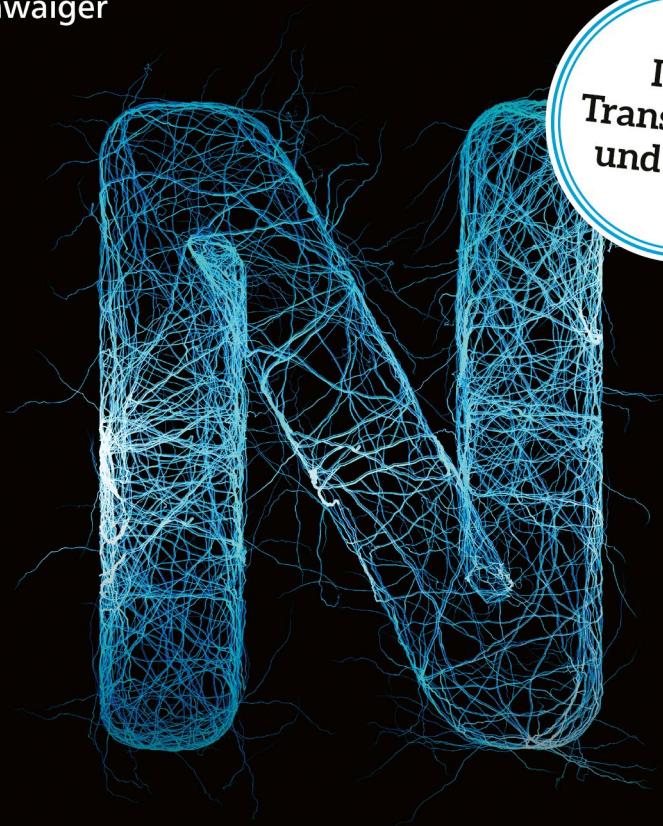


Joachim Steinwendner
Roland Schwaiger



Inkl.
Transformer
und LLMs

Neuronale Netze programmieren mit Python

- ▶ Schneller Einstieg mit allen Python- und Mathe-Grundlagen
- ▶ Lernalgorithmen, Aktivierungsfunktionen, Backpropagation u. v. m.
- ▶ Inkl. Online-Lernumgebung und Einstieg in TensorFlow



Alle Beispielprojekte zum Download



Rheinwerk
Computing

Inhalt

Vorwort	13
---------------	----

1 Einleitung 19

1.1 Wozu neuronale Netze?	19
1.2 Über dieses Buch	20
1.3 Der Inhalt kompakt	22
1.4 Ist diese Biene eine Königin?	25
1.5 Ein künstliches neuronales Netz für den Bienenstaat	26
1.6 Von der Biologie zum künstlichen Neuron	32
1.6.1 Das biologische Neuron und seine technische Kopie	32
1.6.2 Das künstliche Neuron und seine Elemente	34
1.7 Einordnung und der Rest	36
1.7.1 Big Picture	37
1.7.2 Artificial Intelligence (künstliche Intelligenz)	37
1.7.3 Geschichte	39
1.7.4 Machine Learning (maschinelles Lernen)	41
1.7.5 Deep Neural Networks	42
1.7.6 Transformer Neural Networks	43
1.8 Zusammenfassung	45
1.9 Referenzen	45

Teil I Up and running

2 Das minimale Starterkit für die Entwicklung von neuronalen Netzen mit Python 49

2.1 Die technische Entwicklungsumgebung	49
2.1.1 Die Anaconda-Distribution	49
2.1.2 Unser Cockpit: Jupyter Notebook	54
2.1.3 Wichtige Python-Module	66

2.1.4	Die Google Colab-Plattform für Jupyter Notebooks	68
2.1.5	Weitere Jupyter Notebook-Cloud-Ressourcen	71
2.2	Zusammenfassung	72

3 Ein einfaches neuronales Netz 73

3.1	Vorgeschichte	73
3.2	Her mit dem neuronalen Netz!	73
3.3	Neuron-Zoom-in	77
3.4	Stufenfunktion	82
3.5	Perceptron	84
3.6	Punkte im Raum – Vektorrepräsentation	85
3.6.1	Aufgabe: Werte vervollständigen	86
3.6.2	Aufgabe: Den Iris-Datensatz als Scatter-Plot ausgeben	89
3.7	Horizontal und vertikal – Spalten- und Zeilenschreibweise	92
3.7.1	Aufgabe: Das Skalarprodukt mithilfe von NumPy ermitteln	93
3.8	Die gewichtete Summe	95
3.9	Schritt für Schritt – Stufenfunktionen	95
3.10	Die gewichtete Summe reloaded	96
3.11	Alles zusammen	97
3.12	Aufgabe: Roboterschutz	100
3.13	Zusammenfassung	103
3.14	Referenzen	103

4 Lernen im einfachen Netz 105

4.1	Vorgeschichte: Man lässt planen	105
4.2	Lernen im Python-Code	106
4.3	Perceptron-Lernen	107
4.4	Trenngerade für einen Lernschritt	110
4.5	Perceptron-Lernalgorithmus	112
4.6	Die Trenngeraden bzw. Hyperplanes oder auch Hyperebenen für das Beispiel	117

4.7	scikit-learn-kompatibler Estimator	120
4.8	scikit-learn-Perceptron-Estimator	127
4.9	Adaline	129
4.10	Zusammenfassung	139
4.11	Referenzen	141

5 Mehrschichtige neuronale Netze 143

5.1	Ein echtes Problem	143
5.2	XOR kann man lösen	145
5.3	Vorbereitungen für den Start	150
5.4	Der Plan für die Umsetzung	152
5.5	Das Setup (»class«)	153
5.6	Die Initialisierung (»__init__«)	155
5.7	Was für zwischendurch (»print«)	158
5.8	Die Auswertung (»predict«)	159
5.9	Die Verwendung	161
5.10	Zusammenfassung	162

6 Lernen im mehrschichtigen Netz 163

6.1	Wie misst man einen Fehler?	163
6.2	Gradientenabstieg an einem Beispiel	165
6.2.1	Gradientenabstieg – die Idee	165
6.2.2	Algorithmus für den Gradientenabstieg	167
6.3	Ein Netz aus sigmoiden Neuronen	174
6.4	Der coole Algorithmus mit Vorwärts-Delta und Rückwärts-Propagation	176
6.4.1	Die __init__-Methode	176
6.4.2	Die predict-Methode	179
6.4.3	Die fit-Methode	184
6.4.4	Die plot-Methode	186
6.4.5	Alles im Konzert	186

6.5	Ein fit-Durchlauf	189
6.5.1	Initialisierung	191
6.5.2	Forward	192
6.5.3	Output	193
6.5.4	Hidden	194
6.5.5	Delta W_kj	195
6.5.6	Delta W_ji	197
6.5.7	W_ji	197
6.5.8	W_kj	197
6.6	Zusammenfassung	198
6.7	Referenz	198

7 Beispiele für tiefe neuronale Netze 199

7.1	Convolutional Neural Networks	199
7.1.1	Die Architektur von Convolutional Networks	201
7.1.2	Der Kodierungsblock	203
7.1.3	Der Prädiktionsblock	209
7.1.4	Trainieren von Convolutional Neural Networks	211
7.2	Transformer Neural Networks	216
7.2.1	Die Netzwerkstruktur	217
7.2.2	Embeddings	219
7.2.3	Positional Encoding	219
7.2.4	Encoder	223
7.2.5	Decoder	224
7.2.6	Trainieren von Transformer Neural Networks	226
7.3	Das Optimierungsverfahren	226
7.3.1	Momentum-Optimierung	227
7.3.2	Adam-Optimierung	228
7.4	Overfitting verhindern	228
7.4.1	Early Stopping	228
7.4.2	Dropout	229
7.5	Zusammenfassung	230
7.6	Referenzen	231

8 Programmierung von Deep Neural Networks mit TensorFlow 2	233
8.1 Convolutional Networks zur Handschriftenerkennung	233
8.1.1 Der Datensatz	233
8.1.2 Ein einfaches CNN	237
8.1.3 Die Ergebnisse	242
8.2 Transfer Learning mit Convolutional Neural Networks	249
8.2.1 Das vortrainierte Netzwerk	251
8.2.2 Datenvorbereitung	253
8.2.3 Das vortrainierte Netz	253
8.2.4 Die Ergebnisse	256
8.3 Transfer Learning mit Transformer Neural Networks	259
8.3.1 Die Transformer-Bibliothek	259
8.3.2 Tokenizer und Modelle	261
8.3.3 Der Model Hub von Hugging Face	262
8.4 Zusammenfassung	264
8.5 Referenzen	265

Teil II Deep Dive

9 Vom Hirn zum Netz	269
9.1 Ihr Gehirn in Aktion	270
9.2 Das Nervensystem	270
9.3 Das Gehirn	271
9.3.1 Die Teile	272
9.3.2 Ein Ausschnitt	273
9.4 Neuronen und Gliazellen	274
9.5 Eine Übertragung im Detail	276
9.6 Darstellung von Zellen und Netzen	279
9.7 Zusammenfassung	280
9.8 Referenzen	281

10 Die Evolution der künstlichen neuronalen Netze	283
10.1 Die 1940er-Jahre	284
10.1.1 1943: McCulloch-Pitts Neurons	284
10.1.2 1949: Donald Hebb	285
10.2 Die 1950er-Jahre	286
10.2.1 1951: Marvin Minsky und Dean Edmonds – SNARC	286
10.2.2 1955/1956: Artificial Intelligence	286
10.2.3 1958: Rosenblatts Perceptron	287
10.2.4 1960: Bernard Widrow und Marcian Hoff – Adaline und Madaline	287
10.3 Die 1960er-Jahre	288
10.3.1 1969: Marvin Minsky und Seymour Papert	288
10.4 Die 1970er-Jahre	288
10.4.1 1972: Kohonen – assoziativer Memory	288
10.4.2 1973: Lighthill Report	288
10.4.3 1974: Backpropagation	289
10.5 Die 1980er-Jahre	289
10.5.1 1980: Fukushimas Neocognitron	289
10.5.2 1982: John Hopfield	291
10.5.3 1982: Kohonens SOM	301
10.5.4 1986: Backpropagation	302
10.5.5 1987: NN-Konferenz	302
10.5.6 1989: Yann LeCun: Convolutional Neural Networks	302
10.6 Die 1990er-Jahre	303
10.6.1 1997 Sepp Hochreiter und Jürgen Schmidhuber – LSTM	303
10.7 Die 2000er-Jahre	303
10.7.1 2006: Geoffrey Hinton et al.	304
10.8 Die 2010er-Jahre	304
10.8.1 2014: Ian J. Goodfellow et al. – Generative Adversarial Networks (GAN)	304
10.8.2 2017: Ashish Vaswani et al. – Attention Is All You Need	306
10.9 Zusammenfassung	307
10.10 Referenzen	307

11 Der Machine-Learning-Prozess	309
11.1 Das CRISP-DM-Modell	309
11.1.1 Geschäfts(prozess)-Verständnis / Business Understanding	310
11.1.2 Datenverständnis / Data Understanding	311
11.1.3 Datenvorbereitung / Data Preparation	311
11.1.4 Modellierung / Modeling	312
11.1.5 Evaluation	312
11.1.6 Einsatz / Deployment	313
11.2 Ethische und rechtliche Aspekte	313
11.2.1 Algorithmische Fairness und Bias	314
11.2.2 Erklärbarkeit und Interpretierbarkeit	317
11.2.3 Ökologische Aspekte	321
11.2.4 Juristische Aspekte	322
11.3 Feature Engineering	324
11.3.1 Feature-Kodierung	326
11.3.2 Feature-Extraktion	337
11.3.3 Der Fluch der Dimensionalität	347
11.3.4 Feature-Transformation	347
11.3.5 Feature-Auswahl	352
11.4 Zusammenfassung	353
11.5 Referenzen	354
12 Lernverfahren	355
12.1 Lernstrategien	355
12.1.1 Überwachtes Lernen (Supervised Learning)	356
12.1.2 Unüberwachtes Lernen (Unsupervised Learning)	360
12.1.3 Verstärkendes Lernen (Reinforcement Learning)	374
12.1.4 Teilüberwachtes Lernen (Semi-supervised Learning)	383
12.2 Werkzeuge	384
12.2.1 Confusion Matrix	385
12.2.2 ROC-Curves	386
12.3 Zusammenfassung	389
12.4 Referenzen	390

13 Anwendungsbereiche und Praxisbeispiele 391

13.1 Warmup	391
13.2 Bildklassifikation	394
13.2.1 Begriffsbestimmung	394
13.2.2 Von Bienen und Hummeln	396
13.2.3 (Vor-)Trainierte Netze	407
13.3 Erträumte Bilder	415
13.3.1 Der Algorithmus	416
13.3.2 Die Implementierung	418
13.4 Deployment mit vortrainierten Netzen	426
13.4.1 Eine Webapplikation für ein neuronales Netz zur Erzeugung von Bildbeschreibungen	426
13.4.2 Eine Webapplikation für die Bilderzeugung	428
13.5 Zusammenfassung	430
13.6 Referenzen	431

Anhang 433

A Python kompakt	435
B Mathematik kompakt	465
C TensorFlow 2 und Keras	485
Index	497

Kapitel 2

Das minimale Starterkit für die Entwicklung von neuronalen Netzen mit Python

»Am Ende sieht's ein Tor, ein Klüger in der Mitte,
und nur der Weise sieht das Ziel beim ersten Schritte.«
– Friedrich Rückert (deutscher Schriftsteller, 1788–1866)

Damit der erste Schritt leichter fällt, werden wir uns in diesem Kapitel zuerst ein Arbeitsumfeld schaffen. Aus technischer Sicht ist das eine Entwicklungsumgebung für die Programmierung von neuronalen Netzen in der Programmiersprache Python.

2.1 Die technische Entwicklungsumgebung

Als Erstes installieren eine sogenannte *integrierte Entwicklungsumgebung*, die Werkzeuge zum Editieren des Quellcodes, zum Testen, zum Debuggen (d. h. zur Fehlersuche) und natürlich zum Kompilieren (d. h. zum Übersetzen in computerverständlichen Code) enthält. Im Englischen ist dafür auch die Abkürzung *IDE* gebräuchlich, die für *Integrated Development Environment* steht.

Es gibt unzählige Arten, Python auf Rechnern mit unterschiedlichen Betriebssystemen zu installieren. Auf der Website der Python Software Foundation (<https://docs.python.org/3/using/index.html>) finden Sie Installationsanweisungen für die drei wichtigsten Betriebssysteme macOS, Windows und Linux.

2.1.1 Die Anaconda-Distribution

Die Installationsanweisungen auf der Website der Python Software Foundation sind sehr lehrreich, dennoch empfehlen wir die Verwendung der Anaconda-Distribution (<https://www.anaconda.com>). Anaconda ist eine der beliebtesten Data-Science-Plattformen und bietet einige sehr hilfreiche Eigenschaften:

- ▶ eine große Anzahl an wichtigen Data-Science-Modulen, mit den für uns wichtigen Modulen NumPy, scikit-learn, TensorFlow und Keras

- ▶ den Virtual Environment Manager für Windows, Linux und macOS. Er erlaubt das Arbeiten in Entwicklungsumgebungen mit ganz spezifischen Versionen von Modulen.
- ▶ das Conda-Paket zur Installation und zum Upgraden von Modulen

Wichtig ist zu erwähnen, dass wir in diesem Buch durchgängig Python Version 3 verwenden, also sollten Sie auch die Anaconda-Version für Python 3.x herunterladen (x steht hier für eine beliebige Zahl größer gleich 0). Die detaillierten Installationsanleitungen finden Sie unter der URL <https://docs.anaconda.com/anaconda/>.

Diese Webseite enthält noch zusätzliche Informationen und Tausende von Paketen, die bei Bedarf sehr einfach installiert werden können.

Aufgabe

Bevor Sie nun mit dem Buch fortfahren, installieren Sie zuerst die Anaconda-Distribution auf Ihrem Rechner. Die Distribution ist so gestaltet, dass die Installation wenig Aufwand erfordert und selbstständig durchläuft. Laden Sie auch gleich das *Cheat Sheet* für Anaconda (<https://docs.anaconda.com/anaconda/user-guide/cheatsheet>) herunter, das eine kompakte Übersicht zur Anaconda-Distribution bietet. Danach starten Sie den Anaconda Navigator.

Anaconda Navigator

Für die Verwaltung und das Starten von Entwicklungsumgebungen bietet die Anaconda-Distribution den sogenannten *Anaconda Navigator* an (siehe Abbildung 2.1). Dieses Programm erlaubt die Administration der verschiedenen Python-Pakete, und Sie können daraus unter anderem unser Hauptwerkzeug für die Python-Programmierung, das Jupyter Notebook, starten.

Sobald der Anaconda Navigator gestartet ist, ist es möglich, unterschiedliche Applikationen unter der aktuellen Umgebung (bzw. Environment) zu starten – am Beispiel von Abbildung 2.1 wären das QT Console, Spyder, Glueviz, Orange 3, RStudio.

Deutsch oder Englisch?

Für viele Moduldistributionen, die wir hier verwenden, gibt es oft nur englische Dokumentation und Begriffe. Wir werden versuchen, diese Begriffe zu erklären, aber im Text den englischen Begriff verwenden. Wir opfern also die Schönheit der deutschen Sprache gezwungenermaßen der Eindeutigkeit. Das sei an folgendem Beispiel illustriert: Im Anaconda Navigator kann man unterschiedliche *Umgebungen* einrichten, im Navigator werden sie aber als *Environments* bezeichnet.

