

ЗАДАНИЯ
по лабораторному практикуму
по курсу «Основы компьютерной лингвистики»

Для выполнения лабораторных работ создается собственный датасет, на базе которого ставятся эксперименты. Предварительно датасет согласуется с преподавателем и далее студент приступает к выполнению лабораторного практикума.

Перечень лабораторных работ, подлежащих выполнению, выглядит следующим образом:

1. Лабораторная работа №1 «Графематический анализ»

Задание. Изучить пакеты для токенизации текста библиотеки NLTK, подсоединить библиотеку стоп слов и удалить их. Операции проделать на личном дата сете. Объяснить проделанные операции на конкретных примерах.

2. Лабораторная работа №2 «Морфологический анализ»

Задание. Изучить пакеты для морфологического разбора слов библиотеки NLTK. Провести морфологический разбор слов предложений. Уметь пояснить основные понятия из морфологии и полученных результатов. Операции провести на личном дата сете.

3. Лабораторная работа №3. «Синтаксический анализ».

Задание. Изучить пакеты для синтаксического разбора слов библиотеки NLTK и Natasha. Провести синтаксический разбор предложений. Уметь пояснить основные понятия из синтаксического анализа, полученные результаты. Операции провести на предложениях личного дата сета. Сравнить результаты NLTK и Natasha.

4. Лабораторная работа №4. «Проверка проективности предложений»

Задание. Выбрать пять предложений из личного дата сета. Проверить их проективность, построить для них деревья зависимостей и отрезочные представления в виде системы составляющих. Проверить проективность на ПО и вручную

5. Лабораторная работа №5. «Функционал смысла»

Задание. Выбрать пять предложений из личного дата сета. Построить для них семантические схемы.

6. Лабораторная работа №6. «Функционал смысла»

Задание. Выбрать пять предложений из личного дата сета (результаты предыдущей лабораторной работы). По их семантическим схемам определить смысловую (семантическую) близость для каждой пары предложений.

7. Лабораторная работа №7. «Нечеткая смысловая классификация предложений в семантическом информационном поиске».

Задание. Выполнить нечеткую классификацию предложений на соответствие запросу.

Составил:

д.т.н., профессор

Ю.М. Вишняков