

## **SOSE 2017,CIS,LMU,München**

Protokoll zur Sitzung am 15.05.2017

Präsentation von Joseph Birkner

Das Thema der Präsentation lautet als “Ranking with neural networks derived document vectors”, der Betreuer der Bachelorarbeit ist Wenpeng Yin. Die Arbeit basiert auf einem Projekt IROM (Intelligent Recommendation of Massive Open online course). Das Grundprinzip dafür ist Ubiquitous Vertical Search, welcher gründet sich auf der Suche mit der Hilfe von “Vertikal search engine”. Dazu hat Joseph eine Abbildung dargestellt, wo “vertical search” angezeigt wurde – es besteht aus zwei Unterbereichen: Repräsentationsoptimierung – DSSM (Deep semantic structured matching) und Passende optimierung – DRMR (Deep relevance matching model). Das alles stellt neural Information Retrieval dar.

Das Ziel dazu ist Optimierung der Dokumentationsdarstellung nämlich Entwicklung des System, die eine gute Dokumentendarstellung für eine effiziente Ranking erzeugen kann. Dazu entstehendes Problem liegt daran dass die Reihenfolge der Wörter ignoriert wurde und die Abhängigkeit von Wörtern wurde oft falsch interpretiert.

Als Objektiv hat er das “semantic space” für Dokumenten definiert nämlich Latent Semantic Spaces (k-dimensionaler space, wo jede Dimension ein orthogonales semantisches Konzept kodiert), Word2Vec (semantic space für Wörter) und Doc2Vec (semantic space für Dokumente).

Der Tasks Prototyp soll generierte Dokumentenvektoren (embeddings) implementieren und die 30-dimensionale Dokumentenvektoren erstellen die nur mit circa 1200 Beschreibungen getestet werden sollen.