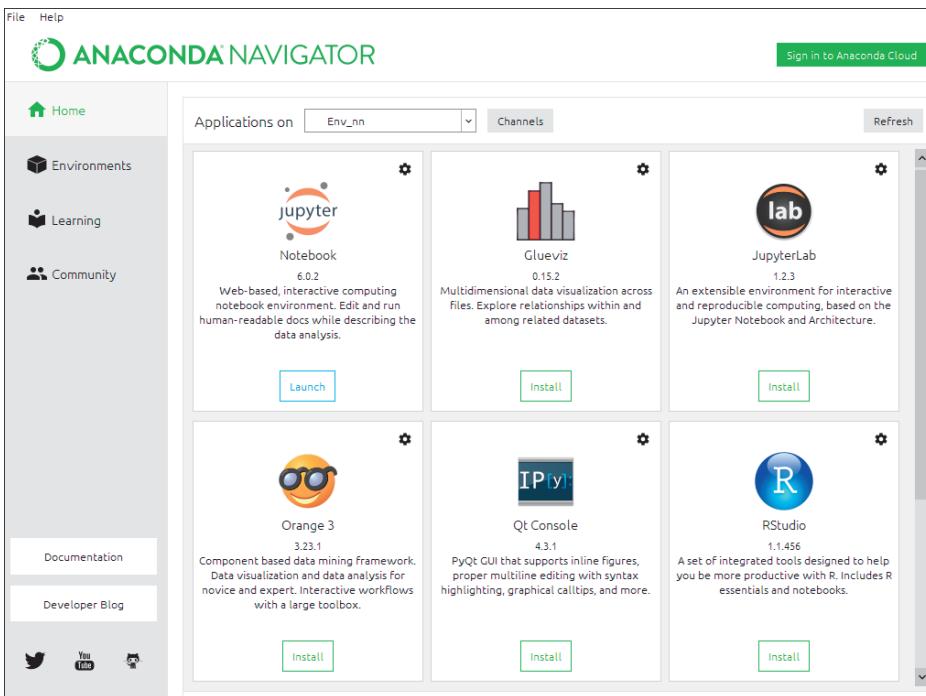


Abbildung 2.1 Der Anaconda Navigator

Das aktuelle Environment ist bei ausgewähltem HOME-Menü in der obersten Zeile angegeben. In unserem Beispiel in Abbildung 2.2 ist das die Umgebung mit der Bezeichnung ENV_NN.

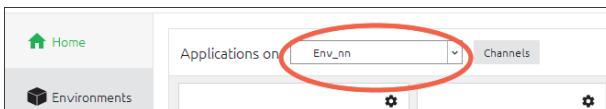


Abbildung 2.2 Die aktuelle Umgebung (Environment)

Erstellen von neuen Environments

Welchen Sinn eigene *Environments* haben, wird an folgenden Beispielen deutlich:

- ▶ Für manche Projekte brauchen Sie nur eine kleine Anzahl an Modulen. Da wäre es sinnlos, alle vorhandenen Pakete in einem Environment zu installieren.
- ▶ Die Open Source Community ist sehr schnell. Das bedeutet auch, dass unterschiedliche Versionen des gleichen Moduls existieren. Manchmal brauchen Module, die aufeinander aufbauen, aber ganz bestimmte Versionen.

- Manche Module haben geräteabhängige Versionen (zum Beispiel *TensorFlow* und *TensorFlow-gpu* für PCs mit einer oder mehreren GPUs). Auch dafür lassen sich eigene Environments einrichten.

Im Anaconda Navigator befindet sich im linken Bereich der Menüpunkt ENVIRONMENTS, der uns im mittleren Bereich eine Übersicht der Environments zeigt und im rechten Bereich eine vollständige Liste der pro Environment installierten Pakete (siehe Abbildung 2.3).

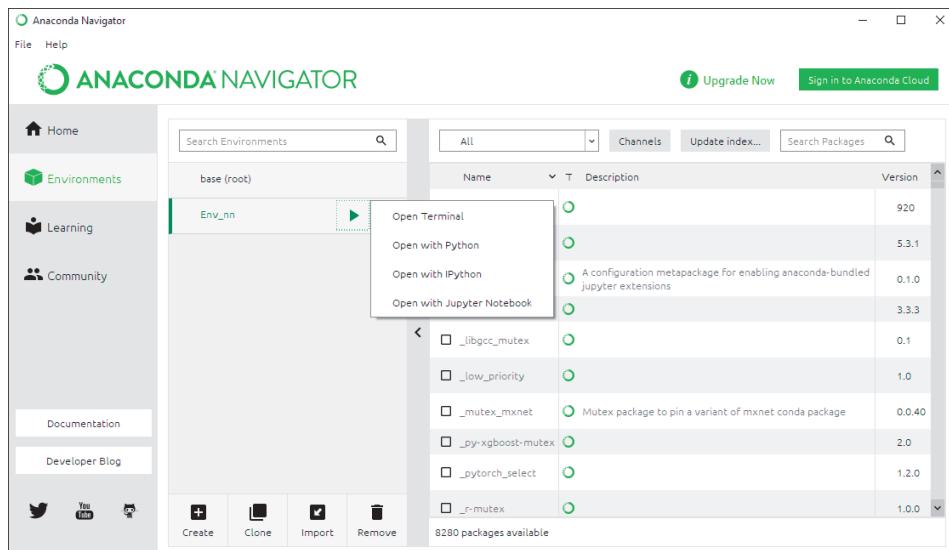


Abbildung 2.3 Das »Environment«-Fenster

Indem Sie auf den CREATE-Button klicken (mittlerer Bereich unten), können Sie ein neues Environment erstellen und einen passenden Namen vergeben.

Installation von neuen Paketen

Für die Installation von neuen Pakten in einem Environment gibt es mehrere Methoden – hier werden wir zwei Möglichkeiten zeigen. Eine richtet sich mehr an die Fans von grafischen Benutzeroberflächen (Graphical User Interface, GUI), während die andere für Verfechter der Kommandozeile (Command-Line Interface, CLI) gedacht ist.

► Installation über die grafische Benutzeroberfläche (GUI)

Zuerst müssen Sie ein Environment auswählen oder neu erstellen. Wählen Sie dazu den Menüpunkt ENVIRONMENTS im Explorer-Teil, der sich links im Anaconda Navigator befindet (siehe Abbildung 2.3). Dann klicken Sie im nun erscheinenden mittleren Teil auf ein Environment (oder erstellen ein neues). Anschließend wählen oder

suchen Sie im rechten Teil ein Paket, das installiert werden soll. Nach der Auswahl eines Pakets (oder mehrerer Pakete) erscheint rechts unten eine Schaltfläche **APPLY**. Das Anklicken dieser Schaltfläche startet die Installation.

► Installation über die Kommandozeile (CLI)

Eine weitere Installationsmöglichkeit ergibt sich über die Eingabe des folgenden Kommandos in ein *Terminal*. Das Terminal öffnen Sie, indem Sie in der ENVIRONMENT-Ansicht mit der linken Maustaste auf das Dreiecksymbol neben einem Environment-Namen klicken: Dann öffnet sich ein Untermenü, wie Sie es in Abbildung 2.3 sehen. Der erste Menüpunkt, OPEN TERMINAL, öffnet ein Kommandozeilenfenster. In der Eingabeaufforderung (die auch als *Prompt* bezeichnet wird) ist auch das aktuelle Environment zu erkennen.

Für die Installation eines Pakets (in unserem Beispiel `scipy`) ist folgender Befehl im Terminal einzugeben:

```
conda install scipy (für die neueste scipy-Version) oder  
conda install scipy=0.15.0 (für die scipy-Version 0.15.0)
```

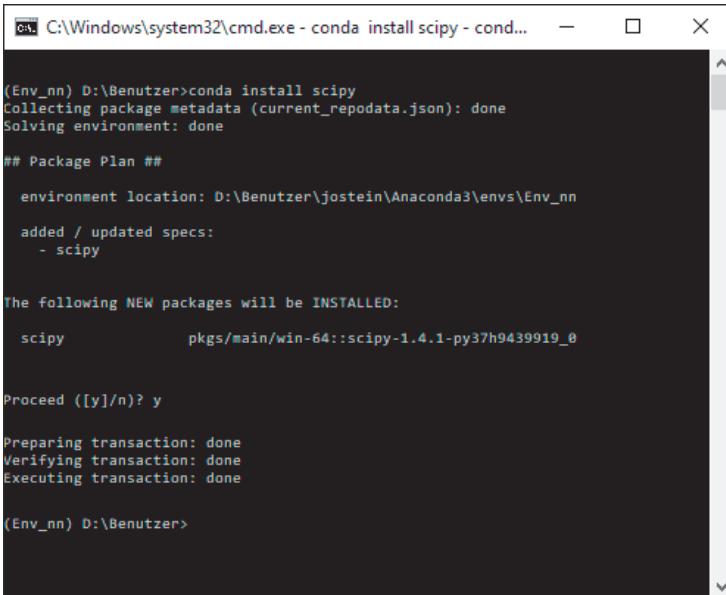
Terminal = Kommandozeilenfenster

Das Fenster für die Eingabe von Kommandozeilen wird oft auch als *Terminal* bezeichnet, so auch hier im Anaconda Navigator. Die Eingabezeile im Terminal beginnt mit einer so genannten *Eingabeaufforderung* oder *Prompt*. Diese Prompts sehen abhängig vom Betriebssystem sehr unterschiedlich aus (Windows: C:\>, Unix: benutzer@maschine:~\$). In der Kommandozeile ist ein Befehl einzugeben, der mit der Eingabetaste abgeschlossen und ausgeführt wird.

Dieser Installationsvorgang kann durchaus ein paar Minuten dauern, da zum einen Pakete aus dem Internet heruntergeladen werden und zum anderen Abhängigkeiten zu anderen Paketen bestehen können, die es erfordern, zusätzliche Pakete zu installieren. Aber keine Angst, die Abhängigkeiten sind vorgegeben, und jedes Paket weiß sozusagen selbst, was es noch braucht. Abbildung 2.4 zeigt die Kommandozeilen-Variante der Installation von `scipy`, und im unteren Drittel des Terminals sind die zusätzlich notwendigen Pakete bzw. Updates angezeigt.

Aufgabe

Erstellen Sie ein neues Environment mit dem Namen »Env_nn«. Installieren Sie das Paket `numpy` über die grafische Benutzeroberfläche (GUI) und das Paket `pandas` über das Kommandozeilenfenster (CLI).



```
(Env_nn) D:\Benutzer>conda install scipy - cond...
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

environment location: D:\Benutzer\jostein\Anaconda3\envs\Env_nn

added / updated specs:
- scipy

The following NEW packages will be INSTALLED:

scipy           pkgs/main/win-64::scipy-1.4.1-py37h9439919_0

Proceed ([y]/n)? y

Preparing transaction: done
Verifying transaction: done
Executing transaction: done

(Env_nn) D:\Benutzer>
```

Abbildung 2.4 Installation via Terminal im Environment »Env_nn«

2.1.2 Unser Cockpit: Jupyter Notebook

Ein *Jupyter Notebook* ist eine Webapplikation, die es erlaubt, Dokumente zu kreieren und mit einer Community zu teilen, und die Python-Code, Gleichungen, Visualisierungen sowie erklärenden Text enthält. Jupyter Notebooks sind also eine Art digitaler Notizblock. Neben dem Vorteil der Kombination von Dokumentation und Code können diese Notizblöcke auch leicht ausgetauscht werden. Die Notebooks werden im sogenannten JSON-Format gespeichert.

Es ist somit das ideale Werkzeug bzw. die ideale Entwicklungsumgebung, um mit neuronalen Netzen und Python-Code zu experimentieren, aber auch um ernsthafte Applikationen zu entwickeln.

Der Begriff Jupyter

Das Jupyter Notebook ist ein Ergebnis des Projekts Jupyter (<http://jupyter.org>). Der Projektname leitet sich von den drei Programmiersprachen **Julia**, **Python** und **R** ab.

Starten von Jupyter Notebook

Sie starten das Jupyter Notebook im Anaconda Navigator entweder im HOME-Bereich oder im ENVIRONMENT-Bereich:

- ▶ Im HOME-Bereich (siehe Abbildung 2.1) wählen Sie zuerst das gewünschte Environment unter APPLICATIONS ON aus und klicken dann auf das Icon von Jupyter Notebook (im Bild links oben).
- ▶ Im ENVIRONMENT-Bereich klicken Sie auf das Dreieckssymbol neben dem entsprechenden Environment und wählen dann OPEN WITH JUPYTER NOTEBOOK (siehe Abbildung 2.3).

Jupyter Notebook in der Cloud

Mit Anaconda haben Sie Jupyter Notebook lokal auf Ihrem Rechner installiert und sind damit mehr oder weniger unabhängig von einer Anbindung an das Internet. Da Jupyter Notebook jedoch eine Webapplikation ist, bietet es sich an, dieses Werkzeug in der Cloud zur Verfügung zu stellen. Das hat den Vorteil, dass Sie mehr Rechenpower einsetzen können, was vor allem beim Arbeiten mit tiefen neuronalen Netzen mit vielen Daten von Vorteil ist.

Erfreulicherweise gibt es eine Unmenge an Anbietern, die Ressourcen via Jupyter Notebook zur Verfügung stellen – allerdings sind diese ab einer gewissen Nutzungsintensität kostenpflichtig. Außerdem müssen Sie sich registrieren und oftmals auch eine Kreditkarte angeben. In Abschnitt 2.1.5, »Weitere Jupyter Notebook-Cloud-Ressourcen«, finden Sie eine (unvollständige) Liste von Anbietern.

Die Zellen eines Jupyter Notebooks

Ein Dokument, das mit Jupyter Notebook erstellt wird, besteht aus einer Liste von aufeinanderfolgenden Zellen. Diese Zellen können verschiedene Ausprägungen annehmen (siehe Abbildung 2.5). Wir interessieren uns hier nur für die zwei wichtigsten:

- 1. Markdown-Zelle:** Dieser Zelltyp erlaubt es, einen erklärenden Text einzugeben. Er ist allerdings viel mächtiger, denn wir können den Text auch formatieren, um die Schriftart zu wechseln oder Fettdruck, Formeln, Weblinks, Bildern etc. darzustellen. Markdown ist eine Auszeichnungssprache, die mit dem Ziel entwickelt wurde, dass schon die Ausgangsform leicht lesbar ist. Diese besteht aus reinem Text inklusive Formatierungszeichen. Zum Beispiel erzeugt die Eingabe ****Fettdruck**** in einer Markdown-Zelle **Fettdruck**.
- 2. Code-Zelle:** Mit Zellen dieser Art können Sie programmieren. Hier können Sie Python-Code eingeben. Natürlich können Sie auch hier Python-Kommentare eingeben, allerdings ohne Formatierungsmöglichkeit.

Beiden Zelltypen ist gemeinsam, dass ein Text eingegeben wird (entweder Text mit Formatierungszeichen oder Python-Code). Um den Inhalt dann als formatierten Text auszugeben oder den Code ablaufen zu lassen, muss die Zelle erst »ausgeführt« werden. Das

heißt, der Prozess zur Übersetzung in formatierten Text bzw. zur Durchführung der Codeanweisung muss erst angestoßen werden, entweder durch eine Menüauswahl (CELL • RUN CELLS) oder durch eine Tastenkombination ($\square + \leftarrow$).

Das Jupyter-Notebook-Dokument wird im sogenannten JSON-Format abgespeichert und hat die Dateierweiterung *.ipynb*. Das Dokument kann auf diese Weise einfach weitergegeben werden. Informationen zum JSON-Format finden sich auf der Website <http://www.json.org>. Wir gehen hier nicht näher darauf ein; gesagt sei nur so viel, dass es sich um ein schlankes Datenaustauschformat handelt, das für Menschen einfach zu lesen und zu schreiben ist und für Computer leicht zu übersetzen und zu erzeugen ist.

Markdown-Zelle

Im folgenden zeigen wir zwei identische Markdown-Zellen, einmal in Ausgangsform und einmal übersetzt.

Das ist ein **Text** mit **Formatierungszeichen**

1. nummerierte Listen beginnen mit einer Nummer
2. wobei die Nummer beliebig gewählt werden kann
3. nämlich so - beim Übersetzen wird das richtig gestellt
 1. Sublisten werden mit Einrückung (Tabulator) erstellt

Absätze werden durch eine Leerzeile erzeugt
Zeilenumbrüche durch zwei Leerzeichen am Ende einer Zeile.

Auch Weblinks können durch Rundklammern schnell eingefügt werden,
zum Beispiel der Link zur [Anaconda-Seite](#) (<https://anaconda.org>).

Das ist ein **Text** mit **Formatierungszeichen**

1. nummerierte Listen beginnen mit einer Nummer
2. wobei die Nummer beliebig gewählt werden kann
3. nämlich so - beim Übersetzen wird das richtig gestellt
 - A. Sublisten werden mit Einrückung (Tabulator) erstellt

Absätze werden durch eine Leerzeile erzeugt
Zeilenumbrüche durch zwei Leerzeichen am Ende einer Zeile.

Auch Weblinks können durch Rundklammern schnell eingefügt werden,
zum Beispiel der Link zur [Anaconda-Seite](#).

Code-Zelle

```
In [1]: a = "das ist unser erster Python-Code und wir grüßen die Welt und besonders unsere Leser"
print(a)
```

```
das ist unser erster Python-Code und wir grüßen die Welt und besonders unsere Leser
```

Abbildung 2.5 Markdown- und Code-Zellen in einem Jupyter-Notebook-Dokument

Aufgabe

Und los geht's. Starten Sie ein Jupyter Notebook, und versuchen Sie, die Inhalte von Abbildung 2.5 nachzubauen. Experimentieren Sie mit dieser Entwicklungsumgebung. Wer des Programmierens noch nicht so mächtig ist bzw. noch keine Erfahrung mit Python hat, kann sich jetzt Anhang A, »Python kompakt«, zu Gemüte führen und schon einige Codebeispiele im Jupyter Notebook ausprobieren.

