

Тест начат	среда, 10 января 2024, 21:53
Состояние	Завершены
Завершен	среда, 10 января 2024, 22:15
Прошло времени	21 мин. 42 сек.
Оценка	7,00 из 7,00 (100%)

Вопрос **Инфо**

Вопрос 1
Выполнен
Баллов: 1,00 из 1,00

Отметьте формальные языки.

Выберите один или несколько вариантов ответа.

Выберите один или несколько ответов:

- ☐ а. Латинский язык
- ☒ b. Язык программирования Lisp
- ☒ с. Язык программирования Java
- ☐ d. Русский язык, на котором написаны юридические документы

Ваш ответ **верный**.

Вопрос 2

Выполнен

Баллов: 1,00 из 1,00

У Вас есть коллекция текстов и Вам нужно написать решающее правило, чтобы выбрать из них те, в которых идёт речь про кошек.

Выберите вариант решения задачи, наиболее экономный с точки зрения количества ручной работы и при этом надёжный.

Подразумевайте, что у Вас имеется реализация лемматизации и стемминга для русского языка.

В настоящем задании приведена крайне упрощённая задача классификации текста. Постарайтесь поразмышлять, как разные этапы лингвистического анализа могут упростить нам, как разработчикам, жизнь, вне зависимости от того, используем мы нейросети или нет.

Выберите один ответ:

- ☒ a. Разбить текст на токены, применить лемматизацию и сформулировать правило как "Текст про кошек, если после лемматизации в тексте есть токен 'кошка'."
- ☐ b. Взять исходный текст как есть и сформулировать правило как "Текст про кошек, если в тексте встречается одна из следующих подстрок: кошка, кошки, кошке, кошкой, ... кошками, ..."
- ☐ c. Разбить текст на токены, применить стемминг и сформулировать правило как "Текст про кошек, если в тексте после стемминга есть токен 'кош'."

Ваш ответ верный.

Вопрос **3**

Выполнен

Баллов: 1,00 из 1,00

Выберите начальную форму (лемму) для каждой словоформы, учитывая информацию о морфологических признаках.

Словоформа "сорок", часть речи "существительное"

Словоформа "три", часть речи "числительное"

Словоформа "три", часть речи "глагол"

Словоформа "сорок", часть речи "числительное"

Ваш ответ **верный**.

Вопрос 4

Выполнен

Баллов: 1,00 из 1,00

Допустим, Вы разрабатываете поисковую машину и в процессе разработки и отладки оцениваете качество работы алгоритма на некоторых примерах "запрос - текст - оценка сходства":

Запрос	Текст	Оценка сходства
Как прибраться квартиру быстро?	В этой статье мы рассмотрим, как быстро прибраться квартиру.	Высокое сходство
Как прибраться квартиру быстро?	Мошенники быстро решили прибраться к рукам бесхозную квартиру.	Низкое сходство

Какой набор методов обработки текста мог бы обеспечить правильную оценку сходства для приведённых примеров в Вашей поисковой машине?

Выберите один ответ:

- ☐ a. Лемматизация, оценка сходства по количеству общих слов
- ☐ b. Стемминг, оценка сходства по количеству общих слов
- ☐ c. Синтаксический анализ и лемматизация, оценка сходства по количеству общих слов и словосочетаний
- ☒ d. Word Sense Disambiguation, оценка сходства по количеству общих смыслов

Ваш ответ верный.

Вопрос Инфо

Вопрос 5

Выполнен

Баллов: 1,00 из 1,00

Отметьте предложения, в которых имеет место **частеречная** омонимия.

Выберите один или несколько вариантов ответа.

Выберите один или несколько ответов:

- ☐ а. На дворе дрова.
- ☒ b. Ведро дало течь, вода стала течь.
- ☒ с. Сорок сорок сидели на дереве.
- ☐ d. Глокая куздра штеко будланула бокра.
- ☐ e. Пятнадцать птиц сидели на дереве.
- ☒ f. Мама любит печь и поэтому поставила в печь пирожки.

Ваш ответ верный.

Вопрос Инфо

Вопрос 6

Выполнен

Баллов: 1,00 из 1,00

Сопоставьте особенности языков и их описания.

Флективность

“интенсивность” словоизменения в зависимости от ситуации посредством суффиксов и окончаний

Смысловая
омонимия

одна и та же словоформа может обозначать разные объекты или явления (денотаты) в зависимости от контекста

Частеречная
омонимия

слова с разной частью речи могут иметь одинаковую словоформу

Ваш ответ верный.

Вопрос 7

Выполнен

Баллов: 1,00 из 1,00

Предположим, Вы делаете систему анализа отзывов о мобильных телефонах в Интернете. Такая система собирает отзывы по нужной модели телефона, а потом из каждого отзыва извлекает пары "характеристика телефона - оценка" (это называется аспектным сентимент анализом).

Как обычно, в процессе разработки и отладки Вы пользуетесь выборкой примеров.

Текст	Характеристика	Оценка
Экран, в отличие от ёмкости батареи, отличный.	экран	отличный
Экран, в отличие от ёмкости батареи, отличный.	ёмкость батареи	низкая
Экран отличный, но плохая батарея	экран	отличный
Экран отличный, но плохая батарея	батарея	плохая

Выберите группу методов обработки текстов, алгоритмы из которой помогут правильно связать характеристику и её оценку, несмотря на свободный порядок слов.

Выберите один ответ:

- ☐ a. Лемматизация
- ☐ b. Word sense disambiguation
- ☐ c. POS-теггинг
- ☐ d. Стемминг
- ☒ e. Синтаксический анализ

Ваш ответ верный.

◀ 1.2 В общих чертах: естественный язык и текст

Перейти на...

1.4 В общих чертах: Лингвистический анализ ►