperNEAT

1.13.10 Convolutional neural network

1.13.11 Elman Network

1.13.12 Jordan Network

1.13.13 Recurrent Network

1.14 Domänen und Typen Matrix

1.15 Tensorflow Typen Unterstützung



## Kapitel 2

# TensorFlow Bibliothek

### 2.1 Python API

### 2.2 C++ API

### 2.3 Go API

### 2.4 TensorFlow Python

#### 2.4.1 Graphs / Dataflowgraph

#### 2.4.2 Operation

#### 2.4.3 Tensor

#### 2.4.4 Operationen

Konstanten, Zufallswerte

Variables

Transformationen

Mathematik

Flusskontrolle

Images / FFmpeg

Input und Readers

Neural Network

Running Graphs

Training

#### 2.4.5 Probleme

NaN Problem

## Kapitel 3

# Facial Keypoints Detection

### 3.1 Ausgangssituation

### 3.2 Vorbereitung

#### 3.2.1 Daten vorbereiten und normalisieren

#### 3.2.2 Evaluation- und Errorfunktion

### 3.3 Neuronale Ebenen vorbereiten

### 3.4 Neuronale Ebenen verknüpfen

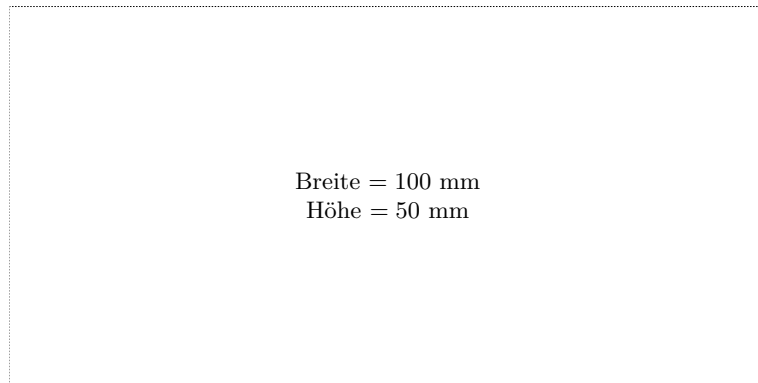
### 3.5 Trainieren

### 3.6 Validierungsergebnisse

# Quellenverzeichnis

# Messbox zur Druckkontrolle

— Druckgröße kontrollieren! —



— Diese Seite nach dem Druck entfernen! —