Interaktives Python

Jupyter ist so konzipiert, dass man es mit verschiedenen Programmiersprachen verwenden kann, daher auch die Herkunft des Namens (siehe den Kasten »Der Begriff Jupyter«). Um andere Sprachen verwenden zu können, müsste der entsprechende Kernel installiert werden. Wir verwenden den Python-Kernel oder genauer gesagt den IPython-Kernel, wobei das »l« in IPython für »Interactive« steht. Dieser IPython-Kernel bietet uns noch erweiterte Funktionalitäten, die wir in den Notebooks zur Verfügung haben. Zumeist sind das Kommandos, die mit speziellen Zeichen, wie `%`, `?`, `!` oder auch der `Shift-Tab`-Taste aufgerufen bzw. kombiniert werden. Die Kommandos, die mit dem Präfix `%` aufgerufen werden, nennt man auch *Magic Functions*. Die Magic Function `%quickref`, die einfach in eine Code-Zelle eines Jupyter Notebooks eingegeben wird, liefert eine sehr gute Übersicht über die Interaktivitätsfunktionalitäten von Jupyter. Wir beschreiben im Folgenden einige wenige und wichtige Beispiele zum Zugriff auf die Dokumentation von Python-Funktionen bzw. zum Debuggen von Python-Code in Jupyter Notebooks.

Zugriff auf Hilfe und Dokumentation

In Bereich Data Science ist eine unglaubliche Fülle an Informationen vorhanden, die es auch einem erfahrenen Data-Science-Experten kaum möglich macht, alles im Kopf zu behalten. Viel wichtiger ist das effektive und schnelle Auffinden von Informationen, sei es im Internet, in Bedienungsanleitungen oder innerhalb der Entwicklungsumgebung.

Jupyter Notebooks bietet für den Benutzer einige einfache und schnelle Methoden, um Fragen zu beantworten, beispielsweise:

- ▶ Welche Argumente braucht dieser Funktionsaufruf?
- ▶ Wie wurde eine Funktion oder Klasse implementiert?
- ▶ Was enthält das importierte Modul an Funktionen, Methoden etc.?

Eine schnelle und nützliche Informationsquelle ist der *Docstring*, der eine knappe und präzise Zusammenfassung eines Objekts enthält. Die folgenden Codebeispiele zeigen, wie man schnell zu dieser Information kommt:

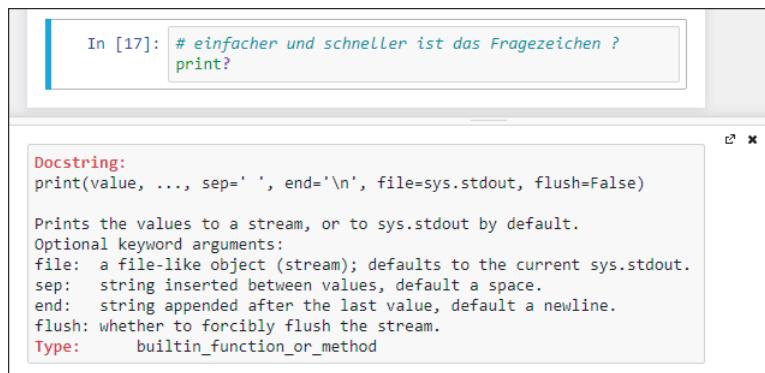
```
# Der Docstring kann mit der Funktion help ausgegeben werden
help(print)
# Ausgabe:
Help on built-in function print in module builtins:

print(...)
    print(value, ..., sep=' ', end='\n', file=sys.stdout, flush=False)

    Prints the values to a stream, or to sys.stdout by default.
    Optional keyword arguments:
        file: a file-like object (stream); defaults to the current sys.stdout.
        sep: string inserted between values, default a space.
        end: string appended after the last value, default a newline.
        flush: whether to forcibly flush the stream.
```

Listing 2.1 Docstring-Ausgabe

Anstatt die Funktion `help()` aufzurufen, können Sie aber auch ganz einfach ein Fragezeichen ans Ende des Funktionsnamens setzen. Je nach Entwicklungsumgebung wird die Beschreibung in einer Ausgabezelle angezeigt, in einem Pop-up-Fenster oder in Jupyter Notebook in einem eigenen Ausgabebereich (siehe Abbildung 2.6).

**Abbildung 2.6** Ausgabebereich bzw. Pager im Jupyter Notebook

Das Fragezeichen kann sehr vielfältig eingesetzt werden. Es kann an Funktionen angehängt werden oder aber auch an Objekte, wie in Listing 2.2 zu sehen ist. Dabei gibt es im ersten Fall Information zur Funktion `myList.append` im Docstring-Format aus und im zweiten Fall Informationen über das Objekt `myList`, wie Typ, Form oder Länge.

```
# das Fragezeichen ? kann vielfältig eingesetzt werden
myList = ['NumPy', 'scikit', 'Keras', 'TensorFlow']
myList.append?
# Ausgabe: (im Pager-Fenster - ganz unten im Webbrowser)
Signature: myList.append(object, /)
Docstring: Append object to end of the list.
Type:      builtin_function_or_method

# oder einfach, um Informationen über ein Objekt zu erhalten
myList?
# Ausgabe: (im Pager-Fenster - ganz unten im Webbrowser)
Type:      list
String form: ['NumPy', 'scikit', 'Keras', 'TensorFlow']
Length:     4
Docstring:
Built-in mutable sequence.
```

If no argument is given, the constructor creates a new empty list.
The argument must be an iterable if specified.

Listing 2.2 Das Fragezeichen

Hinweis

Da sich die Open-Source-Welt sehr schnell ändert, können die hier dargestellten Ausgaben auf Ihrem Rechner je nach installierter Python-Version etwas anders aussehen. Die Ausgaben in diesem Listing zeigen Ausgaben in der Python-Version 3.7.

Jupyter Notebook bietet natürlich auch über die grafische Benutzeroberfläche eine Möglichkeit, schnell an Informationen über Funktionen und Objekte zu kommen. Setzen Sie den Cursor auf den Funktions- oder Objektnamen und drücken Sie die Tastenkombination **Shift** + **Tab**, erscheint eine Ballonhilfe wie in Abbildung 2.7.

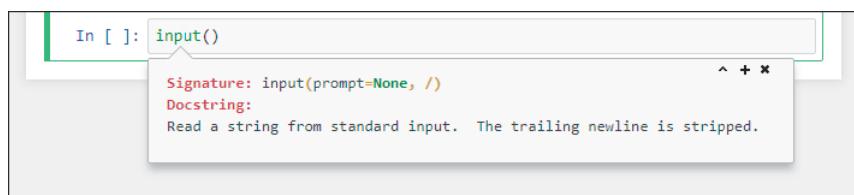


Abbildung 2.7 Kontexthilfe zur Built-in-Funktion »input()«

Aber auch die Tabulator-Taste lässt Sie Inhalte von Modulen, Klassen oder Objekten erforschen. In Abbildung 2.8 sehen Sie eine Auflistung der Funktionen des Moduls `math` bei Eingabe von `m.` in eine Code-Zelle von Jupyter.

```
In [2]: # Module erforschen mit der Tab-Taste
import math as m

In [ ]: m.
m.acos
m.acosh
m.asin
m.asinh
m.atan
m.atan2
m.atanh
m.ceil
m.copysign
m.cos
```

Abbildung 2.8 Mit der Tabulator-Taste lassen Sie sich die Attribute und Funktionen von Modulen auflisten ...

Das gilt natürlich auch für unser Listenobjekt `myList` aus Listing 2.2. In Abbildung 2.9 sehen Sie die Liste, die Sie durch Eingabe von `myList.` erhalten.

```
In [3]: # das Fragezeichen ? kann vielfältig eingesetzt werden
myList = ['numpy', 'scikit', 'keras', 'tensorflow']
myList.

In [ ]: myList.
myList.append
myList.clear
myList.copy
myList.count
myList.extend
myList.index
myList.insert
myList.pop
myList.remove
myList.reverse
```

Abbildung 2.9 ... und auch von Objekten.

Fehler und Debugging

Für eine einfache Fehleranalyse gibt es die Magic Function `%xmode`, die es Ihnen erlaubt, die Menge der ausgegebenen Fehlerinformationen zu steuern. Diese Funktion verlangt eines der Argumente `Plain`, `Context` oder `Verbose`, die den Detaillierungsgrad an Debug-Informationen im Falle eines Fehlers steuern.

Debugging – Fun Fact

Unter *Debugging* versteht man in der Welt der Programmierung das Nachverfolgen des Anwendungsablaufs. Das wird vor allem dann gemacht, wenn die Anwendung nicht das tut, was sie sollte, beziehungsweise einen Fehler (also einen *Bug*) generiert. Der Ursprung des Wortes wird oft auf den 9. September 1945 datiert.

Man fand Ungeziefer (engl. *bug*), genauer gesagt eine Motte, in einem Relais eines Computers. Diese Motte war für einen Fehler verantwortlich, der diesen Computer lahmlegte. Abbildung 2.10 zeigt das damalige Logbuch mit der eingelegten Motte. Allerdings wurde der Begriff schon vorher in diesem Sinne verwendet; die damalige Mitarbeiterin fand es aber wie wir witzig, dass ein echter Bug zu einem Computerausfall führte.

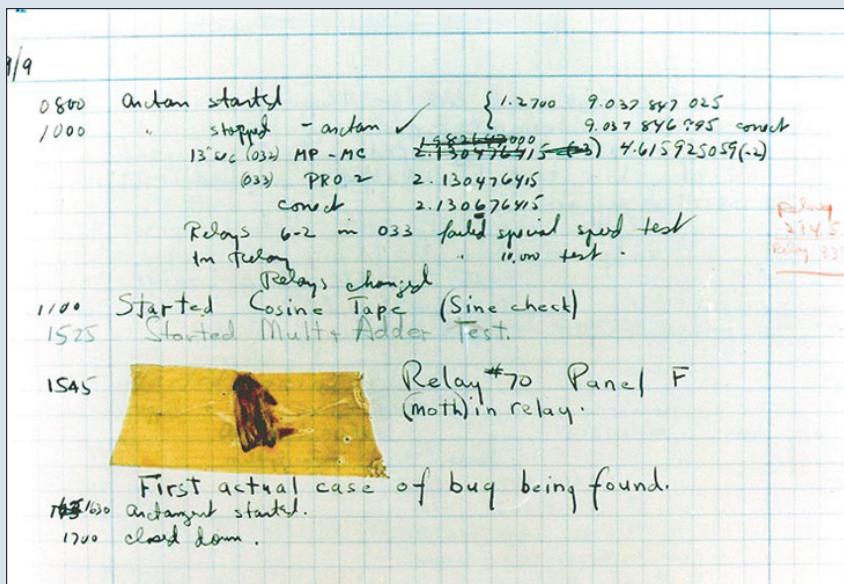


Abbildung 2.10 Der erste Bug in der Computergeschichte (Quelle: U.S. Naval Historical Center Online Library)

In den folgenden Beispielen sehen Sie die Auswirkungen der Argumente. Wir greifen hier schon ein bisschen auf die Python-Programmierung vor. Wenn Sie sich darin noch unsicher fühlen, verweisen wir Sie auf Anhang A.

In Listing 2.3 definieren wir zuerst eine einfache Funktion `avg_bee_length()`, die die Aufgabe hat, die Mittelwerte der Längen unserer Bienen zu berechnen. Innerhalb dieser Funktion rufen wir dann die Funktion `fehlerfunc()` auf, wohl wissend, dass der Aufruf einen Fehler generieren könnte.

```

def fehlerfunc(x,y):
    return x/y

def avg_beelength(valueList):
    total = sum(valueList)
    count = len(valueList)

    return fehlerfunc(total,count)

bees = (3, 3.5, 4.1, 5.2, 5, 2)
nobees = ()
avg_beelength(bees)
# Ausgabe:
3.800000000000003

```

Listing 2.3 Die Funktion »avg_beelength« ohne Fehler

So weit, so gut: Listing 2.3 definiert unsere beiden Funktionen. Mit der Liste der Bienenlängen, bees, erhalten wir ein passables Ergebnis und vor allem keinen Fehler.

Sehen wir uns nun in Listing 2.4 an, welche Fehlermeldung angezeigt wird (abhängig vom eingestellten %xmode), wenn wir die gleiche Funktion mit nobees, also einer Liste ohne Werte, aufrufen:

```

# Das ist die Standardeinstellung
%xmode Context
avg_beelength(nobees)
# Ausgabe:
-----
ZeroDivisionError                                 Traceback (most recent call last)
<ipython-input-11-be3dced12b16> in <module>
      1 # Das ist die Standardeinstellung
      2 get_ipython().run_line_magic('xmode', 'Context')
----> 3 avg_beelength(nobees)

<ipython-input-10-185b8ad2cde6> in avg_beelength(valueList)
      6     count = len(valueList)
      7
----> 8     return fehlerfunc(total,count)
      9
     10 bees = (3, 3.5, 4.1, 5.2, 5, 2)

```

```
<ipython-input-10-185b8ad2cde6> in fehlerfunc(x, y)
  1 def fehlerfunc(x,y):
----> 2     return x/y
      3
  4 def avg_beelength(valueList):
  5     total = sum(valueList)

ZeroDivisionError: division by zero
# etwas kompaktere Fehlermeldung
%xmode Plain
avg_beelength(nobees)
# Ausgabe:
Exception reporting mode: Plain
Traceback (most recent call last):

File "<ipython-input-12-fe49876a1c72>", line 3, in <module>
    avg_beelength(nobees)

File "<ipython-input-10-185b8ad2cde6>", line 8, in avg_beelength
    return fehlerfunc(total,count)

File "<ipython-input-10-185b8ad2cde6>", line 2, in fehlerfunc
    return x/y

ZeroDivisionError: division by zero

# ausführliche Fehlermeldung
%xmode Verbose
avg_beelength(nobees)
# Ausgabe:
Exception reporting mode: Verbose
-----
ZeroDivisionError                               Traceback (most recent call last)
<ipython-input-13-3d86e356f330> in <module>
      1 # ausführliche Fehlermeldung
      2 get_ipython().run_line_magic('xmode', 'Verbose')
----> 3 avg_beelength(nobees)
      global avg_beelength = <function avg_beelength at 0x000000004FE1B88>
      global nobees = ()
```

```

<ipython-input-10-185b8ad2cde6> in avg_beelength(valueList=())
    6     count = len(valueList)
    7
----> 8     return fehlerfunc(total,count)
    global fehlerfunc = <function fehlerfunc at 0x000000000500BDC8>
    total = 0
    count = 0
    9
10 bees = (3, 3.5, 4.1, 5.2, 5, 2)

<ipython-input-10-185b8ad2cde6> in fehlerfunc(x=0, y=0)
    1 def fehlerfunc(x,y):
----> 2     return x/y
    x = 0
    y = 0
    3
    4 def avg_beelength(valueList):
    5     total = sum(valueList)

```

`ZeroDivisionError: division by zero`

Listing 2.4 Fehlermeldungen mit unterschiedlichem Detailgrad

In allen Modi wird ein sogenannter *Stacktrace* bzw. *Traceback* angezeigt, der den Fehlerverlauf anzeigt. Er beginnt mit der Funktion auf der obersten Ebene und geht bis zum letzten Element, das den Fehler ausgelöst hat. In unserem Fall ist das `return x/y` in der Funktion `fehlerfunc()`.

Der Standardmodus `Context` liefert aber nicht nur die Funktionenfolge wie im Modus `Plain`, sondern auch die unmittelbare Umgebung im Programmcode. Damit findet man zum einen die Programmstellen leichter, zum anderen liegt die Fehlerursache oft in der unmittelbaren Umgebung.

Der Modus `Verbose` zeigt uns noch mehr, nämlich den Wert der lokalen Variablen. Es zeigt in den letzten Zeilen, dass die Werte für `x` und `y` gleich 0 sind. Damit ist uns die Ursache schon klar – wir haben durch die Zahl 0 dividiert (*Division by Zero*), was natürlich nicht geht.

Wird der Code allerdings komplexer, kann die Analyse der Fehlermeldung sehr schwierig werden, da der Traceback extrem lang werden kann.

Traceback (bzw. Stacktrace)

Der Traceback ist ein Werkzeug zur Rückverfolgung eines Programmfehlers. Ausgehend vom Auftreten des Fehlers werden alle daran beteiligten Funktionen angezeigt, was die Rückverfolgung bis zur Fehlerursache ermöglicht. Je komplexer und verschachtelter der Programmcode ist, desto länger werden solche Stacktraces.

Manchmal liefert der Code keinen Fehler, tut aber trotzdem nicht ganz das, was man eigentlich erreichen wollte. Um dann genauere Informationen zu erhalten, was alles bei der Ausführung des Programmcodes passiert, verwendet man einen *Debugger*. Im Jupyter Notebook wird der interaktive Debugger mit dem Befehl `%debug` gestartet.

Rufen wir nun diesen interaktiven Debugger auf, dann wird ein Kommandozeilenfenster geöffnet mit dem Prompt `ipdb>`. In dieser Kommandozeile können wir Befehle eingeben, die uns unter anderem die Inhalte unserer Variablen anzeigen. Somit können wir hoffentlich erkennen, wo der Fehler liegt. Listing 2.5 zeigt, wie der interne Debugger zu verwenden ist.

