



Во всех вопросах может быть несколько правильных ответов

В тестовых заданиях первая галочка — правильный ответ, вторая галочка — выбранный ответ. Цвет обозначает, правильно ли в данном пункте поставлена галочка. Если все пункты верные (галочки совпадают / все пункты зеленые), то за задание ставится полный балл, в противном случае ставится 0 баллов.

1. Рассмотрим задачу, где по отзыву человека на фильм нужно определить, является ли отзыв положительным, отрицательным или нейтральным. Какую схему рекуррентной нейросети нужно использовать?

☐ ☐ one-to-many

☒ ☒ many-to-one

☐ ☐ many-to-many

☐ ☐ one-to-one

Балл: 0.75

Комментарий к правильному ответу:

2. Зачем нужна обрезка градиентов (Gradient clipping) при настройке рекуррентных нейронных сетей?

☐ ☐ для борьбы с малыми по норме градиентами

☒ ☒ для борьбы с большими по норме градиентами

Балл: 0.75

Комментарий к правильному ответу:

3. Передача информации от кодировщика к декодировщику в many-to-many рекуррентной сети осуществляется через:

☐ ☐ вектор, размер которого пропорционален квадрату длины входной последовательности

☒ ☒ вектор фиксированного размера

☐ ☐ вектор, размер которого пропорционален длине входной последовательности

Балл: 0.75

Комментарий к правильному ответу:

4. Выходы кодировщика в сети трансформера зависят:



☒ ☒ от всех входов кодировщика

☐ ☐ только от входов на соответствующих позициях

☐ ☐ только от входов кодировщика в небольшой окрестности позиции каждого выхода

Балл: 0.75

Комментарий к правильному ответу:

5. Выберите верные утверждения для эмбедингов слов, генерируемых нейросетевыми языковыми моделями (такими как word2vec, glove, fastText и т.д.):

☒ ☒ эмбединги имеют меньшую размерность, чем one-hot кодирование слов, что приводит к упрощению использующей их нейросети

☐ ☐ по норме эмбединга можно судить о том, насколько слово редкое

☐ ☐ эмбединги имеют интерпретируемые компоненты

☒ ☒ эмбединги можно использовать в другой задаче, не связанной с той, на которой эмбединги обучались

Балл: 0.75

Комментарий к правильному ответу:

6. Проблема взрывающегося градиента (exploding gradient) в рекуррентных сетях приводит к тому, что:

☐ ☐ сеть помнит ограниченное число наблюдений в истории

☒ ☒ настройка сети происходит нестабильно

☐ ☐ сеть не помнит предыдущего наблюдения

☐ ☐ сеть помнит все наблюдения, но информация о них забывается экспоненциально быстро

Балл: 0.75

Комментарий к правильному ответу:

7. Недостаток подхода teacher forcing при настройке генерирующих рекуррентных сетей, по сравнению с настройкой её, подавая на вход ранее сгенерированные выходы (free run), заключается в том что в общем случае при teacher forcing:



- ☒ ☒ в отличие от free run, сеть не учится восстанавливаться после совершения ошибок при генерации последовательности
- ☐ ☐ сеть настраивается медленнее, чем при free run
- ☐ ☐ сеть хуже помнит историю входных данных, чем при free run

Балл: 0.75

Комментарий к правильному ответу:

8. Увеличение смещения (bias) в forget gate сети LSTM позволяет сети:

- ☒ ☐ дольше помнить входную информацию, полученную много шагов назад
- ☐ ☒ быстрее забывать входную информацию, полученную много шагов назад
- ☐ ☐ не влияет на запоминание прошлой информации

Балл: 0

Комментарий к правильному ответу: