



13.09.2018

Sentiment Analysis Workshop

INNOQ

Agenda

- Willkommen bei OpenCH an der ETH
- Vorstellung
- Überblick Unterlagen + Infrastruktur
- Arbeitsumgebung + IDE
- Machine Learning Walkthrough

Vorstellungsrunde

- Ganz kurz
- Vorkenntnisse
- Erwartungen an den Workshop

Michael Krämer

- Michael Krämer, INNOQ Schweiz GmbH
- Fokus auf Java, Scala und maschinelles Lernen
- Architekt und Entwickler
- Sonst: Rad fahren, Reisen, draussen sein und stolzer Papa



Unterlagen

- <https://github.com/mkraemerx/sentiment-workshop>
- Präsentationsunterlagen
- Setup-Guide für Dev-Maschinen
 - <https://github.com/mkraemerx/sentiment-workshop/blob/master/infra/setup.sh>
- Daten
- Code

IDE

- Entwicklung über Jupyter Notebook
- <https://cx.int.postlab.de:8443>
- SSH-Zugriff falls nötig / gewünscht
- Text-Editor auch möglich

IDE (2)

- Esc / Enter um Modus zu wechseln
- Ctrl-Enter / Shift-Enter zum Ausführen
- a / b für neue Zelle
- Tab für Completion / Shift-Tab für Help

Toolchain

- Alles ist OpenSource!
- Python 3
- conda – Paketmanager für python
- <https://conda.io/docs/>
- Pytorch – Deep Learning Library von Facebook
- <https://pytorch.org/>

Deep Learning vs. Statistical Methods

- Deep Learning
 - deutlich weniger Vorwissen eincodiert
 - Braucht mehr Rechenleistung
 - Kann auf mehr Probleme angewandt werden
- Klassisches Machine Learning
 - performt teilweise besser
 - Spezifischere Lösungen
 - Feature-Engineering

DL – Frameworks (1)

- Framework implementiert die nötige Mathematik
- Stellt Datenstrukturen (Tensoren) zur Verfügung
- Bietet GPU-Unterstützung

DL – Frameworks (2)

- Warum pytorch und nicht Tensorflow?
 - + Ansprechenderes API
 - + dynamischer Graph
 - weniger Umgebungen (z.B. Mobile)
 - etwas weniger verbreitet

Trainingsdaten DE

- Leider wenig deutsche Trainingsdaten
- IGGSA - Interest Group on German Sentiment Analysis
- GermEval 2018 als shared Task hat Daten veröffentlicht
- Lustig: Grösster annotierter Trainingsdatensatz in deutscher Sprache ist aus Österreich

Weitere Infos

- <https://github.com/mkraemerx/sentiment-workshop/blob/master/infos.md>

Danke! Fragen?

INNOQ
www.innoq.com

Michael Krämer
michael.kraemer@innoq.com



innoQ Deutschland GmbH

Krischerstr. 100
40789 Monheim am Rhein
Germany
+49 2173 3366-0

Ohlauer Str. 43
10999 Berlin
Germany

Ludwigstr. 180E
63067
Offenbach
Germany

Kreuzstr. 16
80331
München
Germany

innoQ Schweiz GmbH

Gewerbestr. 11
CH-6330 Cham
Switzerland
+41 41 743 01 11

Albulastr. 55
8048 Zürich
Switzerland