

Python in der Wildnis: Maschinelles Lernen mit Tensorflow

Sven Schiffner

August 27, 2019

Inhaltsverzeichnis

Einleitung

Unser Workshop

Handschrifterkennung

Methoden des Maschinellen Lernens

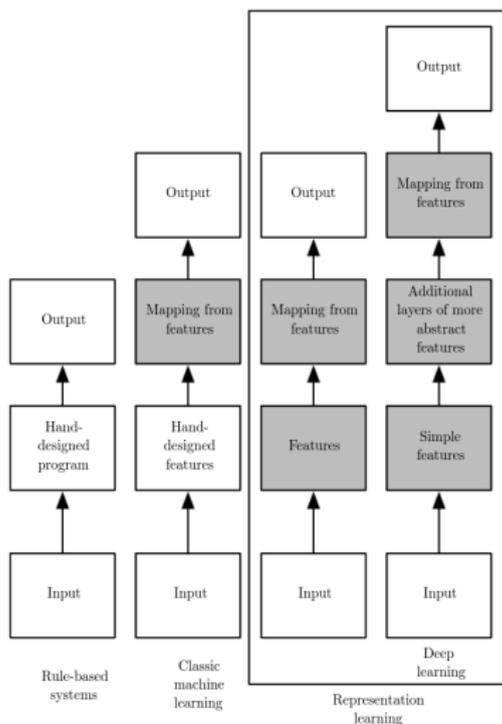
Einleitung

Lernen

Bedeutung nach Duden:

- a) sich Wissen, Kenntnisse aneignen
- b) sich, seinem Gedächtnis einprägen
- c) **Fertigkeiten erwerben**
- d) im Laufe der Zeit [durch Erfahrungen, Einsichten] zu einer bestimmten Einstellung, einem bestimmten Verhalten gelangen

Überblick: Maschinelles Lernen (ML)



- Rule-based systems: "übliche Programmierung"
- Classic machine Learning: Anfänge von ML, sehr zeitaufwendig bei Datenaufbereitung
- Representation Learning: Das wollen wir heute machen
- Deep Learning: Vielleicht lieber nicht zu viel auf eurer CPU

Figure: Übersicht ML (Quelle: <http://www.deeplearningbook.org/contents/intro.html>)

Tensorflow - Heransgehensweise

- Tensor - **Matrix**
- keine Betrachtung der biologischen Seite (zu zeitaufwendig)
- Erklärung mit (hoffentlich) bekannten Grundlagen
- Spezialfall: Handschrifterkennung für einzelne Zahlen (MNIST)
→ "Hello World!" des Maschinellen Lernens

Maschinelles Lernen, wie wir es verstehen wollen

- **Modell:** Aufbau der Berechnungen zur Lösung des gestellten Problems
- Ziel meist nicht unmittelbar erreichbar \Rightarrow Ersatzzielfunktion
 \Rightarrow Minimierung der sog. Fehlerfunktion (Ersatzzielfunktion)
- **Optimierungsproblem:** Optimierung der Modellparameter im Bezug auf die Fehlerfunktion für die diskreten Punkte des Datensatzes

Unser Workshop

Eure Vorstellungen/Wünsche

- Tensorflow in Python
- Verarbeitung von Daten
- Laden von Daten
- erste Berührungen
- keine Einführung
- Interesse

Meine Vorstellungen/Wünsche

- Eure Mitarbeit
- Interesse (auch geheuchelt)
- Spass zusammen

- Tiefe durch euch bestimmt
- für Fragen immer offen

Handschrifterkennung

Gedankenexperiment

- alter Physikkollege → 10^9 ganzzahlige Messwerte auf Papier
- pro Kästchen des karierten Papiers nur eine Zahl
- gute, saubere, aber sehr spezielle Handschrift
- zufällig ein Datensatz mit 60000 seiner handgeschriebenen Zahlen plus Label vorhanden
- eigene Benutzung der Messwerte
- Zeitdruck vom Prof, aber viel andere Arbeit

Nähere Betrachtung

- Versagen üblicher OCR durch spezielle Handschrift
- manuell zu zeitaufwendig
- viel Rechenleistung vorhanden (Bier an Informatiker)
- Datensatz zum Training zufällig vorhanden
- Segmentierung durch Kästchen (automatisch durch Institusscanner)
- Machine Learning: Der neue Hype

Unsere Werkzeuge

Tensorflow:

- spezielle Bibliothek von Google für das maschinelle Lernen
- sowohl Low-Level- als auch High-Level-Programmierung möglich
- gute GPU-Unterstützung und auch TPU-Nutzung möglich

```
import tensorflow as tf
```

Matplotlib:

- sehr gute Plot-Bibliothek mit NumPy-Unterstützung

```
import matplotlib.pyplot as plt
```

Datensatz



Figure: Beispiele von MNIST(Quelle:
<https://upload.wikimedia.org/wikipedia/commons/2/27/MnistExamples.png>)

Methoden des Maschinellen Lernens

Genereller Aufbau

- Dimension des Input-Layer (erste Schicht) je nach Daten
- $\sigma(W \cdot x + b)$ für innere Schicht
 - σ ... Nichtlinearität
 - W ... Matrix mit zu optimierenden Parametern
 - x ... Datenpunkt bzw. Ausgabe vorheriger Schicht
 - b ... Vektor mit zu optimierenden Parametern
- Dimension der Ausgabeschicht je nach gewünschter Ausgabe
- Je nach Form der Ausgabe, Wahl der Nichtlinearität

Gradientabstieg

- Verfahren zur Optimierung
- Abstieg in negativer Richtung des Gradienten
- Verkürzung der Vektorenlänge mit einem Parameter (Lernrate)
- Lernrate als Hyperparameter
- verbesserte Version: Adam

Backpropagation

- Methode zur effizienten Berechnung des Gradienten
- "nur" Kettenregel in effizienter Art und Weise
- Speicherung der Ableitungen im Forward Pass mit Berechnung des Wertes
- backward pass zur Aufteilung des Fehlers auf die einzelnen Parameter
- Eigentlich in allen ML-Bibliotheken integriert

Beste Optimierungsergebnisse

- Dann legt mal los. Aber überfordert nicht euren PC. ;-)

Vielen Dank für Eure Aufmerksamkeit!
Und noch Viel Spass mit Maschinellen Lernen!