In unserem Traceback erkennen wir, dass wir zumindest zwei Stufen haben. Die letzte zeigt in der Funktion `fehlerfunc(x,y)`, wo genau der Fehler passiert ist. Hier haben wir Zugriff auf alle Variablen innerhalb der Funktion, also `x`, `y`, und natürlich auf eventuelle globale Variablen. Wollen wir eine Stufe höher gehen, in diesem Fall in die Funktion `avg_beelength()`, so müssen wir innerhalb des Debuggers den Befehl `up` eingeben. Mit der Ausgabe der Variablen `count` und `total` auf dieser Stufe ergibt dies beide Male 0 und führt uns zum eigentlichen Fehler, nämlich zur Eingabe einer leeren Liste.

```
%debug
```

```
# Ausgabe:
> <ipython-input-1-78d6430f450b>(2)fehlerfunc()
    1 def fehlerfunc(x,y):
----> 2     return x/y

ipdb> print(x, y)
0
ipdb> print(count)
*** NameError: name 'count' is not defined

ipdb> up
> <ipython-input-10-185b8ad2cde6>(8)avg_beelength()
    6     count = len(valuelist)
    7
```

```

----> 8      return fehlerfunc(total,count)
         9
10 bees = (3, 3.5, 4.1, 5.2, 5, 2)

ipdb> print(total, count)
0 0
ipdb> quit

```

Listing 2.5 Die Verwendung des interaktiven Jupyter Debuggers

Natürlich könnte man wie immer Bücher über das Thema Fehleranalyse und Debugging schreiben. Für unsere Zwecke reichen die hier vorgestellten Werkzeuge aber aus.

2.1.3 Wichtige Python-Module

In diesem Abschnitt beschreiben wir in aller Kürze die für dieses Buch wichtigen Python-Module. Für jedes dieser Module gibt es sogenannte *Cheat Sheets*, die auf einer DIN-A4-Seite die wichtigsten Funktionen und Klassen zu einem Modul beispielhaft zeigen. Sie finden sie unter <https://www.datacamp.com/community/data-science-cheatsheets>. Dennoch sei an dieser Stelle bereits gewarnt, dass wir im Verlauf des Buches Elemente aus Modulen verwenden, die wir nicht extra vorstellen oder genauer beschreiben. Diese Arbeitsweise ist durchaus typisch in Python-Projekten, da es fast zu jeder Aufgabe eine Bibliothek gibt, die einfach zu verwenden ist und uns eine Menge Arbeit erspart, genauer gesagt: Wir müssen nicht selbst programmieren.

NumPy

NumPy steht für *Numerical Python* und hat für das wissenschaftliche Rechnen eine große Bedeutung. Dieses Modul enthält eine optimierte Datenstruktur für mehrdimensionale Matrizen und die dazugehörigen Funktionen der linearen Algebra.

Das Herzelement dieses Moduls ist die Klasse `ndarray` für mehrdimensionale Matrizen.

Sie finden NumPy unter <http://www.NumPy.org>.

matplotlib

Dieses Modul erlaubt die Erstellung von wissenschaftlichen Diagrammen und Visualisierungen. Es stellt Funktionen für Liniendiagramme, Histogramme und Streudiagramme zur Verfügung. Wird matplotlib mit Jupyter Notebook verwendet, so können die Visualisierungen direkt im Notebook dargestellt werden.

Sie finden matplotlib unter <http://www.matplotlib.org>.

scikit-learn

scikit-learn ist ein sehr beliebtes Werkzeug und enthält eine große Anzahl an modernen Machine-Learning-Algorithmen. Es ist gleichermaßen in der Industrie und im Universitätsbereich in Verwendung.

Sie finden scikit-learn unter <https://scikit-learn.org>.

pandas

pandas erlaubt die Datenvorverarbeitung und Zusammenführung von Daten aus unterschiedlichen Quellen. Im Besonderen können Daten durch eine SQL-ähnliche Sprache mittels Abfragen und dem Zusammenführen von Tabellen zu den Strukturen Series (eindimensional), DataFrames (zweidimensional) und Panels (dreidimensional) eingelesen werden. Vor allem die DataFrames sind für uns von Bedeutung.

Sie finden pandas unter <https://pandas.pydata.org>.

TensorFlow

TensorFlow ist eigentlich ein ganzes Paket an Bibliotheken, Online-Ressourcen etc., das wir genauer in Anhang C beschreiben werden. Google ist es zu verdanken, dass wir damit für das Deep Learning entwickelte Machine-Learning-Bibliotheken zur Verfügung haben. Die Berechnungen erfolgen mit sogenannten *Datenflussgraphen* (engl. *data flow graphs*). Ein Graph wiederum besteht aus Knoten, die durch Kanten verbunden sind. Jeder mit TensorFlow berechnete Algorithmus besteht aus mathematischen Operationen (repräsentiert als *Knoten*) und Daten (repräsentiert als *Kanten*). Damit lassen sich alle Berechnungen von einer einfachen Summe bis zu hochkomplexen Matrixoperationen durchführen und darstellen.

Im Buch verwenden wir die Version TensorFlow 2, die eine von Francois Chollet entwickelte Keras-Bibliothek zur einfacheren Handhabung von Deep Neural Nets enthält.

Sie finden das Framework unter <http://www.tensorflow.org>.

Keras

Keras ist eine eigenständige Bibliothek, die mit verschiedenen Neuronale-Netze-Modulen arbeiten kann, wie zum Beispiel mit TensorFlow, MXNet oder Theano. Diese Module werden als *Backends* bezeichnet. Mit dem Release TensorFlow 2 hat Google Keras vollständig integriert. Keras wird aber auch als eigenständige Bibliothek weitergeführt, obwohl auch Francois Chollet den Benutzern vorschlägt, in Zukunft nur mit TensorFlow 2 zu arbeiten.

Sie finden die Bibliothek und ihre Dokumentation hier:

- ▶ <http://www.keras.io>
- ▶ <https://www.tensorflow.org/guide/keras>

2.1.4 Die Google Colab-Plattform für Jupyter Notebooks

Sie wissen nun, dass Jupyter Notebooks eine webbasierte Entwicklungsumgebung ist. Der dazu notwendige Webserver kann lokal auf Ihrem eigenen Rechner laufen. Anaconda hilft Ihnen bei der Bereitstellung der notwendigen Applikationen. Als webbasierte Anwendung können Sie diese Entwicklungsumgebung aber auch in der Cloud laufen lassen. Wir empfehlen Ihnen dazu Google Colab – nicht, weil wir von Google bezahlt werden, sondern weil es folgende Vorteile bietet:

- ▶ **Einheitliche Plattform:** Alle Leser und Leserinnen haben die gleiche Plattform und Installation. Wichtige Bibliotheken, wie TensorFlow, Keras, pandas etc. sind bereits mit ihren aktuellen Versionen installiert.
- ▶ **GPU/TPU-Unterstützung:** Google Colab bietet einen kostenlosen Zugang zu leistungsfähigen Hardwarebeschleunigern (GPUs/TPUs), die für das Trainieren von neuronalen Netzen unerlässlich sind.
- ▶ **Freigabe und Zusammenarbeit:** Ähnlich wie Google Docs können Jupyter Notebooks geteilt und gemeinsam bearbeitet werden. Finden sich Lesergruppen, die gemeinsam ein Projekt angehen?
- ▶ **Integration in Google Drive:** Notebooks und Daten können in Google Drive gespeichert werden und dort direkt geöffnet bzw. verwendet werden.

Sehen wir uns dazu erst einmal in Abbildung 2.11 an, wie Abbildung 2.5 im Google Colab-Kleid aussieht:

An dieser Abbildung fällt auf, dass Google Colab für die Markdown-Zellen einen Editor integriert hat. Somit können wir wie gehabt Markdown-Elemente verwenden, müssen sie aber nicht auswendig lernen, denn mit dem Editor würden die Markdown-Kommandos automatisch eingefügt.

Wir raten Ihnen, alle Eigenschaften und Funktionen von Google Colab selbstständig zu erforschen. Hier weisen wir nur auf die wichtigsten Elemente hin, die das Arbeiten mit diesem Werkzeug vereinfachen.

The screenshot shows a Jupyter Notebook interface in Google Colab. At the top, there's a toolbar with various icons for file operations, search, and sharing. The notebook title is "Kapitel_02-01.ipynb". Below the toolbar, there are three sections:

- Markdown-Zelle**: Contains text about Markdown syntax, including numbered lists, bullet points, and links. It also includes a rich text editor toolbar.
- Code-Zelle**: Contains Python code that prints a message. The output cell shows the printed message.

Abbildung 2.11 Jupyter Notebook in Google Colab

Verzeichnis und Google Drive-Verbindung

Die vertikale Toolbar ganz links enthält einen Button mit einem Aktenordner-Symbol. Dieser Button blendet die Verzeichnisstruktur, in der dieses Jupyter Notebook läuft, ein oder aus.

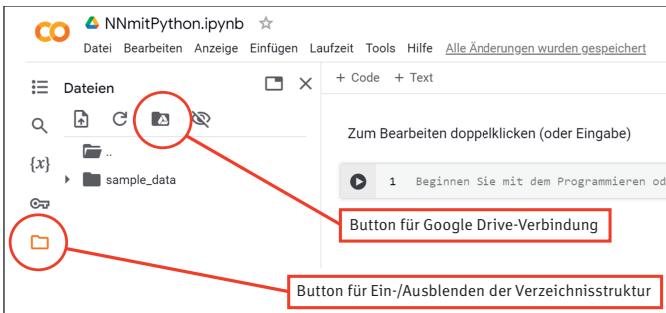


Abbildung 2.12 Buttons für die Verzeichnisstruktur und die Google Drive-Verbindung

Um eigene Daten in diese Verzeichnisstruktur einzufügen, reicht es, die gewünschte Datei in den ausgeklappten Bereich der Verzeichnisstruktur zu ziehen. Sie können das-selbe auch über folgende Programmzeilen in einer Code-Zelle erreichen:

```
from google.colab import files
uploaded = files.upload()

for fn in uploaded.keys():
    print('Datei "{}" mit Länge {} bytes'.format(name=fn, length=
len(uploaded[fn])))
```

Listing 2.6 Öffnen eines Auswahlfensters zum Laden von lokalen Dateien

Ähnlich funktioniert auch das Speichern von Dateien von der Google Colab-Umgebung auf den lokalen Rechner: entweder manuell, indem Sie mit der rechten Maustaste auf die Datei klicken und den Menüpunkt HERUNTERLADEN wählen, oder über im Programm über eine Code-Zelle:

```
from google.colab import files

with open('Beispiel.txt', 'w') as f:
    f.write('Neuronale Netze sind beeindruckend!')

files.download('Beispiel.txt')
```

Listing 2.7 So speichern Sie Dateien von der Google Colab-Umgebung auf das lokale Dateisystem.

Wollen Sie Ihr Notebook mit Google Drive verbinden, klicken Sie auf den in Abbildung 2.12 markierten Button. Diese Verbindung wird durch den Ordner *drive* in der Verzeichnisstruktur angezeigt, der in Abbildung 2.13 zu sehen ist.

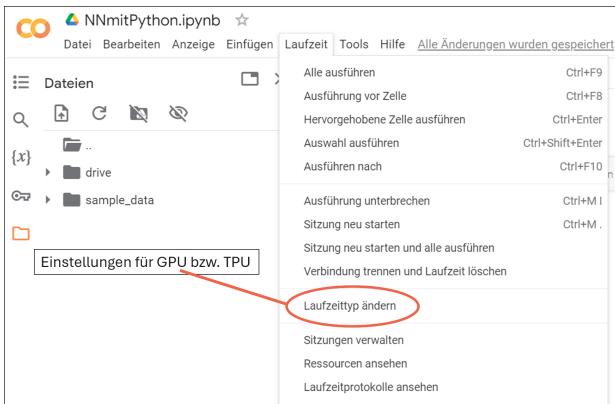


Abbildung 2.13 Menüpunkt zur Einstellung von GPU oder TPU

Nun haben Sie viel Speicherplatz für die oft großen Trainingsdaten für neuronale Netze zur Verfügung.

Kostenlose GPU bzw. TPU

Möchten Sie in den Genuss einer kostenlosen GPU oder TPU kommen, so rufen Sie in der horizontalen Menüleiste den Menüpunkt LAUFZEIT auf, der Ihnen die Möglichkeit bietet, den Laufzeittyp zu ändern (siehe Abbildung 2.13). Wichtig ist, diese Änderung vor dem Arbeiten mit dem Notebook durchzuführen. Wird der Laufzeittyp geändert, so startet das Notebook neu und löscht alle bisherigen Abläufe und Ausführungen.

Wir müssen noch erwähnen, dass die GPU Ihnen nur für begrenzte Zeiträume von ein paar Stunden zur Verfügung steht. Ein Trainieren von neuronalen Netzen ist daher nicht möglich, es sei denn, Sie wählen eine bezahlte Version von Google Colab. Für die Beispiele in diesem Buch ist die kostenlose Version jedoch völlig ausreichend.

Ausführen von Code-Zellen

Abbildung 2.13 zeigt das Untermenü des Menüpunkts LAUFZEIT. In seinem oberen Bereich können Sie durch Auswahl des entsprechenden Menüpunkts oder durch die entsprechende Tastenkombination einzelne oder alle Zellen ausführen. Jede Code-Zelle ist zusätzlich mit einem Button mit dem Zeichen für das Starten ausgestattet, der sich links von der Code-Zelle befindet – ein weiterer Weg, um eine Code-Zelle auszuführen.

2.1.5 Weitere Jupyter Notebook-Cloud-Ressourcen

Die folgende Liste ist ein guter Startpunkt, aber aus mehreren Gründen unvollständig. Zum einen gibt es eine fast unüberschaubare Anzahl an Anbietern, zum anderen bieten

die angegebenen Webseiten viel mehr als nur Jupyter Notebooks an. Wir haben einige ausgewählt, die bis zu einem gewissen Grad die Ressourcen gratis zur Verfügung stellen.

Jupyter Community

Zuerst muss natürlich die Webseite der *Jupyter Community* erwähnt werden, die Jupyter-Ressourcen zur Verfügung stellt – allerdings ohne GPU-Unterstützung und eigentlich nur zum Ausprobieren.

- ▶ <https://jupyter.org/try>

Kaggle

Kaggle ist eine Data-Science-Plattform, die sich ursprünglich auf Machine Learning Competitions spezialisiert hatte. Sie ist inzwischen eigentlich ein Muss, wenn Sie Expertenwissen in der Data Science erwerben wollen.

- ▶ <https://www.kaggle.com/notebooks>

Microsoft Azure

Auch Microsoft ist mit *Azure* sehr stark in der Machine-Learning-Szene vertreten.

- ▶ <https://notebooks.azure.com>

Amazon SageMaker

Wie viele Cloud-Anbieter hat auch Amazon mit Amazon Web Services (AWS) – einer davon ist der SageMaker – Jupyter Notebooks im Portfolio.

- ▶ <https://aws.amazon.com/de/sagemaker>

2.2 Zusammenfassung

Dieses Kapitel bot Ihnen eine Einführung in den Anaconda Navigator, in Jupyter Notebook und dazugehörige Jupyter Notebook-Cloud-Ressourcen, die zusammen eine komfortable technische Entwicklungsumgebung für die Programmiersprache Python darstellen. Um eine kleine Einführung in Python zu erhalten, können Sie jetzt zu Anhang A, »Python kompakt«, vorblättern.

Python bietet eine Unmenge an Bibliotheken. Die wichtigsten für die Programmierung von neuronalen Netzen haben wir hier kurz beschrieben, mit Verweisen auf sehr gute Tutorials. Eine Liste mit weiteren Anbietern von Jupyter Notebooks in der Cloud vervollständigte dieses Kapitel.

Kapitel 3

Ein einfaches neuronales Netz

Ein erster Schritt: ein einfaches Netz, das die Einordnung von »Dingen« in zwei Kategorien ermöglicht.

Dieses Kapitel beschreibt die einfachste Form neuronaler Netze, die sogenannten *Perceptrons*, die auch als *lineare Klassifizierer* bezeichnet werden. Wir werden Schritt für Schritt mit Bildern und Codes die Funktionsweise der Perceptrons erarbeiten. Einfache Beispiele erläutern die Materie, und am Ende können Sie einem fahrenden Roboter beibringen, Löcher zu erkennen oder es zu vermeiden, in die Wand zu fahren.

3.1 Vorgeschichte

Stellen Sie sich vor, Sie müssen keine Planung mehr machen, weil ein neuronales Netz diese für Sie übernimmt. Lassen Sie uns diese Art von Planung als »automatische Planung« bezeichnen. Das Szenario sieht folgendermaßen aus: Personen sollen nach individuellen Terminwünschen und unternehmerischen Bedarfen so geplant werden, dass der unternehmerische Zweck mithilfe ebendieser Personen umgesetzt werden kann. Beispiele dafür sind die *Personaleinsatzplanung* eines Tante-Emma-Ladens oder der Stundenplan in einer Schule.

3.2 Her mit dem neuronalen Netz!

Eine Frage vorweg: Wie würden Sie die zwei Gruppen in Abbildung 3.1 voneinander trennen, wenn Sie einen Stift und ein Lineal verwenden dürften? Und wenn Sie das geschafft haben – finden Sie noch weitere Möglichkeiten?

