

**Билеты:**

1. Рекуррентные нейронные сети. Алгоритм обратного распространения ошибки сквозь время (backpropagation through time, BPTT). Вычисление градиента функции ошибки по параметрам рекуррентной нейросети в явной форме. Проблемы обучения рекуррентных нейронных сетей (затухание и взрыв градиента). [[ссылка](#), [ссылка](#)]
2. Борьба с затуханием и взрывом градиента. Архитектуры LSTM и GRU. Gradient clipping [[ссылка](#), [ссылка](#), [ссылка](#)]
3. Генеративные модели. Задача моделирования последовательностей (языковая модель): представление совместной вероятности через chain rule, методы оценки условных вероятностей через n-gram. Использование нейронных сетей для обучения языковой модели. Truncated backpropagation through time. [[ссылка](#), [ссылка](#)]
4. Задача обучения векторов представления. One-hot encoding, counts, embedding matrix. Word2Vec (CBOW и Skip-gramm). [[ссылка](#)]
5. Задача машинного перевода. Оценка качества перевода (метрики BLEU, ROUGE). Модель Sequence-to-sequence (Seq2Seq). [[ссылка](#), [ссылка](#)]
6. Механизмы внимания в нейросетях (attention). Soft и hard attention. Метод REINFORCE (log-derivative trick) для обучения hard attention. Применение attention для задачи классификации изображений. Применение attention для задачи машинного перевода [[ссылка](#), [ссылка](#)]

**Теоретический минимум:**

1. Архитектура LSTM
2. Архитектура GRU
3. Языковая модель
4. Модель Seq2Seq
5. Метод REINFORCE
6. Задача машинного перевода

**Домашняя работа:**

1. Реализация языковой модели для генерации цитат умных людей
2. Написать рецензию на статью [Attention Is All You Need](#). Рецензия должна содержать краткое описание статьи, ее преимущества и недостатки. Текст желательно писать на английском, но можно и на русском. Если пишете на английском, советую воспользоваться <https://app.grammarly.com/> для проверки орфографии. О том как писать рецензии можно послушать тут: <https://www.coursera.org/learn/sciwrite/lecture/mu4ZE/6-7-doing-a-peer-review>.