Project Details for ECTS Credits

Eva Maria Vecchi
Embeddings in Natural Language Processing

September 9-13, 2019

For ECTS credit, you are required to submit a 2-part project in which you demonstrate a theoretical and practical understanding of a chosen topic using or related to embeddings in NLP. Each student should select a topic, submit a literature review of scientific literature on the selected topic, and implement a practical study on the topic.

Credits   A student can receive 3 ECTS credits for the course, corresponding to circa 90 hours of work.

Deadlines

1. I must receive, via email (evamariavecchi@gmail.com), your (i) topic choice; (ii) selection of literature and research material; and (iii) summary of the practical study by **October 4, 2018** at the latest for approval.
   Note! If I do not receive an email by this date, your project cannot be considered for credit.

2. All projects must be submitted by **December 18, 2019**, without exception.

Choosing your topic   Each student will select their own topic, either from the list provided below or something else approved by me.

Literature and research material selection   Once you have selected your topic, you are expected to select 1-3 pieces of literature (articles, chapters, etc.) which you will review and which will drive your practical study. The selection of material can be done alone (with approval by me once you've selected the material) or it can be discussed with me.

Literature review   The review should be a general and critical overview of the selected literature/research of your topic. You are expected to cover the theoretical aspects of your topic (background, motivation, theoretical implications, etc.) as well as outline experimental design and results presented in the literature. There is no set length for the review, but I would expect 2-4 pages.

Practical study   Each student will implement a practical study relevant to the research discussed in the selected literature and the selected topic. The type of study can vary, including (but not limited to) reimplementation of a system, detailed analysis of results, application of an approach to a new dataset/problem/language, comparative analysis of parameters or embedding types for a task, etc. As with the choice of material, the practical study must be approved by me. You are to submit a summary of the practical study you intend to implement along with your topic choice by October 4th.

Proposed Topics

1. **Lexical Semantics**

   (a) Computational approaches to word meaning
   (b) Lexical relations in computational tasks

2. **Formal Semantics**

   (a) Model-Theoretic semantic spaces
   (b) Entailment in computational semantics

3. **Cognitive Semantics**

   (a) Feature norms and Conceptual semantic spaces
   (b) Metaphor in NLP
   (c) Modeling Concrete vs. Abstract Concepts

4. **Distributional Semantics**

   (a) Applications and evaluation of DS Models

   (b) Approximation of human judgments

   (c) Multi-modal semantic models (also as topic in Machine Learning)

5. **Compositional Distributional Semantics**

   (a) Applications of CDS Models

   (b) Composition functions: performance and implications for syntax, word order, semantic creativity, etc

   (c) Morphology and CDS

   (d) Multi-word expressions and non-compositionality

6. **Machine Learning for Modeling Meaning**

   (a) Improving word representations with recursive neural networks

   (b) Morphological smoothing and extrapolation of word embeddings

   (c) Approaches to de-bias word embeddings

7. **Other**