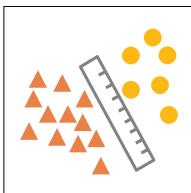


Abbildung 3.1 Lineare Trennung

Für uns Menschen ist es denkbar einfach, diese Aufgabe zu lösen. Wir sehen sofort, dass zwei Gruppen vorhanden sind, die wir einfach durch eine *Gerade* trennen können. Doch wie kann dieser Vorgang maschinell durchgeführt werden, welche Ansätze stehen dafür zur Verfügung? Die sperrige Antwort lautet je nach Fachgruppe:

- ▶ Computerwissenschaftler: durch einen *Klassifikator*
- ▶ Statistiker: durch einen *linearen Diskriminatoren*
- ▶ Data Scientist: durch ein *Perceptron*, ein spezielles KNN
- ▶ [Wählen Sie bitte Ihre Fachgruppe]: [Wählen Sie bitte Ihren Ansatz]

Wir werden natürlich die Linie des Data Scientists weiterverfolgen. Aber was ist mit *Perceptron* gemeint? Diese Frage wird uns bis zum Ende dieses Kapitels beschäftigen, und zwar mit folgendem Beispiel:

Die Filialleiterin (Frau Äppel) eines kleinen Ladens hat zwei Mitarbeiter zu planen: Herrn Lauch und Frau Karotte. Beide haben von Montag bis Freitag prinzipiell Zeit zum Arbeiten. Samstags und natürlich am Sonntag ist der Laden zu. Frau Äppel hätte gerne jeden Tag **mindestens einen** Mitarbeiter zur Seite, also zum Beispiel am Montag

- ▶ Frau Karotte allein
- ▶ oder Herrn Lauch allein
- ▶ oder beide gemeinsam, aber sicher mindestens einen von beiden

Wenn man das so hört, klingt es ein wenig kompliziert. Einfacher ist das Problem zu verstehen, wenn wir es in Form von Tabelle 3.1 aufschreiben:

Frau Karotte	Herr Lauch	-->	Montag
nicht anwesend	nicht anwesend	-->	nicht okay
anwesend	nicht anwesend	-->	okay
nicht anwesend	anwesend	-->	okay
anwesend	anwesend	-->	okay

Tabelle 3.1 Erster Personalplan

Wenn Frau Karotte und Herr Lauch nicht anwesend sind, dann ist der Wunsch von Frau Äppel nicht erfüllt. Es ist keiner da zum Helfen, und nach so einem Tag ist Frau Äppel »nicht okay«, da sie alles allein machen muss. Besser ist es, wenn zum Beispiel Frau Karotte anwesend ist, denn das ist auf alle Fälle »okay« für Frau Äppel.

Mit dieser Tabellendarstellung sind die möglichen Kombinationen schon viel einfacher zu verstehen. Aber optimal ist das immer noch nicht. Ein KNN kann mit den Begriffen »nicht anwesend«, »anwesend«, »nicht okay« und »okay« nicht rechnen, und daher müssen wir uns eine andere Darstellung mit Zahlen überlegen.

Anmerkung am Rande: Datenaufbereitung

Dieser Schritt der Datenaufbereitung ist einer der schwierigsten Schritte im Gesamtprozess des Arbeitens mit KNN. Woher soll man bitte wissen, in welches Format die Daten für das KNN zu transformieren sind? Woher soll man wissen, was aus »anwesend«/»nicht anwesend« werden soll? In Teil II werden wir uns dieses Thema noch genauer vornehmen.

Ohne jetzt großartig eine Theorie zur Datenvorverarbeitung für das Arbeiten mit KNN zu entwickeln, stellen wir aber trotzdem die Frage, wie man Text wie »okay« oder »nicht okay« als KNN-Input passend aufbereitet: Haben Sie dazu eine Idee? Wie könnte die Vorverarbeitung aus Ihrer Sicht aussehen? Bitte kurz nachdenken!

Danke für das Nachdenken, wir haben uns dazu folgenden Vorschlag überlegt: Wenn ein Mitarbeiter *anwesend* ist, dann schreiben wir »1«, und wenn der Mitarbeiter *nicht anwesend* ist, »0«. Diese Festlegung scheint möglicherweise willkürlich, und in der Tat ist sie das auch. Jedoch ist der Wert 1, wenn etwas »an« oder »gut« ist, passend für das positive Ereignis, und 0 für »aus« oder »schlecht« auch intuitiv nachvollziehbar. Daher passt diese Zuordnung in diesem Kontext aus unserer Sicht ganz gut. Andererseits ersetzen wir »nicht okay« durch »0« und »okay« durch »1«, um darzustellen, wie das gewünschte Ergebnis der Planung aussehen sollte.

Wenn wir unsere Überlegungen wieder im Tabellenformat darstellen, bekommen wir die Inhalte in Tabelle 3.2:

Frau Karotte	Herr Lauch	-->	Montag
0	0	-->	0
1	0	-->	1
0	1	-->	1
1	1	-->	1

Tabelle 3.2 Zweiter Personalplan mit KNN-geeigneter Codierung

Damit können wir nun rechnen. Moment einmal, wieso rechnen? Wir wollen einem KNN beibringen, dass es **nicht okay** ist, wenn **kein Mitarbeiter** da ist, und dass die Situation in allen anderen Anwesenheitskonstellationen immer **okay** ist. Das heißt, dass wir uns eine Berechnung überlegen müssen, um das gewünschte Ergebnis zu bekommen. Fürs Erste beginnen wir ohne ein KNN, also so richtig mit Nachdenken und Selberrechnen. Das wird superwichtig für das Verständnis der eigentlich sehr einfachen Berechnungen, die in einem KNN durchgeführt werden.

Der erste Versuch könnte darin bestehen, dass Frau Karotte und Herr Lauch zusammengezählt werden – es werden nicht wirklich die Personen summiert, sondern die Anwesenheit der beiden. Das Ergebnis wird dann von dieser Summe mit einer einfachen Regel abgeleitet, die da lautet: »Wenn die Summe größer gleich 1 ist, dann ist das Ergebnis 1, ansonsten ist das Ergebnis 0.«

Sehen wir uns das Ganze nun in Form von Tabelle 3.3 an:

Frau Karotte	Berechnung	Herr Lauch	-->	Summe	Ergebnis
0	+	0	=	0	0
1	+	0	=	1	1
0	+	1	=	1	1
1	+	1	=	2	1

Tabelle 3.3 Errechneter Personalplan aus den Anwesenheiten

Diese Überlegungen könnten wir natürlich in Form eines Netzes darstellen, so wie es in Abbildung 3.2 zu sehen ist, um eine andere Sicht als die der Tabellendarstellung zu erhalten. Wir wollen uns ja schrittweise der KNN-Visualisierung nähern.

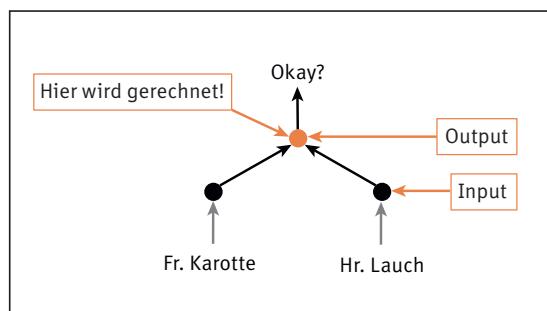


Abbildung 3.2 Die Entscheidung für Frau Karotte und Herrn Lauch berechnen

Diese Art der Darstellung ist nicht mehr weit weg von der Darstellung eines KNN. Wir müssen noch ein paar Details ergänzen, aber im Prinzip sind wir schon fertig – fast.

3.3 Neuron-Zoom-in

Eines der Details, die noch zu ergänzen sind, betrifft die Art der Berechnung. Sehen wir uns das Neuron genauer an.

Aus Tabelle 3.3 wissen wir, dass wir eine Berechnung und dann eine Entscheidung benötigen. Diese Erkenntnis ergänzen wir in der Darstellung des Berechnungsknotens in Abbildung 3.3.

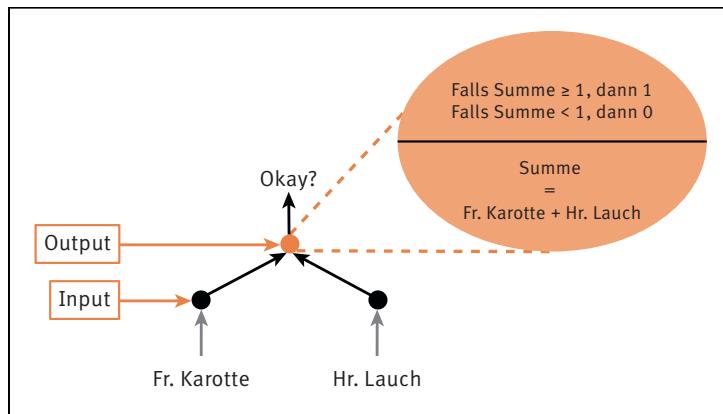


Abbildung 3.3 Die Ergebnisberechnung für Frau Karotte und Herrn Lauch im Detail

Um die Details zum Berechnungsknoten so in Form zu bringen, dass ein Python-Programm damit umgehen kann, benötigen wir eine Formel bzw. eine Rechenvorschrift, und die könnte so aussehen:

$$\text{Summe} = \text{Fr. Karotte} + \text{Hr. Lauch}$$

und

$$\text{Ergebnis} = \text{Entscheidung}(\text{Summe})$$

wobei

$$\text{Entscheidung} = \begin{cases} 1, & \text{falls Summe} \geq 1 \\ 0, & \text{falls Summe} < 1 \end{cases}$$

Noch ein Wort zu der Zeile $\text{Ergebnis} = \text{Entscheidung}(\text{Summe})$: Wir haben die sogenannte *Funktionsschreibweise* für die Berechnung der Entscheidung verwendet. Diese besagt nur, dass das Ergebnis der Funktion *Entscheidung* von der *Summe* abhängt. Oder anders gesagt, dass das Ergebnis der Funktion *Entscheidung* in Abhängigkeit von der *Summe* ermittelt wird. Sehr oft werden die Funktionen mit $f(x)$ bezeichnet, wie zum Beispiel $f(x) = x^2$. Das besagt nur, dass die Funktion f in Abhängigkeit von x definiert ist.

Die Funktionsschreibweise kann direkt in die Python-Programmierung übernommen werden. Wir machen das nun und definieren uns eine Funktion mit der eben beschriebenen Funktionalität. Falls Sie Ihr Jupyter Notebook noch nicht gestartet haben, dann bitten wir Sie, das nun zu erledigen. Sie können auch gerne nochmals zu Kapitel 2, »Das minimale Starterkit für die Entwicklung von neuronalen Netzen mit Python«, blättern und die Details nachlesen. Wir werden die Programmbeispiele in unterschiedlichen Notebooks organisieren, wobei wir pro Kapitel ein Notebook vorgesehen haben. Ihr nächster Schritt besteht somit darin, dass Sie sich ein neues Notebook anlegen (siehe Abbildung 3.4).

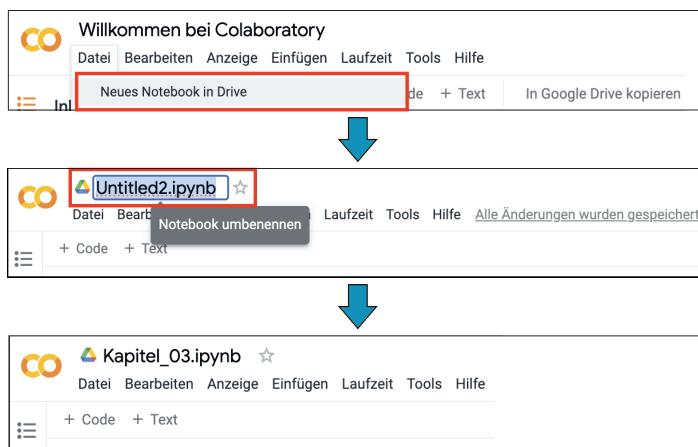


Abbildung 3.4 Das Notebook »Kapitel_03« anlegen

Nutzen Sie dazu im Menü DATEI den Menüeintrag NEUES NOTEBOOK IN DRIVE. Daraufhin erzeugt Colab das neue Notebook, und Sie können mittels Mausklick auf den Text UNTITLED.IPYNB einen Namen für das Notebook vergeben, zum Beispiel »Kapitel_03«.

Sie können den Text direkt überschreiben. Damit wird das Notebook in »Kapitel_03« umbenannt, und Sie können mit dem Programmieren beginnen. Wir werden zwei Zellen verwenden; die erste Zelle enthält die Überschrift und die zweite Zelle den Code.

Wir werden die Zelle als TEXT-Zelle einfügen (siehe Abbildung 3.5), um dem Code in der folgenden Zelle eine Überschrift zu spendieren. Klicken Sie dazu auf den Button + TEXT.



Abbildung 3.5 Eine Zelle mit »Text« einfügen

Nachdem dies geschehen ist, können Sie die Auszeichnungselemente für eine Überschrift in der Zelle eingeben, zum Beispiel ein »#« für eine Überschrift der Ebene 1 und, durch ein Leerzeichen getrennt, den Text für die Überschrift (siehe Abbildung 3.6). Natürlich können Sie auch die Texteditor-Toolbar von Colab verwenden, um die Formatierungen durchzuführen. Rechts neben der Eingabe erscheint bereits die Vorschau (*Preview*) des Textes mit den Formatierungsauswirkungen.

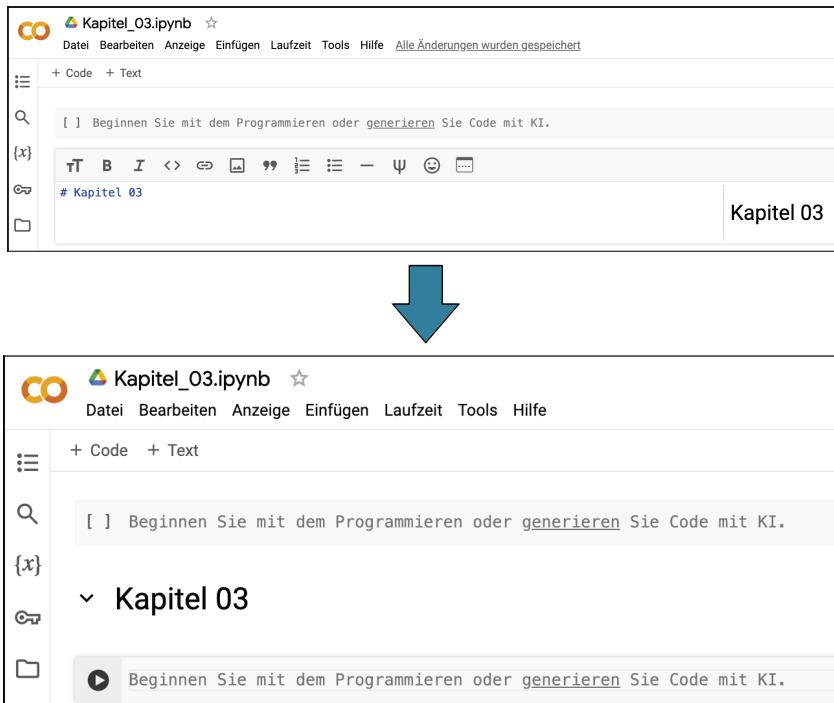


Abbildung 3.6 Notebook mit Überschrift

Wenn Sie Ihre Eingabe mit der Tastenkombination $\boxed{\text{Shift}}$ + $\boxed{\text{Enter}}$ bestätigen, wird der Text in der Markdown-Zelle formatiert dargestellt.

Immer wiederkehrender Header

Wir haben in den Notebooks zum Herunterladen noch zwei Aspekte eingefügt, die beinahe in allen Notebooks vorkommen:

- ▶ einen Hinweis, dass Sie beim erstmaligen Start den Menüpunkt LAUFZEIT • ALLE AUSFÜHREN aufrufen sollten. Dies empfehlen wir, da auf diese Weise gewährleistet ist, dass die Scripts in allen Code-Zellen der Reihe nach ausgeführt und die besprochenen Outputs erzeugt werden.

- eine MARKDOWN-Zelle und eine CODE-Zelle, die eine etwaige Warnung zum Beispiel hinsichtlich zukünftiger *deprecated*-Anweisungen (dt. *veralteter Anweisungen*) unterdrücken. Wenn Sie die Warnung sehen wollen, dann setzen Sie einfach ein Kommentarzeichen vor den Code in der Zelle.

Klicken Sie nun auf eine Zelle und bewegen Sie die Maus zum oberen oder unteren Rand der Zelle. Dadurch erscheint eine Toolbar, mit deren Hilfe Sie sehr einfach eine Code- oder Text-Zelle einfügen können. In dieser Zelle platzieren Sie nun weitere Markdowns bzw. Ihren Code, zum Beispiel den aus Listing 3.1:

```
def entscheidung( summe ):
    """ Berechnung der Entscheidung zum Wert summe
    Input: summe
    Output: 1, falls summe >= 1,
            0 sonst
    """
    if summe >= 1:
        return 1
    else:
        return 0
#-----
# Berechnung der entscheidung
ergebnis = entscheidung(1)
# Ausgabe in Zelle als String
print('Input: ', 1)
print('Output:', ergebnis)
# Ausgabe:
Input: 1
Output: 1
```

Listing 3.1 Die Entscheidung als Stufenfunktion

Die vorher besprochene Berechnung der Entscheidung bilden wir als Funktion `entscheidung` mit dem Parameter `summe` ab. Danach folgt im Code ein Funktions-*Docstring* (Dokumentationsstring), der die Funktion erläutert, so wie in Kapitel 2 beschrieben. Sehen Sie sich bitte Abbildung 3.7 an.

Nach dem Docstring folgt der Code für die Berechnung der Entscheidung. Falls der übergebene Wert größer als oder gleich 1 ist, dann wird als Ergebnis der Berechnung 1 zurückgegeben, ansonsten 0.

Wir haben direkt nach der Definition den Aufruf der Funktion eingefügt und mit dem Wert 1 getestet. Der Rückgabewert wird in der Variable `ergebnis` gespeichert und am Ende mit dem `print`-Befehl nach der Code-Zelle ausgegeben. Die Ausgaben der Programme werden wir hinter dem Code platzieren und fett markieren. Das Ergebnis in Ihrem Jupyter Notebook sollte dann so ähnlich aussehen wie in .

