Skoltech
Skolkovo Institute of Science and Technology

Course Syllabus

Course Title (in English)	Neural Natural Language Processing
Course Title (in Russian)	Нейросетевые методы обработки естественного языка
Lead Instructor(s)	Panchenko, Alexander

Contact Person	Alexander Panchenko
Contact Person's E-mail	a.panchenko@skoltech.ru

1. Annotation

Course Description

The course is about neural models for natural language processing. The new generation of neural network-based methods based on deep learning has dramatically improved the performance of a wide range of natural language processing tasks, ranging from text classification to question answering. The course covers the basics and the details of successful models and methods for natural language processing based on neural networks, starting from the simple word embedding models, such as word2vec, all the way to more sophisticated language models, such ELMo and BERT. Besides, the course contains a small introduction to basic NLP methods. The course involves a substantial practical component with a number of practical assignments.

Course Prerequisites / Recommendations	Linear algebra, Machine learning, Python programming

2. Structure and Content

Course Academic Level	Master-level course suitable for PhD students
Number of ECTS credits	3

Topic	Summary of Topic	Lectures (# of hours)	Seminars (# of hours)	Labs (# of hours)
An introduction to NLP and text categorisation	Text preprocessing and input representations in NLP. Text classification and the task of sentiment analysis. Bag of word model, one-hot vectors and embeddings, TF-IDF weighting. Bernoulli and multinomial Naive Bayes baseline model.	3	3	0
A reminder on neural models	A short reminder on neural networks. Logistic regression and feedforward neural network. Training of neural networks with stochastic gradient descent and its variants. Logistic regression and feedforward networks for text classification.	3	3	0
Word and document embeddings	Word embeddings: word2vec, GloVe and related models. Document embeddings: doc2vec, DV-ngram and related models.	3	3	0
Recurrent neural networks for NLP	Recurrent neural networks. LSTM and GRU models. Language modeling task.	3	3	0
Pretraining	Pretraining for text classification: word2vec and language models. Large-scale LM pretraning. Knowledge transfer from language models. ELMo and ULMFiT. Adversarial and virtual adversarial training for text classification.	3	3	0
Sequence to sequence models	Sequence2sequence models. Neural machine translation. Encoder-decoder architecture.	3	3	0
Attention	Attention in neural machine translation. Self-attention mechanism. Transformer model.	3	3	0
BERT model	Knowledge transfer using BERT model. Fine-tuning of BERT model for various NLP tasks.	3	3	0

3. Assignments

Assignment Type	Assignment Summary
Final Project	1) Find an interesting task and propose an neural NLP model to solve it. 2) Propose a new NLP task or a variant of some existing one and come up with a baseline for its solution. 3) Get a recently published NLP paper and replicate its results. Discuss the outcomes.
Homework	Sentiment analysis using Naive Bayes classifier.
Homework	Sentiment analysis using logistic regression.
Homework	Sentiment analysis using feed-forward neural network.
Homework	Sentiment analysis using LSTMs.
Homework	Sentiment analysis using document embeddings.
Homework	Sentiment analysis using ELMo or BERT.

4. Grading

Type of Assessment	Graded

Grade Structure

Activity Type	Activity weight, %
Homework Assignments	40
Final Project	40
Final Exam	20

Grading Scale

A:	86
B:	76
C:	66
D:	56
E:	46
F:	0
Attendance Requirements	Mandatory with Exceptions

5. Basic Information

Maximum Number of Students

	Maximum Number of Students
Overall:	35
Per Group (for seminars and labs):	5

Course Stream	Science, Technology and Engineering (STE)
Course Delivery Frequency	Every year

Students of Which Programs do You Recommend to Consider this Course as an Elective?

Masters Programs	PhD Programs
Data Science Information Science and Technology	Computational and Data Science and Engineering

Course Tags	Math Programming Engineering

6. Textbooks and Internet Resources

Required Textbooks	ISBN-13 (or ISBN-10)
Yoav Goldberg. Neural Network Methods in Natural Language Processing. Synthesis Lectures on Human Language Technologies. https://doi.org/10.2200/S00762ED1V01Y201703HLT037	9781627052986
Kyunghyun Cho. Natural Language Understanding with Distributed Representation. Lecture note for the course DS-GA 3001 on Natural Language Understanding with Distributed Representation at the Center for Data Science, New York University. https://arxiv.org/abs/1511.07916	N/A

Recommended Textbooks	ISBN-13 (or ISBN-10)
Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning. MIT Press. http://www.deeplearningbook.org	0262035618
Aston Zhang, Zack C. Lipton, Mu Li, Alex J. Smola. Dive into Deep Learning. Sep 25, 2019. https://d2l.ai	N/A

Papers	DOI or URL
Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI Blog, 1(8).	https://www.techbooky.com/wp-content/uploads/2019/02/Better-Language-Models-and-Their-Implications.pdf
Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Melbourne, Australia	https://www.aclweb.org/anthology/P18-1031/
Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal	https://www.aclweb.org/anthology/D15-1166/
Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis, Minnesota	https://www.aclweb.org/anthology/N19-1423/
Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). New Orleans, Louisiana	https://www.aclweb.org/anthology/N18-1202/
Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. Journal of machine learning research, 12(Aug), 2493-2537.	http://www.jmlr.org/papers/volume12/collobert11a/collobert11a.pdf
Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. ICLR 2015	https://arxiv.org/pdf/1409.0473.pdf

Web-resources (links)	Description
http://web.stanford.edu/class/cs224n/	Stanford class on deep learning for natural language processing
https://www.coursera.org/learn/language-processing	Coursera class from HSE on natural language processing
https://github.com/yandexdataschool/nlp_course	A course on natural language processing in Yandex School of Data Analysis

7. Facilities

Equipment
Laptop
Access to a server with GPU in Skoltech

Software
Linux Ubuntu 18.04
Python 3.6
PyTorch
iPython
Google CoLab
Tensorflow

Labs for Education	Computer Lab

8. Learning Outcomes

Knowledge
Obtain a big picture of various neural models used for natural language processing.
Know the architectural peculiarities of modern neural models applied for natural language processing problems.
Know the theoretical basis and conceptual tools needed for the development of new neural models for natural language processing;
Be aware of the recent progress in natural language processing and ways to efficiently familiarise yourself with the further developments.

Skill
Prototype, train, and apply neural architectures to natural language processing problems.
Be able to solve a natural language processing problem using modern methods: identify the corresponding type of model, select an appropriate method, evaluate the model, interpret the obtained results;
Be able to implement NLP methods from scratch and apply existing software libraries for NLP;
Ability to read and discuss research papers in NLP and follow developments in the state-of-the-art in this field.

Experience
Obtain a sufficient experience during practical exercises and project activities to become a qualified user of neural natural language processing methods.

9. Assessment Criteria

Input or Upload Example(s) of Assigment 1:

Select Assignment 1 Type	Computer Labs
Input Example(s) of Assignment 1 (preferable)	N/A
Assessment Criteria for Assignment 1	A filled-in Jupyter notebook with a trained network yielding sufficiently good result

Input or Upload Example(s) of Assigment 2:

Input or Upload Example(s) of Assigment 3:

Input or Upload Example(s) of Assigment 4:

Input or Upload Example(s) of Assigment 5:

10. Additional Notes