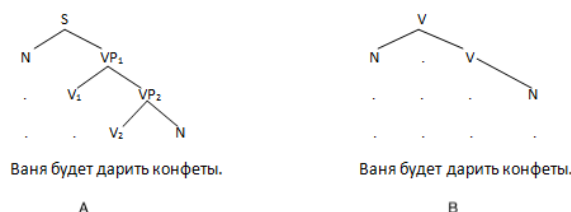


Анализ неструктурированных данных
Самостоятельная работа
Максимальная оценка: 10 баллов

Имя и фамилия: _____
Группа: _____

Ответьте на вопросы ниже. Можно использовать компьютер для вычислений и для просмотра лекций и учебников. Запишите решения, выделите ответы и прокомментируйте каждый шаг решения.

Каждый вопрос оценивается в 2 балла.



1. На каком рисунке (А или В) изображено дерево зависимостей?

Ответ: На рисунке А дерево составляющих, на рисунке В – дерево зависимостей.

2. Даны два предложения. Отобразите их структуру зависимостей (линейно, над словами) и напишите какое из них непроективное и почему.

Я хочу изучать синтаксис.

Я синтаксис хочу изучать.

Ответ: правое предложение непроективное. Стрелки (хочу – я) и (изучать – синтаксис) пересекаются.

3. Зачем нужен beam search для transition-based парсинга?

- А Beam search увеличивает количество возможных разборов и позволяет выбрать один из лучших
- В Beam search уменьшает количество возможных разборов и сокращает время, затрачиваемое на построение дерева зависимостей
- С Beam search – это аналог ЕМ алгоритма, используется для обучения синтаксического парсера без учителя
- D Beam search нужен для поиска по трибанкам.

Ответ: А

4. Может ли метрика качества синтаксического парсера LAS быть больше метрики UAS?

Ответ: Нет, не может.

5. Что из перечисленного ниже не входит в пайплайн UDPipe?

- А Сегментация на предложения
- В Токенизация
- С Извлечение именованных сущностей

D Лемматизация

E Стемминг

F Морфологический теггинг

G Построение дерева зависимостей

H Построение дерева составляющих

Ответ: C, E, H