- ▼ Kapitel 03
- ACHTUNG:** Bitte zum Starten im Menü Laufzeit • Alle ausführen ausführen.
- ▼ Deaktivieren der Warnungen

```
[ ] import warnings
warnings.filterwarnings('ignore')
```

- ▼ Die Entscheidung
- ▼ Listing 3.1, Abbildung 3.7

```
[ ] def entscheidung( summe ):
    """ Berechnung der Entscheidung zum Wert summe
    Input: summe
    Output: 1, falls summe >= 1,
            0 sonst
    """
    if summe >= 1:
        return 1
    else:
        return 0
#-
# Berechnung der entscheidung
ergebnis = entscheidung(1)
# Ausgabe in Zelle
print('Input: ' + str(1))
print('Output: ' + str(ergebnis))

→ Input: 1
Output: 1
```

- ▼ Dokumentationsstring

```
[ ] # Ausgabe des Docstring mittels help-Funktion
print(entscheidung.__doc__)

→ Berechnung der Entscheidung zum Wert summe
Input: summe
Output: 1, falls summe >= 1,
        0 sonst
```

Abbildung 3.7 Programm und Ausgaben für die Funktion »entscheidung«

Wir hoffen, dass Sie mit diesen ausführlich beschriebenen Schritten gut gewappnet an die nächste Programmierung gehen, in der wir kurz die Anwendung des Moduls `matplotlib` besprechen werden. Dieses Modul kann, wie wir bereits in Kapitel 2 erwähnt haben, für die Erstellung von Diagrammen verwendet werden. Naheliegenderweise werden wir den Output der Funktion `entscheidung` mit ihm visualisieren.

3.4 Stufenfunktion

In der Funktion `entscheidung` haben wir unterschieden, ob die Summe kleiner als 1 ist oder nicht (dann ist sie größer oder gleich 1). Der Wert 1 wird als *Schwellenwert* bezeichnet, da wie bei einer Türschwelle ein Höhensprung – oder Tiefensprung, falls ich von oben komme – eingebaut ist. Schlagartig ändert sich der berechnete Wert der Funktion von 0 in 1.

Die Mathematiker haben sich für solch eine Funktion die Bezeichnung *Stufenfunktion* ausgedacht, im Englischen findet man die Bezeichnungen *step function*. Dabei kann der Schwellenwert einen beliebigen Wert annehmen. Falls der Schwellenwert 0 ist, wird die Stufenfunktion als *Heaviside-Funktion* bezeichnet, benannt nach dem Mathematiker *Oliver Heaviside*.

Wir verwenden für das einfache KNN, das wir gleich bauen werden, genau diese Stufenfunktion. Die Python-Funktion `entscheidung` dafür haben wir ja bereits programmiert. Bevor wir uns dem KNN zuwenden, werden wir aber noch ein weiteres Python-Script programmieren, um das Erstellen von Grafiken zu üben. Das Python-Modul `matplotlib` eignet sich hervorragend dafür, da es Funktionen für Liniendiagramme, Histogramme, Streudiagramme usw. zur Verfügung stellt. Außerdem haben wir als Jupyter-Notebook-Verwender den Vorteil, dass die Visualisierungen der Diagramme direkt im Notebook dargestellt werden.

Legen Sie für das neue Script wieder eine Markdown-Zelle mit der Überschrift »Stufenfunktion« an, so wie Sie es bereits vorher gemacht haben, und zusätzlich eine Code-Zelle für das folgende Script in Listing 3.2.

Noch eine Anmerkung, bevor es losgeht: Wir werden die Funktion `entscheidung` in diesem Script wiederverwenden. Damit dies gelingen kann, ist es nötig, dass Sie die Zelle mit der Funktion `entscheidung` ausführen. Erst dann kann Ihr Script in einer Code-Zelle eine Funktion in einer anderen Code-Zelle aufrufen.

```
# Import der Module für Plots
import matplotlib.pyplot as plt
# Ganz wichtig, sonst wird der Plot nicht angezeigt
%matplotlib inline

def entscheidung( summe ):
    """ Berechnung der Entscheidung zum Wert summe
    Input: summe
    Output: 1, falls summe >= 1,
            0 sonst
    """

```

```

if summe >= 1:
    return 1
else:
    return 0
#-----
# x-Werte des Graphen
x = [-1,0,0.999,1,2]
# y-Werte mit der Funktion entscheidung berechnen und mithilfe
# einer List Comprehension eine neue Liste erzeugen (siehe Anhang A)
y = [ entscheidung(i) for i in x ]
# Erzeugen des Graphen mit einer orangefarbigen Stufe und der Bezeichnung step
plt.step(x, y, color='Orange', label='step')

# Die Achsen setzen
plt.grid(True)
# Die horizontale und vertikale 0-Achse etwas dicker in Schwarz zeichnen
plt.axhline(0, color='black', lw=1)
plt.axvline(0, color='black', lw=1)

# Achsenbeschriftung und Titel
plt.xlabel('Summe')
plt.ylabel('Ergebnis')
plt.title('Stufenfunktion')

# Legendenplatzierung festlegen
plt.legend(loc='center right')

# Den Graphen anzeigen
plt.show()

```

Listing 3.2 Verwendung von »pyplot«

In Abbildung 3.8 sehen Sie das Ergebnis: eine Stufenfunktion.

Zuerst ist es natürlich nötig, `matplotlib.pyplot` zu importieren und einen passenden Namen im lokalen Namensraum zu definieren; `plt` hat sich als Name dafür eingebürgert. Die Anweisung `%matplotlib inline` ist eine sogenannte *Magic Function*. Mit ihr wird festgelegt, dass der Output der `plot`-Anweisung `inline` direkt unter der produzierenden Zelle im Notebook durchgeführt wird. (Mehr zu den Magic Functions finden Sie in Kapitel 2.)

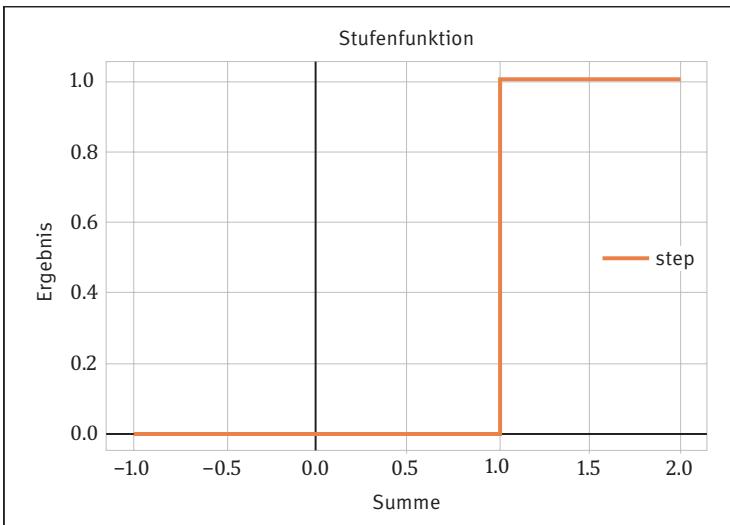


Abbildung 3.8 Stufenfunktion

Danach werden die Daten in Listen zusammengestellt. Wir verwenden eine Liste mit x- und eine mit y-Werten. Die Werte der y-Liste werden durch die Funktion `entscheidung` berechnet, indem für alle Elemente der x-Liste die Funktion aufgerufen wird. Die Funktion `plt.step` erzeugt die Stufe, wobei zum Beispiel die Farbe zum Zeichnen angegeben werden kann. Dazwischen folgen einige Detailbearbeitungen der `figure`, und schlussendlich wird die `figure` mit `plt.show()` angezeigt.

3.5 Perceptron

Durch unsere Vorüberlegungen wird es jetzt sehr einfach, auf das sogenannte *Perceptron* zuzusteuern. Das Perceptron ist eines der ältesten und einfachsten KNN-Modelle. Es entstand im Jahre 1957 und wurde von Frank Rosenblatt erfunden. Die Knoten im Perceptron werden als *Linear Threshold Unit* (LTU) bezeichnet, weil das Perceptron als Ausgabefunktion die Stufenfunktion verwendet und damit eine lineare Trennung der Eingabedaten durchführt. Zur besseren Veranschaulichung haben wir in Abbildung 3.9 eine Schautafel mit allen Bestandteilen des Perceptrons und der Berechnungsbausteine zusammengestellt. Sie ist wie ein kleiner Experimentierkasten, aus dem wir nun nach und nach Teile herausholen und besprechen werden.

Tritt Ihnen der Schweiß auf die Stirn, wenn Sie diese ganzen mathematischen Symbole sehen? Unserer Meinung nach verschleiert die kompakte und elegante mathematische Notation die Einfachheit des Modells.

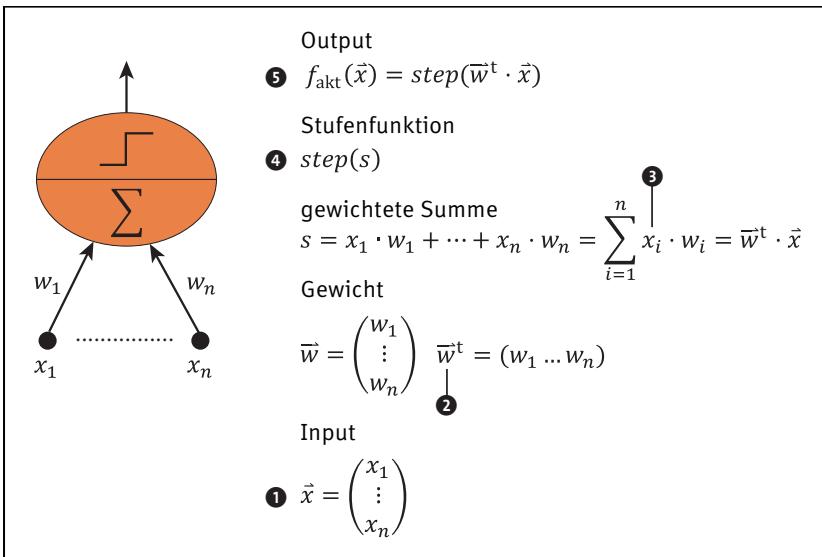


Abbildung 3.9 Perceptron – Bestandteile und Berechnungsbausteine

Wir haben damit nämlich aufgeschrieben, was wir bis jetzt bei der Planung von Frau Karotte und Herrn Lauch gemacht haben – nur allgemeiner. Um Ihnen das zu beweisen, gehen wir die einzelnen Bausteine aus der Schautafel in den nächsten Abschnitten Punkt für Punkt durch:

- ▶ Baustein ①, Input \vec{x} , betrachten wir in Abschnitt 3.6, »Punkte im Raum – Vektorrepräsentation«.
- ▶ Baustein ②, \vec{w}^t , erläutern wir in Abschnitt 3.7, »Horizontal und vertikal – Spalten- und Zeilenschreibweise«.
- ▶ Baustein ③, $\sum_{i=1}^n x_i \cdot w_i$, ist Thema von Abschnitt 3.8, »Die gewichtete Summe«.
- ▶ Baustein ④, $\text{step}(s)$, sehen wir uns in Abschnitt 3.9, »Schritt für Schritt – Stufenfunktionen«, an.
- ▶ Baustein ⑤, den Output, betrachten wir in Abschnitt 3.10, »Die gewichtete Summe reloaded«.

3.6 Punkte im Raum – Vektorrepräsentation

Los geht's: Der Input x_1 steht für die **Anwesenheit** von Frau Karotte; er kann die Werte 0 (dann ist sie leider nicht anwesend) oder 1 (sie kann glücklicherweise mithelfen) annehmen. Der Input x_2 steht für die Anwesenheit von Herrn Lauch und kann natürlich mit

denselben Werten belegt werden. Wenn zum Beispiel Frau Karotte anwesend ist und Herr Lauch noch Urlaub hat, könnten wir das so schreiben:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Die Schreibweise für die kombinierten Werte nennt man übrigens *Vektorschreibweise*. Die Anzahl der Werte nennt man die *Dimension*; in unserem Fall wäre die *Dimension* = 2, dies wird auch *zweidimensional* genannt. Der Punkt liegt also in einer zweidimensionalen Ebene, wenn man von oben aus der dritten Dimension hinuntersieht¹. Punkte lassen sich auch in Diagrammen darstellen, wie das vorherige Beispiel in Abbildung 3.10.

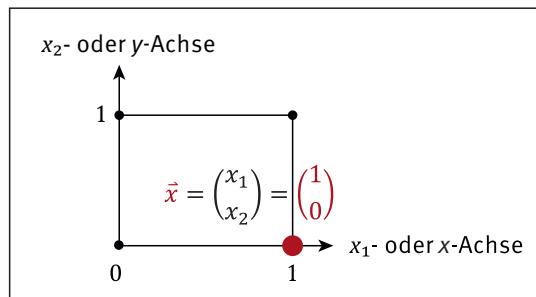


Abbildung 3.10 Punkt im kartesischen Koordinatensystem

Wir zeichnen die Werte in einem sogenannten *Koordinatensystem* ein. (Super wäre jetzt kariertes Papier, da sind schon so schöne Kästchen für das Einzeichnen der Punkte vorgedruckt.) Eine Dimension ist die x_1 - oder einfach x -Dimension, wobei die möglichen Werte dieser Dimension im Koordinatensystem auf der x -Achse dargestellt werden. Und die zweite Achse, die x_2 - oder auch y -Achse, wird darauf im rechten Winkel gezeichnet. Mit diesen zwei Angaben, x_1 und x_2 , ist die Lage eines Punktes eindeutig festgelegt.

3.6.1 Aufgabe: Werte vervollständigen

Nun sind Sie wieder an der Reihe: Zeichnen Sie die anderen möglichen Punkte für unser Planungsbeispiel mit Herrn Lauch und Frau Karotte im Koordinatensystem ein, und beschreiben Sie die Punkte mithilfe der Vektorschreibweise, so wie vorher besprochen.

Lösung:

Das Ergebnis sehen Sie in Abbildung 3.11.

¹ Eine schöne Geschichte zu den Dimensionen ist »Flatland. A Romance of Many Dimensions«, 1884, von Edwin Abbott unter dem Pseudonym A. Square.

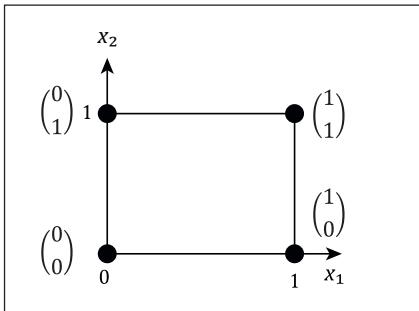


Abbildung 3.11 Alle Punkte zum Planungsbeispiel in Vektorschreibweise

Als Belohnung für den erfolgreichen Abschluss der Übung schreiben Sie ein kleines Python-Script, das für Sie das Zeichnen der Punkte übernimmt (siehe Listing 3.3). Dort sehen Sie, wie in Python ein Vektor deklariert wird, und zusätzlich setzen Sie zur Visualisierung einen *Scatter-Plot* (dt. ein *Streudiagramm*) ein. Ein *Scatter-Plot* stellt Wertepärchen als Punkte in einem Diagramm im kartesischen Koordinatensystem dar.

```
# Mathematik
import numpy as np
# Import der pyplot-Funktionen
import matplotlib.pyplot as plt
# Ganz wichtig, sonst wird der Plot nicht angezeigt
%matplotlib inline

# Die Funktion array wandelt eine Python-Liste in ein numpy-Array um
# Die x1 = x-Koordinaten
x1    = np.array([0, 0, 1, 1])
# Die x2 = y-Koordinaten
x2    = np.array([0, 1, 0, 1])
# Die Farben für die Punkte
color = np.array(['black', 'black', 'red', 'black'])
# Die Punktgröße für jeden Punkt
size  = np.array([100, 100, 500, 100])

# Den Vektor x1 mit allen x1-Koordinaten ausgeben
print('Der Vektor x1: ', x1)

# Die Achsen setzen
plt.grid(True)
```

```
# Den Plot zeichnen für die x1- und x2-Koordinaten, Farbe und Größe
plt.scatter(x1, x2, c=color, s=size)

# Die Achsen setzen
plt.xlabel('x1')
plt.ylabel('x2')

# Die Achsenabschreibungen setzen
plt.xticks([0.0,1.0])
plt.yticks([0.0,1.0])

# Endlich das Diagramm ausgeben
plt.show()
# Ausgabe:
Der Vektor x1: [0 0 1 1]
```

Listing 3.3 Scatter-Plot

Das Ergebnis sehen Sie in Abbildung 3.12.

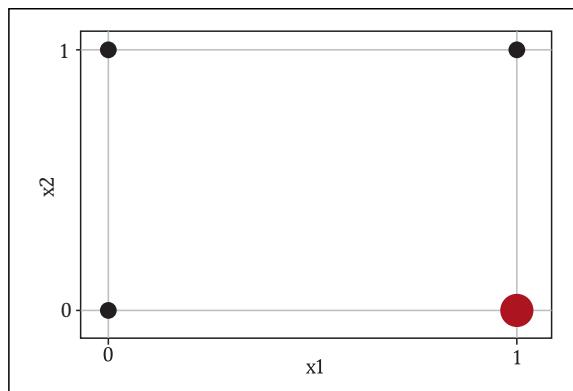


Abbildung 3.12 Scatter-Plot für die Planungspunkte

Wir starten im Script, indem wir das Modul `numpy` für mathematische Operationen importieren. Es eignet sich speziell für das Rechnen mit Vektoren und anderen mathematischen Konstrukten, die wir später noch besprechen werden. Danach importieren wir das `matplotlib.pyplot`-Modul und vergeben dafür den Namen `plt`. Anschließend folgt ein erklärungsbedürftiger Abschnitt im Programm. In den vorherigen Erläuterungen haben wir Punkte mit einem Vektor dargestellt, der aus den zwei Dimensionen x_1 und x_2 besteht. Der Einsatz der `scatter`-Funktion erfordert es, dass die Koordinaten der einzelnen Vektoren in zwei `np.array`-Objekte aufgeteilt werden, also pro Dimension ein Array

und pro Punkt ein Eintrag im Array. Zur Kontrolle geben wir das Array für die x_1 -Dimension mit der `print`-Funktion aus.

