

Университет ИТМО, факультет программной инженерии и компьютерной техники
Двухнедельная отчётная работа по «Информатике»: аннотация к статье

Дата прошедшей лекции	Номер прошедшей лекции	Название статьи/главы книги/видеолекции	Дата публикации (не старше 2021 года)	Размер статьи (от 400 слов)	Дата сдачи
11.09.2024	1	Логические и арифметические операции в системе счисления фибоначчи и их применение в вычислительных системах	26.11.2022	~2002	25.09.2024
25.09.2024	2	Анализ корректирующей способности кодов Рида-Соломона в системах передачи данных	12.04.2021	~1728	09.10.2024
09.10.2024	3	Обработка естественного языка	26.05.2021	~2010	23.10.2024
	4				
	5				
	6				
	7				

Выполнил(а)

Мельник Ф.А.
Фамилия И.О. студента

, № группы P3106, оценка

не заполнять

Прямая полная ссылка на источник или сокращённая ссылка (bit.ly, tr.im и т.п.)

https://elibrary.ru/download/elibrary_47129513_24144570.pdf

Теги, ключевые слова или словосочетания (минимум три слова)

ОБРАБОТКА ТЕКСТА, МАШИННОЕ ОБУЧЕНИЕ, РЕГУЛЯРНЫЕ ВЫРАЖЕНИЯ, PYTHON, NLP

Перечень фактов, упомянутых в статье (минимум четыре пункта)

1. Обработка естественного языка (NLP) используется для анализа текстов и речи с помощью алгоритмов машинного обучения.
2. Этапы обработки текста включают токенизацию, нормализацию, фильтрацию и векторизацию.
3. Существует модель "мешок слов", она определяет частоту встречаемости каждого слова в тексте, игнорируя порядок слов.
4. TF-IDF оценивает важность слов в тексте по их частоте в документе и общем количестве документов.

Позитивные следствия и/или достоинства описанной в статье технологии (минимум три пункта)

1. NLP позволяет эффективно анализировать большие текстовые данные, такие как статьи, отзывы или документы.
2. NLP помогает классифицировать текстовые данные, например, для автоматической сортировки электронных писем или сообщений.
3. NLP может улучшить поиск информации, делая его более точным и релевантным запросам пользователя.

Негативные следствия и/или недостатки описанной в статье технологии (минимум три пункта)

1. Технологии NLP могут неправильно интерпретировать значение слов в сложных контекстах, что приводит к неточному анализу и результатам.
2. Неформальные выражения, сленг и региональные языковые особенности могут быть плохо обработаны, что снижает точность работы моделей.
3. Алгоритмы NLP менее эффективны, если обучаются на малом объеме данных, что ограничивает их применение для узкоспециализированных задач.

Ваши замечания, пожелания преподавателю или анекдот о программистах¹