|  |  |  |
| --- | --- | --- |
|  | | |
| Федеральное государственное бюджетное  образовательное учреждение высшего образования «Новосибирский государственный технический университет» | | |
|  | | |
| Кафедра прикладной математики | | |
| Лабораторная работа № 1 | | |
| по дисциплине «Программные системы статистического анализа» | | |
| **,** | | |
|  | Вариант 52 | ПМИМ-51 Буданцев Дмитрий |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  | Преподаватели | Тимофеева Анастасия Юрьевна |
|  |  |  |
| Новосибирск, 2025 | | |



1. Постановка задачи
   1. Цель

Сравнение метрик, реализованных для вычисления расстояния между строками в stringdist.

* 1. Задачи

**1**. **Часть 1**: Базовое сравнение на простых примерах

**2**. **Часть 2**: Анализ чувствительности к разным типам различий

1. Опечатки

2. Разная длина строк

3. Перестановки символов

4. Разный регистр символов

**3**. **Часть 3**: Производительность метрик

1. Теоретическая часть
2. Введение
3. Расстояние Левенштейна
4. Определение
5. Операции редактирования
6. Пример вычисления
7. Расстояние Дамерау-Левенштейна
8. Определение
9. Дополнительная операция
10. Пример вычисления
11. Расстояние Хэмминга
12. Определение
13. Примеры вычисления
14. Расстояние Джакарда
15. Определение
16. Пример вычисления
17. Параметр q и его влияние
18. Косинусное расстояние
19. Определение
20. Векторное представление строк
21. Q-граммное расстояние
22. Определение
23. Расстояние Jaro
24. Алгоритм вычисления
25. Расстояние Jaro-Winkler
26. Формула вычисления
27. Расстояние LCS
28. Определение
29. Пример вычисления
30. Практическая часть