Im weiteren Verlauf setzen wir für den Plot die gewünschten Eigenschaften, zum Beispiel die Achsenbeschriftungen mit `plt.xlabel('x1')`, `plt.ylabel('x2')` und die Achsen-einteilungen mit `plt.xticks()` und `plt.yticks()`. Am Ende des Scripts wird der Plot mittels `plt.show()` ausgegeben und gezeichnet.

3.6.2 Aufgabe: Den Iris-Datensatz als Scatter-Plot ausgeben

Die folgende Aufgabe dient dazu, Ihnen zu zeigen, wie Sie Daten aus einer Datei in Python importieren, zeilenweise verarbeiten und für die Visualisierung verwenden können. Die ersten Schritte, die das Einlesen betreffen, müssen wir natürlich detailliert erläutern, da wir dieses Thema noch nicht besprochen haben.

Als Datenmaterial für den Plot haben wir den berühmten *Iris-Datensatz* auserkoren, der einen typischen Testfall für Klassifikationstechniken darstellt. In diesem Datensatz befindet sich eine Menge von 150 Beobachtungen von je vier Eigenschaften von Schwertlilien. Die vier Eigenschaften sind die Breite und die Länge des Kelchblatts (*Sepalum*) sowie die Breite und die Länge des Kronblatts (*Petalum*), gemessen in Zentimeter. Weiterhin ist zu jeder vorhandenen Eigenschaftskombination die Art von Schwertlilie angegeben. Das sind *Iris-setosa*, *Iris-virginica* oder *Iris-versicolor*.

Sepal Length	Sepal Width	Petal Length	Petal Width	Class
5.1	3.5	1.4	0.2	Iris-setosa
7.0	3.2	3.5	1.0	Iris-versicolor
6.3	3.3	6.0	2.5	Iris-virginica
...

The diagram shows a blue Iris flower with its parts labeled: 'petal' pointing to the upper petal, 'sepal' pointing to the lower sepal, and 'Kronblatt' (Crown leaf) pointing to the three upper petals. Brackets above the table group 'Sepal Length' and 'Sepal Width' under 'Kelchblatt' (Sepal), and 'Petal Length' and 'Petal Width' under 'Kronblatt' (Petal).

Abbildung 3.13 Blattmaße für Schwertlilien (© Kaggle)

Ihre Aufgabe besteht nun darin, einen *Scatter-Plot* mit den Python-Mitteln zu zeichnen, die wir bisher besprochen haben. Natürlich müssen wir Ihnen noch verraten, woher Sie den Datensatz für die Schwertlilien beziehen können. Besuchen Sie die Webseite zum Buch, und laden Sie von dort das Material zum Buch und damit den Datensatz *iris.csv* herunter (Verzeichnis *Kapitel_03*). Gehen Sie nun wie folgt vor:

1. Bevor Sie mit dem Coding beginnen, laden Sie bitte das *iris.csv*-File aus unserem Download in den Sitzungsspeicher Ihrer Colab-Sitzung hoch, damit die Datei dann ohne Pfadangabe eingelesen werden kann.
2. Lesen Sie zunächst das *iris.csv*-File ein. Verwenden Sie dazu die Anweisung `with open("iris.csv", "r") as fobj:`, die Ihnen ein File-Handle zum Auslesen (Parameter "`r`") zurückgibt. Mit der Anweisung `for line in fobj:` können Sie den Datensatz Zeile für Zeile abarbeiten, und nachdem Sie alle Zeilen verarbeitet haben, wird das File-Handle `fobj` automatisch geschlossen.
3. Erzeugen Sie drei Python-Listen, `x1`, `x2` und `colors`, um die Daten für Sepal Length, Sepal Width und die Farben für die Datenpunkte aufzunehmen. Um diese Daten ermitteln zu können, benötigen Sie noch den Spaltenaufbau des Datensatzes pro Zeile, den Sie in Abbildung 3.13 sehen können.

Dabei erfolgen die Längenangaben in den Spalten in Zentimeter, und die Spalte `Class` kann die folgenden Werte annehmen:

- Iris-setosa
- Iris-versicolor
- Iris-virginica

Die einzelnen Spalten sind durch ein Komma getrennt.

4. Nutzen Sie `matplotlib.pyplot` für die Erzeugung des Scatter-Plots.

Lösung:

```
# Mathematik
import numpy as np
# Import der pyplot-Funktionen
import matplotlib.pyplot as plt
# Ganz wichtig, sonst wird der Plot nicht angezeigt
%matplotlib inline

# x1 sind die Koordinaten der x-Achse, x2 die der y-Achse
x1    = []
x2    = []
# Farben für die Datenpunkte
colors = []
```

```

# Mapping der Schwertlilien zu Farben mit
# einem python dictionary
iris_colors = { 'Iris-setosa' : 'red',
                 'Iris-versicolor' : 'green',
                 'Iris-virginica' : 'blue'
             }
# File-Inhalt einlesen
# Voraussetzung: Das File iris.csv ist in den Sitzungsspeicher hochgeladen
# File Handle fobj wird automatisch geschlossen
with open("iris.csv", "r") as fobj:
    # Den Datensatz zeilenweise verarbeiten
    for line in fobj:
        # Split in einzelne Worte
        words = line.rstrip().split(",")
        # Leerzeilen auslassen
        if len(words) != 5:
            continue
        # Sepallength
        x1.append(float(words[0]))
        # SepalWidth
        x2.append(float(words[1]))
        # Farbe
        colors.append(iris_colors[words[4]])
# Gitter im Scatter-Plot zeichnen
plt.style.use('seaborn-whitegrid')
# Achsenbeschriftung und Titel
plt.xlabel('Sepal Length')
plt.ylabel('Sepal Width')
plt.title('Scatter Plot')
# Den Plot anzeigen
plt.scatter(np.array(x1), np.array(x2), color=colors )
# Plot anzeigen
plt.show()

```

Listing 3.4 »Sepal Length« und »Sepal Width« als Scatter-Plot

Anfangs importieren wir wieder die benötigten Module für Plot und Arrays. Danach wird der Datensatz zeilenweise gelesen und verarbeitet. Mit der Funktion `line.rstrip()` wird eine Kopie des Strings `line` erzeugt, aus dem die Leerzeichen am Ende entfernt wurden. Mittels `split(",")` wird dann der String in seine Einzelwerte zerlegt, immer am Komma getrennt, und die Einzelwerte werden zum Array `words` zusammengefügt. Die

ersten zwei Werte im `words`-Array, `words[0]` und `words[1]`, definieren die Koordinaten für den Plot. Der letzte Wert im Array an der Position 4 ist der Name der Lilie, die für die Ermittlung der Farbe verwendet wird. Mit dem Python-Dictionary `iris_colors` findet die Zuordnung zwischen den Namen und den Farben statt. Danach setzen wir noch die Ausgabe eines Rasters im Plot, übergeben die Daten an den Scatter-Plot und bekommen damit die Ausgabe aus Abbildung 3.14.

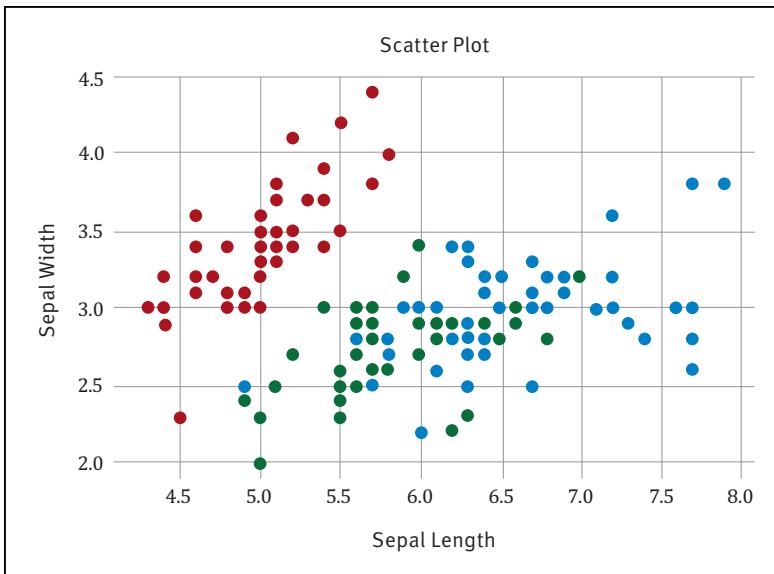


Abbildung 3.14 Die Lilien als Scatter-Plot mit den Koordinaten »Sepal Length« und »Sepal Width«

3.7 Horizontal und vertikal – Spalten- und Zeilenschreibweise

\vec{w} wird als *Gewichtsvektor* bezeichnet. Wenn wir vom *Lernen* in neuronalen Netzen sprechen, dann ist damit die **Anpassung der Gewichte** gemeint. Um zu verstehen, was ein Gewicht ist, stellen Sie sich eine Verbindung zwischen zwei Neuronen vor. Das eine Neuron, nennen wir es A, sendet ein Signal an das zweite Neuron, B. Wie stark das Signal in B ankommt, steuert das Gewicht der Verbindung. Wenn das Gewicht einen Wert zwischen 0 und 1 hat, dann wird die Signalstärke verkleinert; wenn der Wert größer als 1 ist, dann wird das Signal verstärkt, und wenn der Wert kleiner als 0 ist, dann wird das Signal negativ. Somit regelt das Gewicht die Signalstärke, die beim Neuron B ankommt. Details dazu sehen Sie in Kapitel 4, wenn wir das *Perceptron-Lernen* besprechen.

Zusätzlich zum Inputvektor, in dem wir die Werte übereinandergeschrieben haben (die sogenannte *Spaltenschreibweise*), haben wir den Gewichtsvektor auch in *Zeilenschreib-*

weise angeführt. Dabei werden die Werte in einer Zeile nebeneinander geschrieben. Damit man den Unterschied zu einem Spaltenvektor sehr einfach erkennt, wird dem \vec{w} einfach ein t verpasst, und damit wissen wir, dass es in Zeilenschreibweise dargestellt wird. »Verpassen« ist jetzt ein wenig unmathematisch, nennen wir es also *transponieren*. Der Grund, warum wir den ganzen Hokuspokus aufführen, ist, dass wir damit einfach die Multiplikation zwischen den Input-Werten und den Gewichten mathematisch beschreiben und damit die Berechnung in Python optimal implementieren können, nämlich mit $\vec{w}^t \cdot \vec{x}$ anstatt mit $w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n$, wobei die Anzahl der Komponenten in \vec{w} und \vec{x} gleich sein muss. Das ist doch viel netter, oder? Jedenfalls kurz und bündig.

So führen Sie also die Multiplikation zwischen den Vektoren durch: w_1 mal x_1 plus w_2 mal x_2 plus ... bitte selbstständig vervollständigen.

$$\vec{w}^t \cdot \vec{x} = (w_1 w_2 \dots w_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n$$

3.7.1 Aufgabe: Das Skalarprodukt mithilfe von NumPy ermitteln

Python sieht die Multiplikation von Vektoren mithilfe von Funktionen des Moduls `numpy` vor, genauer gesagt mit der Funktion `dot`. Die Multiplikation wird als *Dot-Product*, *Skalarprodukt* oder *inneres Produkt* bezeichnet. Sie macht aus zwei Sequenzen von Zahlen eine Zahl, wie Sie zuvor gesehen haben.

Fun Facts

1. Die Multiplikation zwischen den Vektoren heißt *Skalarprodukt*, weil das Ergebnis ein Skalar, also ein Wert, ist und kein Vektor.
2. Der englische Name *dot product* kommt daher, dass bei der Multiplikation zwischen den Vektoren ein Punkt steht.

Nun wird es Zeit, das Skalarprodukt mithilfe von Python zu berechnen. Dafür benötigen Sie wieder eine Code-Zelle in Ihrem Jupyter Notebook und, wenn Sie wollen, auch eine Markup-Zelle für eine Überschrift.

Bevor Sie mit der Implementierung beginnen, lesen Sie sich das folgende Python-Script Zeile für Zeile genau durch, und analysieren Sie die Berechnungen. Nach Ihrer Analyse werden wir Sie nach **zwei** Möglichkeiten fragen, um mithilfe von NumPy das innere Produkt zu berechnen.

```
# Die Multiplikation von Vektoren mit numpy
# Dot Product, Skalarprodukt, inneres Produkt
# macht aus zwei Sequenzen von Zahlen eine Zahl (algebraisch)
```

```

# Es ist der Cosinus des Winkels zwischen zwei Vektoren (geometrisch),
# multipliziert mit deren Längen
# Mathematik
import numpy as np

# numpy array erzeugen
x = np.array([0,1])
# numpy array erzeugen
w = np.array([0.5,0.7])

# Vektor x ausgeben
print("x =", x)
# Vektor w ausgeben
print("w =", w)

# numpy arrays sind keine Matrizen, und
# die Operatoren *, +, -, / funktionieren Element für Element
print("w*x =", w*x)

# Inneres Produkt. Man muss den Vektor nicht transponieren,
# das erledigt numpy für uns
print("np.dot(w,x) =", np.dot(w,x))

# Alternative Syntax zum Dot Product
print("w.dot(x) =", w.dot(x))
# Ausgabe:
x = [0 1]
w = [0.5 0.7]
w*x = [0.  0.7]
np.dot(w,x) = 0.7
w.dot(x) = 0.7

```

Listing 3.5 Skalarprodukt mit Python

Welche zwei Möglichkeiten würden Sie für die Ermittlung des inneren Produkts mit NumPy in Betracht ziehen? Natürlich haben Sie die NumPy-dot()-Varianten als die passenden Kandidaten gefunden!

Wir machen weiter mit unserer Erläuterung der Elemente im Baukasten. Das dritte Element aus dem Baukasten ist die Darstellung der *gewichteten Summe*.

3.8 Die gewichtete Summe

Die Mathematiker leihen sich für die Summe der $w_i \cdot x_i$ ein Symbol aus dem griechischen Alphabet – das große Sigma: Σ . Außerdem werden einige zusätzliche »Mascherl« am Sigma befestigt:

$$\sum_{i=1}^n$$

Die Mascherl bedeuten, dass die Laufvariable i von einem Anfangswert bis zu einem Endwert nacheinander alle Werte annimmt. (Man sagt, sie *durchläuft* den Bereich in ganzen Schritten). In unserem Beispiel läuft sie vom Anfangswert 1 bis zum Endwert n . Nehmen wir an, dass $n = 3$, dann sind das die Schritte 1, 2 und 3. Mit diesem Wissen können Sie die Summe aller $w_i \cdot x_i$ – nennen wir sie s – jetzt sehr kompakt schreiben:

$$s = \sum_{i=1}^n w_i \cdot x_i$$

Damit ergibt sich der dritte Baustein:

$$s = \sum_{i=1}^n w_i \cdot x_i = \vec{w}^t \cdot \vec{x} = w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n$$

Weiter geht es zum vierten Baustein im Baukasten, der Stufenfunktion. Diese haben wir zwar bereits mehrfach besprochen, jedoch wollen wir sie hier in einer mathematischen Darstellung präsentieren.

3.9 Schritt für Schritt – Stufenfunktionen

Für die Stufenfunktion $step(s)$ definieren wir die folgende Funktion:

$$step(s) = \begin{cases} 0, & \text{falls } s < \theta \\ 1, & \text{falls } s \geq \theta \end{cases}$$

Egal, mit welchem Wert Sie die Stufenfunktion aufrufen, das Ergebnis ist entweder 0 oder 1, jedoch ist es vom Schwellenwert *Theta*, dem griechischen Buchstaben θ , abhängig, wann der Sprung von 0 auf 1 stattfindet. Nachdem der Wert der Schwelle nicht mehr fixiert ist, kann er auch durch Lernen verändert werden.

Nachdem wir die Stufenfunktion verallgemeinert haben, können wir uns das fünfte Element aus dem Baukasten ansehen, die Berechnung der Funktion $f_{\text{akt}}(\quad)$.

3.10 Die gewichtete Summe reloaded

Hier folgt noch eine Überlegung zu der gewichteten Summe und der Stufenfunktion. Das geht doch einfacher zu schreiben als diese komplizierte Unterscheidung zum Beispiel bei der Heaviside-Funktion: »falls größer θ , dann das, sonst das ...«

$$w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n \geq \theta$$

Wenn wir die Ungleichung nun umformen, also einfach das θ auf die andere Seite bewegen, dann erhalten wir den erweiterten Gewichtsvektor, der um die eine Dimension für den Schwellenwert erweitert wird. Diese neue Dimension sollte am Index 0 eingefügt werden, damit der ursprüngliche Vektor mit seinen Indizes erhalten bleibt:

$$w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n - \theta \geq 0$$

Noch ein kleiner Trick aus der Zauberkiste: Damit wir den Schwellenwert an der Stelle 0 einfügen können, nennen wir $-\theta$ nun einfach w_0 . Ist das nicht super? Damit können wir Folgendes schreiben:

$$w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n + w_0 \geq 0$$

Damit haben wir eine schöne einheitliche Form gefunden, um wieder $\vec{w}^t \cdot \vec{x}$ zu schreiben. Natürlich muss der Vektor \vec{x} ebenfalls um ein Element verlängert werden. Das hat aber einen sehr einfachen Wert. Können Sie sich vorstellen, welchen? Kurz nachdenken ...

Das können wir mit folgender Abbildung 3.15 beantworten, in der Sie sehen, dass das Neuron für den Input x_0 immer den Wert 1 hat. Es wird als *Bias-Neuron* bezeichnet.

