

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Суворов Антон Дмитриевич
Должность: Ректор
Дата подписания: 15.07.2025 17:52:54
Уникальный программный ключ:
a39bdb15d680d5b0adb1ced0af5e1efb14747dc0

СКОЛКОВСКИЙ ИНСТИТУТ НАУКИ И ТЕХНОЛОГИЙ (Сколтех)

Рабочая программа дисциплины	Трансформерные архитектуры и большие языковые модели
------------------------------	--

Преподаватель	Панченко Александр Иванович, доцент, PhD
---------------	--

Аннотация

Описание курса

Курс посвящен современным моделям для обработки естественного языка. Новое поколение методов на основе нейронных сетей, основанных на глубоком обучении, значительно улучшило производительность широкого спектра задач обработки естественного языка, начиная от классификации текста и заканчивая ответами на вопросы. Курс охватывает основы и детали успешных моделей и методов обработки естественного языка на основе нейронных сетей, начиная с простых моделей векторных представлений слов, таких как word2vec, и заканчивая более сложными языковыми моделями, такими как GPT.

Особое внимание уделяется моделям основанным на архитектуре Transformer, таких как GPT, BERT, RoBERTa, T5, BART, и т.п.

1. Основная информация

Академический уровень курса	Магистратура Аспирантура
Количество кредитов	3

Отображение оценок в процентах

A:	86
B:	76
C:	66
D:	56
E:	46
F:	0

2. Содержание курса

Тема	Краткое содержание	Лекции (час)	Семинары (час)	Лабораторные	Самостоятельная
------	--------------------	--------------	----------------	--------------	-----------------

				занятия (час)	работа (час)
Вложения в документы Word	Программы для встраивания в Word: word2vec, GloVe и связанные с ними модели. Программы для встраивания документов: doc2vec, DV-ngram и связанные с ними модели	3	3		
Рекуррентные нейронные сети для NLP	Рекуррентные нейронные сети. Модели LSTM и GRU. Задача языкового моделирования	3	3		
Предварительная тренировка	Предварительная подготовка к классификации текстов: word2vec и языковые модели. Предварительная подготовка крупномасштабных фильмов. Передача знаний из языковых моделей. ELMo и ULMFiT. Состязательные и виртуальные состязательные тренировки по классификации текстов	3	3		
Модели от последовательности к последовательности	Модели Sequence2sequence. Нейронный машинный перевод. Архитектура кодера-декодера	3	3		
Внимание	Внимание при нейронном машинном переводе. Механизм саморегуляции внимания. Трансформаторная модель	3	3		
Модель BERT	Передача знаний с использованием модели BERT. Тонкая настройка модели BERT для различных задач NLP	3	3		
Сетевые кодеры	NLP и графические данные. Простые кодеры: node2vec и DeepWalk. Сети свертки графов (GCN) для NLP. Преобразователи графических данных	3	3		

3. Результаты обучения

Знать
общую картину различных нейронных моделей, используемых для обработки естественного языка
архитектурные особенности современных нейронных моделей, применяемых для решения задач обработки естественного языка
теоретические основы и концептуальные инструменты, необходимые для разработки новых нейронных моделей обработки естественного языка
быть в курсе последних достижений в области обработки естественного языка и способов эффективного ознакомления с дальнейшими разработками

Уметь
создавать прототипы, обучайте и применяйте нейронные архитектуры для решения задач обработки естественного языка
решать задачу обработки естественного языка с использованием современных методов: определять соответствующий тип модели, выбирать

подходящий метод, оценивать модель, интерпретировать полученные результаты
внедрять методы NLP с нуля и применять существующие программные библиотеки для NLP
читать и обсуждать исследовательские работы по NLP и следить за последними достижениями в этой области

Владеть
достаточным опытом во время практических занятий и проектной деятельности, чтобы стать квалифицированным пользователем нейронных методов обработки естественного языка

4. Задания и выставление оценок

Тип назначения	Краткое содержание задания	% от итоговой оценки за курс
Домашние задания	Анализ настроений. Цель этой задачи - внедрить систему анализа настроений, которая будет выявлять положительные и отрицательные комментарии. Предоставляется общедоступная таблица лидеров с использованием системы CodaLab. Оценивается отчет в виде Jupyter notebook.	30
Домашние задания	Детоксикация текста. Целью данной задачи является внедрение системы передачи стиля текста. Предоставляется общедоступная таблица лидеров с использованием системы CodaLab. Оценивается отчет в виде Jupyter notebook	30
Итоговый проект	Исследовательский проект. Сначала группа из 5 студентов выбирает тему – набор заранее определенных проектов, соответствующих актуальной программе исследований лаборатории NLP Сколтеха (возможны и другие темы). Команда работает с наставником в группе и представляет свое решение аналогично презентации на конференции по NLP	30
Тест	Вопросы по темам лекций	10

5. Учебники и интернет-ресурсы

Необходимые учебники	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	
Дополнительная литература	
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	
Статьи	
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 finetuning 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

Веб-ресурсы (ссылки)	Описание
-----------------------------	-----------------

http://web.stanford.edu/class/cs224n/	Занятия в Стэнфорде по глубокому обучению обработке естественного языка
https://www.coursera.org/learn/language-processing	Курс Coursera от HSE по обработке естественного языка
https://github.com/yandexdataschool/nlp_course	Курс по обработке естественного языка в Школе анализа данных Яндекса

6. Оборудование

Программное обеспечение
Linux Ubuntu Python 3 PyTorch Google CoLab

Оборудование
Ноутбук со стабильным подключением к Интернету Доступ к серверу с графическим процессором в Сколтехе