**Вопросы и ответы:**

1) Эссе, которое содержит фактическую ошибку имеет ответ True или False?

2) В датасете есть дисбаланс по кол. эссе. Какой он?

В каком колабе он был выявлен? (Я увидел дисбаланс только в кол. слов, что содержатся в выборках).

3) Где указано кол. эссе по каждому предмету

Количество эссе по каждому предмету - указано графой text\_count в файле.json.

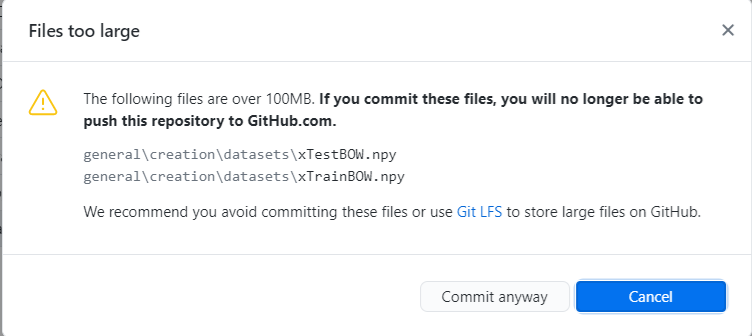
4) Как организовать хранение наших выводов и предположений по обучению нейросети? Как это делают датасентисты?

5) Удачный ли мы выбрали размер словаря (кол. слов, что учитываются при обучении), при формировании датасета?

Указать данные

6) Правильно ли я сделал экспорт датасета Мелании (Словарь частотности)?

7) Как сохранить сформированный датасет, если он весит очень много?



Дела:

Сохранять все датасеты в отдельную папку.