

Ответы на вопросы к лабораторной работе №8.

Головина Е.С.

1. В чем заключается смысл ансамблирования комитетом большинства?

Ансамблирование комитетом большинства – это простое голосование, где некоторое количество моделей предсказывают определенный результат и итоговым выбирается тот, что встречается чаще других.

Смысл данного ансамблирования в повышении точности конечного результата. При увеличении количества моделей, участвующих в ансамбле – их совместная ошибка уменьшается. Но на качество результата также влияет независимость моделей, т.е. чем более независимыми будут модели, входящие в ансамбль, тем более качественный результат будет получен.

2. Какая метрика качества используется в лаб. работе и почему?

Обычные метрики, используемые в прошлых лабораторных работах, не имело смысл использовать в данной работе, потому что нет эталона, и с исходным текстом не сравнить, ведь модель генерирует новый текст.

Поэтому в качестве метрики качества я использовала свое восприятие получившегося текста. В идеале нужно смотреть, насколько текст сгенерировался похожим на оригинал по стилистике. Но так как в данной работе хорошего текста не вышло я смотрела в первую очередь на то:

- сколько разных слов сгенерировано
- сколько слов реально существует в английском языке
- количество повторений одного и того же
- количество сочетаний слов, которые имеют смысл
- получилось ли сгенерировать предложения

3. В чем плюсы и минусы того, что модель предсказывала слова, а не символы?

Если бы мы обучали сеть предсказывать слова, то, конечно, сгенерированные слова получались бы лучше, чем в данной лабораторной. В этом плюс.

Но минус в том, что, так как связь между символами в разных словах не учитывалась, то и связного текста не получилось бы, а в данной лабораторной работе цель была в создании текста, а не отдельных слов.