МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Систем обработки информации и управления»

ОТЧЕТ

**Рубежный контроль № 2**

по дисциплине «Методы машинного обучения»

## Тема: «Методы обработки текстов..»

ИСПОЛНИТЕЛЬ: Ли Лююй

ФИО

группа ИУ5И-21М

подпись"

ПРЕПОДАВАТЕЛЬ: Гапанюк\_Ю .Е

Москва - 2024

## Заданте

## Решение задачи классификации текстов.

Необходимо решить задачу классификации текстов на основе любого выбранного Вами датасета (кроме примера, который рассматривался в лекции). Классификация может быть бинарной или многоклассовой. Целевой признак из выбранного Вами датасета может иметь любой физический смысл, примером является задача анализа тональности текста.

Необходимо сформировать два варианта векторизации признаков - на основе CountVectorizer и на основе TfidfVectorizer.



