

Research Colloquium SS2018



In the context of the research colloquium Business Informatics and Data Science Prof. Dr. Andreas Hotho, Professor of Computer Science - in particular Data Mining and Information Retrieval from the University of Würzburg, will give a lecture on "Learning Semantic Relatedness From Human Feedback Using Metric Learning".

Date and location: April 5, 2018 12:15 UC 11.008

Abstract:

Assessing the degree of semantic relatedness between words is an important task with a variety of semantic applications, such as ontology learning for the Semantic Web, semantic search, recommendation or query expansion. To accomplish this in an automated fashion, many relatedness measures have been proposed. However, most of these metrics only encode information contained in the underlying corpus or in the navigation and thus do not directly model human intuition. In this talk, we show the utilization of metric learning to improve existing semantic relatedness measures by learning from additional information, such as explicit human feedback. Our approach is based on the knowledge that emergent semantic information is collected in Social Media systems in the user's content or navigational traces. We argue to use word embeddings instead of traditional high-dimensional vector representations in order to leverage their semantic density and to reduce computational cost as a first step to improve the extraction of the hidden semantic. We present results on several domains including tagging data as well as publicly available embeddings based on Wikipedia texts and navigation. Second, human feedback about semantic relatedness for learning and evaluation is extracted from publicly available datasets such as MEN or WS-353. We will show that our method can significantly improve semantic relatedness measures by learning from the additional explicit human feedback. For tagging data, we are the first to generate and study embeddings. Our results are of special interest for researchers and practitioners of Semantic Web and show the power of Machine Learning methods for extracting semantics.