# Machine Learning

##### Session 4.1 (SEMANTiCS)

#### Time: Thursday, September 19, 2024 - 13:40 to 15:00

#### Chair:

## **Talks**

### Entity Linking with Out-of-Knowledge-Graph Entity Detection and Clustering using only Knowledge Graphs

Based on our novel dataset, we develop an approach using pre-trained language models and knowledge graph embeddings without the need for a parallel full-text corpus.

Moreover, by assessing the influence of knowledge graph embeddings on the given task, we show that implementing a sequential entity linking approach, which considers the whole sentence, can outperform clustering techniques that handle each mention separately in specific instances.

| Cedric Möller | Ricardo Usbeck |
| --- | --- |

### 

### TWIG-I: Embedding-Free Link Prediction and Cross-KG Transfer Learning using a Small Neural Architecture

Knowledge Graphs (KGs) are relational knowledge bases that represent facts as a set of labelled nodes and the labelled relations between them. Their machine learning counterpart, Knowledge Graph Embeddings (KGEs), learn to predict new facts based on the data contained in a KG -- the so-called link prediction task. To date, almost all forms of link prediction for KGs rely on some form of embedding model, and KGEs hold state-of-the-art status for link prediction. In this paper, we present TWIG-I (Topologically-Weighted Intelligence Generation for Inference), a novel link prediction system that can represent the features of a KG in latent space without using node or edge embeddings. We show that TWIG-I can increase performance on the link prediction relative to KGE models, including a 35 base-point increase in MRR performance on FB15k-237 over the strongest baseline; this represents a 100% relative increase in performance. Unlike KGEs, TWIG-I can be natively used for transfer learning across KGs, even across KGs that come from different knowledge domains. We show that using transfer learning with TWIG-I can lead to notable increases in performance both over KGE baselines and over TWIG-I models trained without finetuning. With finetuning, TWIG-I is able to achieve a 44 base-point increase in MRR over the standard benchmark KG FB15k-237 relative to the strongest baseline, which corresponds to a 126% relative increase in predictive performance.

| Jeffrey Sardina | Alok Debnath |
| --- | --- |
| John D. Kelleher | Declan O'Sullivan |

### Stitching Gaps: Fusing Situated Perceptual Knowledge with Vision Transformers for High-Level Image Classification

We construct the ARTstract Knowledge Graph (AKG), capturing perceptual semantics from over 14,000 cultural images labeled with ACs. We extract perceptual semantic units using off-the-shelf models and integrate them into the AKG, enriching it with high-level linguistic frames. For AC-based image classification, we adopt a hybrid approach, integrating knowledge graphs and visual transformers. Specifically, we compute knowledge graph embeddings (KGE) on AKG and fuse them with visual transformer embeddings. For interpretability, we conduct post-hoc qualitative analyses by probing model similarities with training instances.

| Delfina Sol Martinez Pandiani | Nicolas Lazzari | Valentina Presutti |
| --- | --- | --- |

### ReWise: A Relation-Wise Sampling Framework for Relational Graph Convolutional Networks

Relational graph convolutional networks (RGCNs) have been successful in learning from knowledge graphs. However, training on large-scale knowledge graphs becomes challenging due to the exponential growth of the neighborhood size across the network layers. Moreover, knowledge graphs have multiple relations, and often, the literals can have multimodal content; these properties make it extra challenging to scale up the training of RGCNs to large-scale graphs. Graph sampling techniques have been shown to be effective in scaling learning to large graphs by reducing the number of processed nodes and lowering memory usage. However, only a few studies have focused on sampling for knowledge graphs. In this work, we introduce ReWise, a relation-wise sampling framework that includes a family of sampling methods designed for knowledge graphs. Our experiments demonstrate that sampling reduces memory usage up to 50% lower than the case without sampling while maintaining the same classification accuracy and, in some cases, outperforming it. Additionally, we show that our sampling strategy is compatible with the multimodal RGCN, showing the same behavior as RGCNs.

| Taraneh Younesian | Peter Bloem | Stefan Schlobach |
| --- | --- | --- |