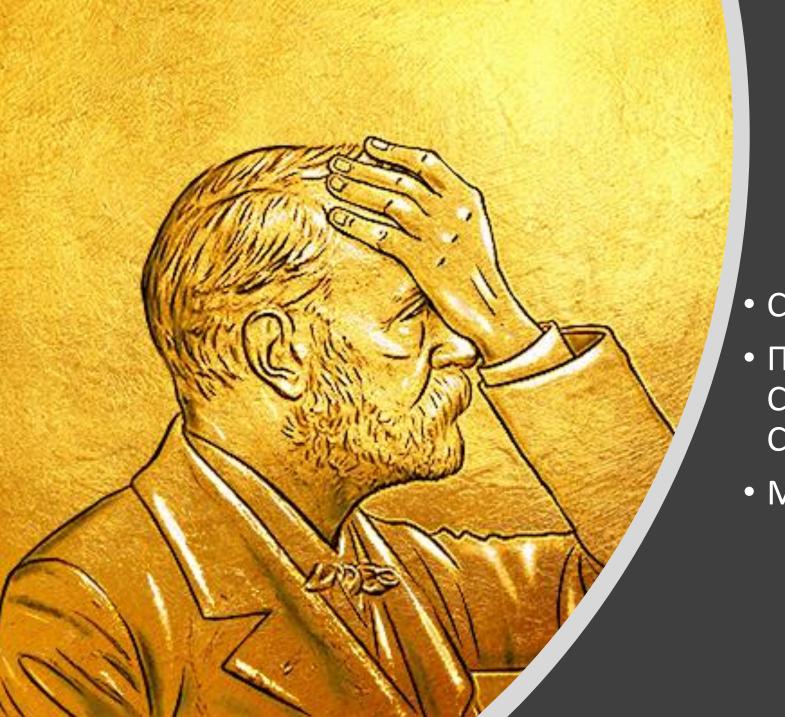


Генерация названий статей для Шнобелевской премии

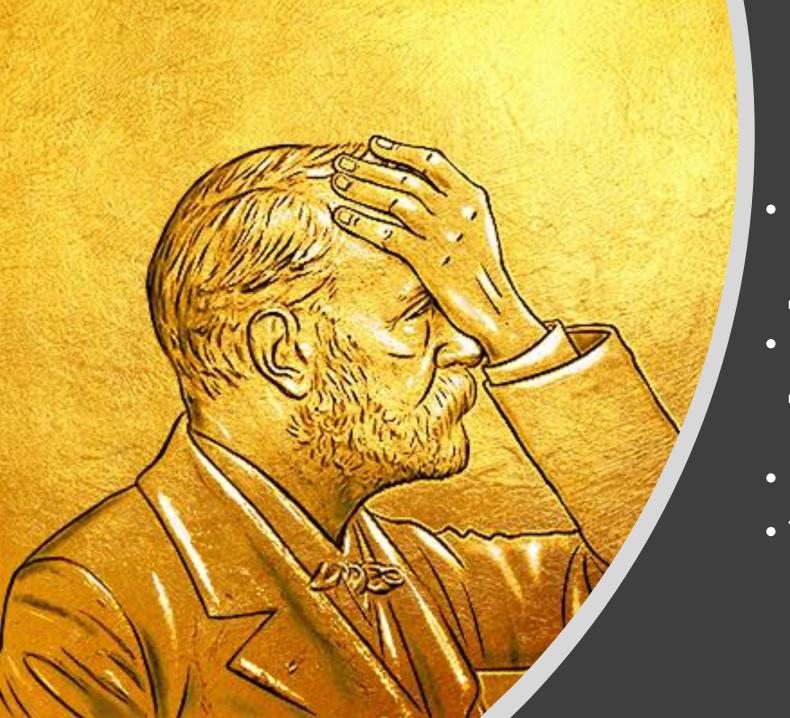
Misfit`s Prize Team ODS Хакатон 3-10 октября 2020



Команда

Misfit's Prize Team

- Скрипт Александр Мирленко
- Поиск данных Владимир Семериков, Ирина Васильева, Сергей Мельник
- Ментор Анастасия Малышева



Данные

- Названия статей на английском языке STEM дисциплин (Archive.org)
- Поиск датасетов из других дисциплин (здесь нас ждал провал, наука не для всех)
- Поиск информации по теме
- Туториалы от индусов



Концепция

Бредогенератор — выбирает следующее слово, основываясь на вероятности появления после предыдущего слова.

Пространство для творчества – какое слово взять следующим?

- самое вероятное
- самое невероятное (несвязный набор слов)
- случайное из наиболее вероятных



Первые результаты

- Обилие неинформативных токенов (нужна фильтрация)
- Длинный текст (не подходит для заголовка)
- Первое слово задается автоматически процесс неуправляемый



Итоговый результат

Space –

Space and criteria in imported malaria, mid-range and mid-quartile range, adaptive hedge, lattice interferometer

Apple –

Apple dapple substantial energy output of population gradients of nitrogen dioxide measurement in same-sign



В продакшен!



Пространство для творчества

- улучшенная рандомизация (подбирается индивидуально под каждую задачу);
- разнообразие данных (модель обучена и работает со STEM-дисциплинами);
- улучшенное быстродействие (за счет применения более изысканных алгоритмов).

