Advanced Natural Language Processing

Course Workload		Assessment form (examination/ graded test/ ungraded test)
ECTS	Hours	
5	180	Exam

The course "Advanced Natural Language Processing" covers a wide range of methods for solving practical problems, ranging from plagiarism detection to text generation. The building blocks of neural networks, ways of modifying them and quality measurement metrics for solving target problems are analyzed.

Course structure:

1. Basic data processing

1.1. The main pipeline for word processing. Separate operations and methods of cleaning the text, conditions of their use and purpose.

2. Approaches to words and text embeddings

2.1. Introduction to Natural Language Processing; an overview of approaches to the presentation of words and texts; basic models for representing words and texts.

3. Topic Modeling

3.1. Introduction to the problem of searching for topics. The main methods of thematic modeling: NMF, LDA, ARTM. Building thematic models using BigARTM.

4. Full Text Search with ElasticSearch

4.1. Text proximity metrics. Implementation of text search using ElasticSearch database. ElasticSearch Appliance Features: Inverse Index, Metrics, and Scoring Functions.

5. Basic neural network architectures for NLP: RNN and CNN for text analysis. ELMo and ULmFit

5.1. Acquaintance with the principles of the structure of the basic neural network architectures and approaches for text processing: vanilla RNN, LSTM, GRU, CNN, ULmFit, Elmo, Siamese architecture etc. Understanding the concepts of pretraining and fine-tuning. Building applied solutions based on them using the PyTorch framework and training them on Kubeflow and Kubernetes.

6. Attention and Transformer architecture

6.1. Learning attention mechanism and it's advantages and disadvantages. Fine-tuning the model for specific tasks.

7. Modern Transformer-based architectures

7.1. Looking at the Transformer-based architectures, understanding the engineering differences between vanilla Transformer (and BERT) with the new architectures. Defining conditions of usage.

8. Question-Answering Systems

8.1. Introduction to question-Answering tasks, BiDAF architecture, ways to prepare the datasets.

9. Knowledge distillation

9.1. Reduction of model size and weights to use the distilled model in production.