

Exposé

Titel der Arbeit: Adaption multilingual vortrainierter Modelle zur automatischen Zusammenfassung von Texten auf die deutsche Sprache

Autor: Daniel Vogel

Forschungsfragen: Wie lassen sich Texte automatisiert zusammenfassen? Wie können bereits existierende Modelle auf eine andere Sprache adaptiert werden? Wie qualitativ und skalierbar ist die Lösung?

Inhaltsverzeichnis

Kapitel	Seiten
Einleitung	3
Deep Learning	-
- Neuronale Netze	2
- Architekturen	8
- Hyperparameter	2
- Transfer Learning	3
Natural Language Processing	-
- Vorverarbeitung	3
- Word Embeddings	3
- Deep Language Representations	3
Datengrundlage	3
Abstraktiver Ansatz	-
- Architektur	5
- Training	2
- Evaluation	2
Sprachtechnische Adaption	-
- Konzeption	8
- Training	2
- Evaluation	2
Zusammenfassung	2
Diskussion & Ausblick	2

Abstract

Zunächst ist eine kurze Einleitung erforderlich, welche den Kontext und die Notwendigkeit der Arbeit offenlegt. Als Ziel wird die abstraktive Zusammenfassung von Textdokumenten verstanden. Darüber hinaus sollen bereits existierende multilinguale Modelle mittels Transfer Learning auf die deutsche Sprache adaptiert werden. Die Zusammenfassung von Dialogen ist somit nicht Teil dieser Arbeit ist, wohl aber eine potenzielle Fortsetzung. Auch der Forschungsstand und die Referenzen gehören offengelegt.

Im Kapitel des Deep Learning werden zuerst neuronale Netze als unmittelbare Grundlage vorgestellt. Anschließend werden verschiedene Architekturen vorgestellt, welche im Verlauf der Arbeit noch an Relevanz gewinnen. Hierbei sind mithin Encoder-Decoder-Architekturen und Attention-Mechanismen zu erläutern sowie Transformer zu definieren. Weiterhin werden Hyperparameter und deren Bedeutung für das Training eines Modells behandelt. Das Transfer Learning bedarf ebenfalls einer Einführung, da es essenzieller Bestandteil der Arbeit sein wird.

Im Kapitel des Natural Language Processing werden weiterhin unmittelbare Grundlagen tangiert. Neben grundlegenden Prozessen der Arbeit mit und Vorverarbeitung von Texten werden hier Word Embeddings und Deep Language Representations im Sinne des Language Modelling vorgestellt. Diese werden in Kombination mit dem bereits bekannten Transfer Learning noch sehr bedeutsam.

Bevor die beschriebenen Architekturen und Methoden tatsächlich Anwendung finden können, ist die Beschreibung der Datengrundlage erforderlich. Hierbei handelt es sich um frei verfügbare allgemeinsprachliche Textdaten. Eine Ergänzung durch diverse fachspezifische kleinere Korpora ist denkbar.

Der abstraktive Ansatz wird in dem entsprechenden Kapitel zunächst vom extraktiven Ansatz abgegrenzt und beschreibt am Beispiel der ausgewählten Architektur, wie eine Zusammenfassung generiert werden kann. Die Architektur wird der Vollständigkeit halber beschrieben, ein entsprechendes Modell trainiert und das Ergebnis als Benchmark evaluiert.

Im letzten inhaltlichen Kapitel der Arbeit wird sodann die Adaption des Modells auf die deutsche Sprache thematisiert. Hierfür werden zunächst Anpassungen im Vergleich zur ursprünglichen Architektur konzipiert. Ein erneutes Training und eine Evaluation schließen sich an. Die Ergebnisse lassen sich nun vergleichen und gewisse Tendenzen erkennen.

Die Zusammenfassung, die Diskussion und der Ausblick runden die Arbeit ab, bevor die Literatur und die Anhänge den formalen Abschluss vornehmen.

Literatur (siehe Literaturverzeichnis, auf Anfrage per E-Mail zustellbar)

Stand: 13. Januar 2021, 55 Seiten.