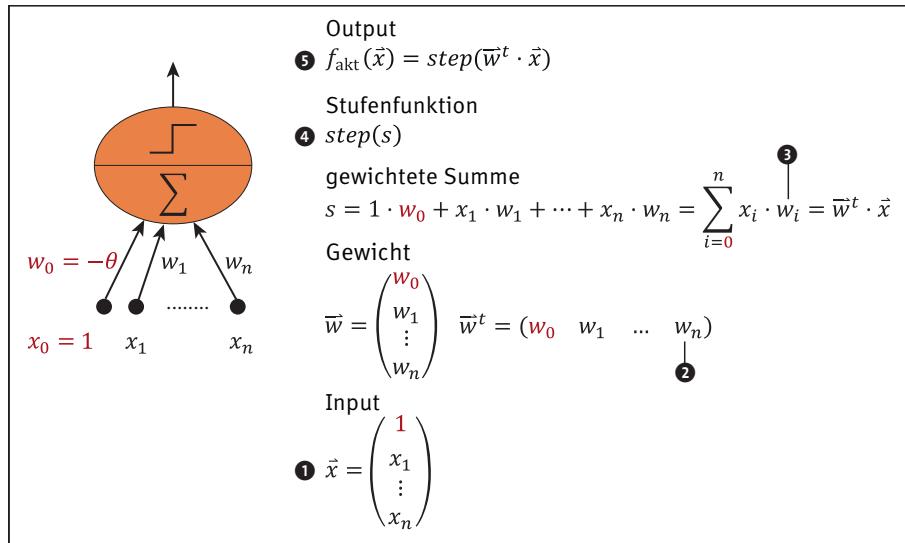


Abbildung 3.15 Bausteine des Perceptrons

Falls Sie das Nachdenken beendet haben, finden Sie die Erweiterungen zu den Vektoren und der gewichteten Summe rot markiert in Abbildung 3.15.

3.11 Alles zusammen

Es wird Zeit, ein kleines Python-Programm zu schreiben, um unsere Erkenntnisse zum Perceptron in Programmcode umzuwandeln. Sehen Sie sich dazu zuerst die folgenden Erläuterungen und Codezeilen an, und übertragen Sie diese in eine neue Code-Zelle in Ihrem Jupyter Notebook, um den gezeigten Output zu produzieren.

Als Input verwenden wir die An- bzw. Abwesenheiten von Herrn Lauch und Frau Karotte. Jede Zeile in Tabelle 3.4 stellt einen Input-Vektor mit drei Werten dar: 1 steht für das Bias-Neuron, für die Anwesenheit Frau Karotte und für die Anwesenheit Herr Lauch.

Bias-Neuron	Frau Karotte	Herr Lauch	-->	Montag	Input-Vektor
1	0	0	-->	0	1
1	0	1	-->	1	2
1	1	0	-->	1	3
1	1	1	-->	1	4

Tabelle 3.4 Der erweiterte Input-Vektor für das Planungsproblem

Im Beispiel $(1, 0, 1)$ steht 1 für das Bias-Neuron; 0 für Frau Karotte bedeutet »nicht anwesend«; 1 für Herr Lauch bedeutet »anwesend«, und das ergibt insgesamt eine 1 für eine akzeptable Planungssituation. Insgesamt haben wir vier Input-Vektoren, die am einfachsten mithilfe eines zweidimensionalen Arrays verwaltet werden. Weil die Input-Vektoren schon mit x bezeichnet werden, verwenden wir ein großes X als Bezeichnung für das Array. Die gewünschten Outputs, die Sie in der Tabelle in der Spalte »Montag« finden, fassen wir als Vektor y zusammen. In der folgenden Implementierung haben wir die Heaviside-Funktion für die Stufenfunktion implementiert, also mit dem Schwellenwert 0, da wir den Schwellenwert zuvor in die gewichtete Summe integriert haben.

Des Weiteren haben wir den Aspekt der **Fehlerberechnung** ergänzt. Für jeden Input wird der berechnete Output mit dem gewünschten Output verglichen. Weil das Perceptron nur 0 oder 1 aufgrund der Heaviside-Funktion ausgeben kann und der gewünschte Output nur 0 oder 1 ist, kann die Differenz aus dem gewünschten und dem errechneten Wert nur die Werte $-1, 0, 1$ annehmen. Um die Einzelfehler summieren und damit die Gesamtaussage über die Ermittlungsgenauigkeit des Perceptrons für die Daten liefern zu können, wenden wir in Listing 3.6 noch den Betrag auf den Einzelfehler an. Ansonsten würde ein Fehler von -1 den Gesamtfehler verringern.

```

# Mathematik
import numpy as np
# Plot
import matplotlib.pyplot as plt

# 3-dimensionaler Input = Bias-Neuron, Fr. Karotte, Hr. Lauch
# 4 Inputvektoren
X = np.array([
    [1,0,0],
    [1,0,1],
    [1,1,0],
    [1,1,1]])
# Die 4 gewünschten Ergebniswerte
y = np.array([0,1,1,1])

# Heaviside-Funktion
def heaviside( summe ):
    """ Berechnung der Entscheidung zum Wert summe
    Input: summe
    Output: 1, falls summe >= 0,
            0 sonst
    """
    if summe >= 0:
        return 1
    else:
        return 0

# Perceptron-Berechnung (Forward Path)
def perceptron_eval(X,y):
    """ Perceptron-Berechnung
    Input: X, Inputvektor
           y, der gewünschte Output
    Output: Der Gesamtfehler, d. h. Summe aus dem Betrag der Differenz
            von errechnetem und gewünschtem Output
    """
    # Der Gesamtfehler
    gesamtfehler = 0;
    # Die Gewichte so wählen, dass das OR-Problem gelöst werden kann
    w = np.array([-1,1,1])

```

```

# Index i und Element x Ermittlung vom Array X
for i, x in enumerate(X):
    # x = Zeile für Zeile verwenden
    # Inneres Produkt zwischen x und w
    summe = np.dot(w,x)
    ergebnis = heaviside(summe)
    # Fehler
    fehler = np.abs(ergebnis - y[i])
    # Gesamtfehler
    gesamtfehler += fehler
    # Ausgabe
    print("Fr. Karotte = {}, Hr. Lauch = {},
          gewünschtes Ergebnis = {}, errechnetes Ergebnis = {}, Fehler = {}".
          format(x[1], x[2], y[i], ergebnis, fehler))
# Gesamtfehler pro Epoche über ganzen Trainingsdatensatz
return gesamtfehler
-----
# Core Function zum Auswerten des Inputs
gesamtfehler = perceptron_eval(X,y)
print("Gesamtfehler = %1d" % (gesamtfehler))
# Ausgabe:
Fr. Karotte = 0, Hr. Lauch = 0, gewünschtes Ergebnis = 0, errechnetes Ergebnis
= 0, Fehler = 0
Fr. Karotte = 0, Hr. Lauch = 1, gewünschtes Ergebnis = 1, errechnetes Ergebnis
= 1, Fehler = 0
Fr. Karotte = 1, Hr. Lauch = 0, gewünschtes Ergebnis = 1, errechnetes Ergebnis
= 1, Fehler = 0
Fr. Karotte = 1, Hr. Lauch = 1, gewünschtes Ergebnis = 1, errechnetes Ergebnis
= 1, Fehler = 0
Gesamtfehler = 0

```

Listing 3.6 Lösung für das sehr einfache Personalplanungsproblem

Der Fehler ist in unserem Beispiel natürlich 0, da wir den Gewichtsvektor optimal gewählt haben. Experimentieren Sie mit den Gewichten im Gewichtsvektor w , sehen Sie sich die möglichen Fehler an, und versuchen Sie, diese anhand der vorher besprochenen und implementierten Berechnungen nachzuvollziehen.

3.12 Aufgabe: Roboterschutz

Wir haben noch eine Aufgabe für Sie vorbereitet, die darin besteht, einen Roboter vor dem Absturz zu schützen. Dieser kleine Roboter fährt vor sich hin und führt nichts Böses im Schilde. Doch die Umgebung ist ihm nicht wohlgesonnen, und große gefährliche Löcher meinen es nicht gut mit ihm. Glücklicherweise besitzt der Roboter zwei Hell-Dunkel-Sensoren, links und rechts vorn montiert, die die Bodenhelligkeit vor ihm bestimmen können.

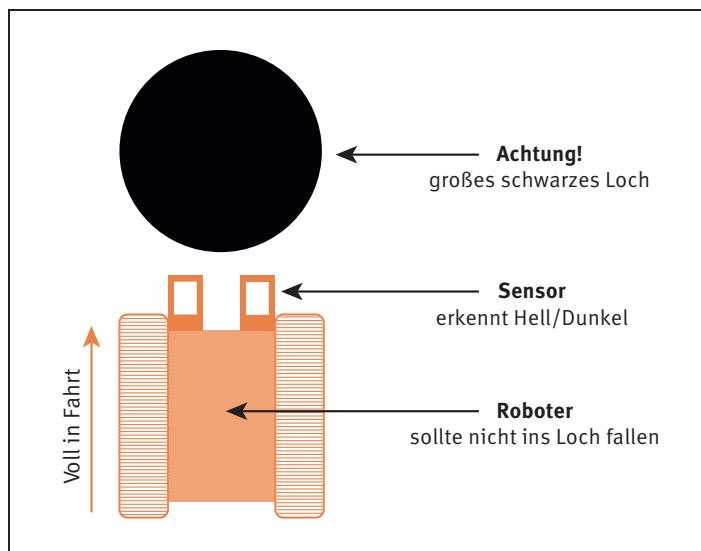


Abbildung 3.16 Einfache Robotersensorik

Wenn ein Sensor **Hell** erkennt, dann liefert dieser den Wert 0, wenn er **Dunkel** erkennt, dann liefert der Sensor 1. Um das Leben des Roboters zu schützen, wäre es schön, wenn Sie ein einfaches Perceptron entwickeln, das erkennt, ob sich ein Loch direkt vor dem Roboter befindet oder ob eine ungefährliche Fahrt möglich ist.

Um das Script zu implementieren, legen Sie bitte eine neue Code-Zelle in Ihrem Jupyter Notebook an und entwickeln dort Ihre Lösung. Orientieren Sie sich bei dem Beispiel an Listing 3.6.

Lösung:

Wenn wir uns die Helligkeitsaussagen der Sensoren ansehen, dann kann mal der linke Sensor anschlagen und mal der rechte oder beide oder gar keiner. Die Kombinationen der Helligkeitsaussagen veranschaulichen wir uns mithilfe von Tabelle 3.5.

Sensor links	Sensor rechts	Loch
0	0	0
1	0	0
0	1	0
1	1	1

Tabelle 3.5 Locherkennung mithilfe der Sensorwerte

Wir haben in Listing 3.7, so wie von uns empfohlen, die Implementierung aus Listing 3.6 verwendet und den gewünschten Output $y = \text{np.array}([0,0,0,1])$ und den Gewichtsvektor auf $w = \text{np.array}([-2,1,1])$ angepasst, sodass das gewünschte Ergebnis aus Tabelle 3.5 berechnet wird. Zusätzlich haben wir den Text der Ausgabe an die Aufgabenstellung angepasst:

```
# Mathematik
import numpy as np
# Plot
import matplotlib.pyplot as plt
# 3-dimensionaler Input = Bias-Neuron, Sensor links, Sensor rechts
# 4 Inputvektoren
X = np.array([
    [1,0,0],
    [1,0,1],
    [1,1,0],
    [1,1,1]])
# Die 4 gewünschten Ergebniswerte
y = np.array([0,0,0,1])
# Heaviside-Funktion
def heaviside( summe ):
    """Berechnung der Entscheidung zum Wert summe
    Input: summe
    Output: 1, falls summe >= 0,
            0 sonst
    """
    if summe >= 0:
        return 1
    else:
        return 0
```

```

# Perceptron-Berechnung (Forward Path)
def perceptron_eval(X,y):
    """ Perceptron-Berechnung
    Input: X, Inputvektor
           y, der gewünschte Output
    Output: Der Gesamtfehler, d. h. Summe aus dem Betrag der Differenz
            von errechnetem und gewünschtem Output
    """
    # Der Gesamtfehler
    gesamtfehler = 0;
    # Die Gewichte so wählen, dass das Roboter-Problem gelöst werden kann
    w = np.array([-2,1,1])
    # Index i und Element x Ermittlung vom Array X
    for i, x in enumerate(X):
        # x = Zeile für Zeile verwenden
        # Inneres Produkt zwischen x und w
        summe = np.dot(w,x)
        ergebnis = heaviside(summe)
        # Fehler
        fehler = np.abs(ergebnis - y[i])
        # Gesamtfehler
        gesamtfehler += fehler
        # Ausgabe
        print("Sensor L = {}, Sensor R = {}, gewünschtes Ergebnis = {}",
              "{}", gesamtfehler)
        format(x[1], x[2], y[i], ergebnis, fehler))
    # Gesamtfehler pro Epoche über ganzen Trainingsdatensatz
    return gesamtfehler
#-----
# Core Function zum Auswerten des Inputs
gesamtfehler = perceptron_eval(X,y)
print("Gesamtfehler = %d" % (gesamtfehler))
# Ausgabe:
Sensor L = 0, Sensor R = 0, gewünschtes Ergebnis = 0, errechnetes Ergebnis = 0,
Fehler = 0
Sensor L = 0, Sensor R = 1, gewünschtes Ergebnis = 0, errechnetes Ergebnis = 0,
Fehler = 0
Sensor L = 1, Sensor R = 0, gewünschtes Ergebnis = 0, errechnetes Ergebnis = 0,
Fehler = 0

```

```
Sensor L = 1, Sensor R = 1, gewünschtes Ergebnis = 1, errechnetes Ergebnis = 1,  
Fehler = 0  
Gesamtfehler = 0
```

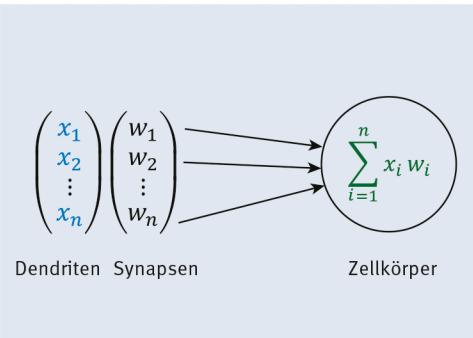
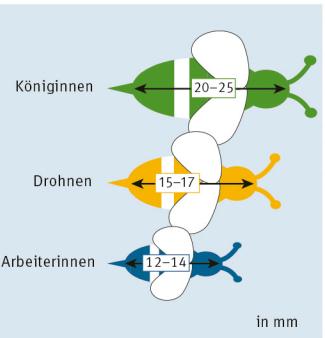
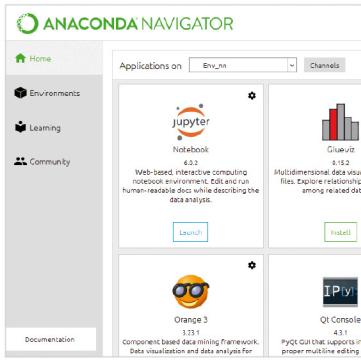
Listing 3.7 Die Perceptron-Steuerung für den Roboter

3.13 Zusammenfassung

Das Perceptron gehört nun Ihnen: Sie haben den Aufbau und die Berechnungen im Perceptron theoretisch und praktisch durchgeführt. Dabei sind so wichtige Dinge wie die gewichtete Summe, Schwellenwertberechnungen, Grafiken und grafische Analysen vorgekommen. Zugegebenermaßen haben wir bis jetzt ausschließlich Auswertungen umgesetzt und das Lernen des Perceptrons noch nicht berücksichtigt. Das werden wir im folgenden Kapitel sofort nachholen.

3.14 Referenzen

- ▶ Rosenblatt, F.: The Perceptron: A Probabilistic Model for Information Storage and Organization in The Brain. 1958. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.588.3775>
- ▶ <https://archive.ics.uci.edu/dataset/53/iris>



Neuronale Netze stehen im Mittelpunkt, wenn es um künstliche Intelligenz geht. Erfahren Sie, welche genialen Ideen dahinter stecken, indem Sie selbst programmieren. Alle Kenntnisse, die Sie brauchen, bekommen Sie genau hier.

Ihr Einstieg in die KI-Welt

Dieser Grundkurs erläutert Ihnen alle wichtigen Konzepte und bietet eine geeignete Lernumgebung. Los geht's mit einfachen neuronalen Netzen und schnellen ersten Erfolgen. Python-Crashkurs und Mathe-Grundlagen inklusive!

Praxis, Praxis, Praxis

Sobald das theoretische Fundament steht, entwickeln Sie Ihre eigenen Netze. Sie erstellen verschiedene Netztypen und wenden sie auf typische Aufgaben wie Bild- oder Handschriftenerkennung an.

So lernen Maschinen

Übung macht den Meister! Das gilt auch für die Netze. Direkt am Beispiel lernen Sie, wie Sie Ihre Netze geschickt trainieren. Dazu gehört: Fallstricke umgehen, Werkzeuge einsetzen und die Trefferquote Ihrer Modelle erhöhen.

 Mit allen Codebeispielen zum Download

Auf einen Blick

Grundlagen

Lernumgebung installieren
Einfache neuronale Netze aufbauen und trainieren
Wichtige mathematische Begriffe

Einblicke

Überwachtes und unüberwachtes Lernen
Reinforcement Learning
Verschiedene Netztypen
Backpropagation
Deep Learning und LLMs
Transformer programmieren

Werkzeuge

Jupyter Notebook
TensorFlow mit Keras
Python-Bibliotheken für Data Science



Dr. Joachim Steinwendner wendet maschinelles Lernen als Forschungsfeldleiter »GeoHealth Analytics« an der Fernfachhochschule Schweiz an. **Dr. Roland Schwaiger** ist Softwareentwickler, Autor und Experte für neuronale Netze und evolutionäre Algorithmen. Beide sind erfahrene Brückenbauer zwischen Theorie und Praxis.

