***Программа курса по NLP, 2020.***

Каждый блок состоит из 1 лекции и 1 семинара.

1. NLP: введение, базовые элементы пайплайна: сегментация предложений, токенизация, морфологический анализ.
2. Современные модели POS-таггинга и морфологического анализа

3-4. Векторные представления слов и языковые модели.  
*После занятия 4 будет выдано дз 1.*

5. Тектовая классификация

6. Извлечение именованных сущностей (NER)

7. Transition-based синтаксический анализ

8. Relation extraction (извлечение фактов и отношений

*После занятия 8 будет выдано дз 2 (проект).*

9. Goal-oriented диалоговые системы.

10. Исправление опечаток.

11-12. Машинный перевод, seq2seq и трансформер.

*После занятия 12 будет выдано дз 3.*

13. General Conversation модели (болталки).

14. Тематическое моделирование.

15. Задача суммаризации текстов.

16. Вопросно-ответные системы  
*После занятия 16 будет выдано дз 4.*