

Направление: «Фундаментальная и прикладная лингвистика»

Профиль: «Компьютерная лингвистика»

Критерии

1. Задание 1. – 20 баллов

Неверный ответ -- 0.

Верный ответ без обоснования -- 5.

Верный ответ с недостаточным обоснованием или два или более неверных утверждения -- 10.

Неверное утверждение существенно используется в рассуждении -- 10.

Есть одно неверное утверждение, после вычеркивания которого остается верное рассуждение -- 15.

Есть одно неверное утверждение, которое по существу не используется в рассуждении -- 17.

Верное рассуждение -- 20.

2. Задание 2. – 45 баллов

2а) выделены все 5 значений – 5 баллов;

за каждое пропущенное значение штраф – один балл;

значения должны быть основаны на тех примерах, которые были приведены в задании

Балл снижается, если выделены не все значения

2б) максимальная оценка – 5 баллов

за правильно выполненное второе задание максимальная оценка – 5 баллов

Правильно выполненное задание предполагает:

а) правильно выделены семантические роли,

б) примеры размечены по семантическим ролям, модели управления глагола в конкретном примере представлены: в каждом из размеченных примеров указана семантическая роль

в) для каждого актанта глагола указана синтаксическая роль этой именной группы, указано ее морфологическое оформление

Комментарий:

должно быть хотя бы отдаленное представление того, что такое семантические роли (можно не знать правильные названия, но тогда одним и тем же семантическим ролям должен быть присвоен один и тот же ярлык, должно быть пояснение, каковы признаки роли с таким ярлыком)

должна быть описана модель управления так, как сказано в задании: указаны синтаксические, морфологические, семантические ограничения

NB: семантические ограничения тоже важны для выполнения задания по разрешению семантической неоднозначности

2в) максимальная оценка – 5 баллов

должны быть перечислены, по крайней мере, следующие модули: токанизация и сегментация на предложения, морфологический анализ, модуль распознавания семантических ролей или аналогичный модуль – тот компонент системы, в котором предполагается решать задачу распознавания семантических ролей и разрешения семантической неоднозначности

в соответствии с заданием должна присутствовать краткая характеристика модуля: что на входе, что на выходе, каков «вклад» в конечную задачу

2г) максимальная оценка – 10 баллов

правила могут быть приведены в произвольной форме, но они должны быть приведены, обоснованы, приведены исключения из правила с конкретными примерами на русском языке, на которых сформулированное правило даст ошибку

NB Для идеального ответа из правил должно быть видно, какова роль предложенных в 2в) модулей

2д) аналогично г)

Бонус: при разрешении семантической неоднозначности важны семантические ограничения на актанты глагола. Например, если подлежащее неодушевленное, то не может быть значения «передавать нечто в собственность другому лицу», если для предложной группы – семантическое ограничение «часть тела» - то значение «физическое воздействие», как в «Дал в нос»

2е) максимальная оценка 10 баллов

в задании указано, что должны быть указаны объекты машинного обучения (например, предложения с глаголом или фрагменты предложений с глаголом);

должны быть четко указаны признаки, которые планируется использовать (например, +2/-2 крокетных словоформы, стоящих слева и справа от глагола, падежи существительных, находящихся в предложении/в контексте длиной n, тип синтаксического отношения..., семантическая роль именной группы и т.п.).

Требования к указанным признакам: должны быть указаны конкретные признаки, расплывчатый ответ штрафуются, должно быть указано, каким из перечисленных в 2в) модулей признаки могут быть получены; для «идеального» ответа про признаки набор признаков должен как-то соотноситься с перечисленными в 2в) модулями

Распространенная ошибка, которая штрафовалась: в в) перечислены модули системы, но ни в г), ни в д), ни в е) не указано, как результаты работы этих модулей участвуют в конечной задаче (например, указано, что нужен синтаксический анализатор, но ни в одном из правил или признаков синтаксическая информация никак не задействована)

нет четкого описания объектов и признаков и объяснений, как они получены.

3 нет упоминания основных моментов: 1) ограничение на память и вычислительные ресурсы, 2) необходимость уменьшения размерности 3) краткое описание алгоритма.

собственные комментарии: 1) нужна конструктивная оценка алгоритма (достоинства и недостатки), 2) предложения по улучшению, 3) применение в других задачах.

Задание 3. – 35 баллов

1) в реферате должны быть отражены следующие идеи:

а) мотивация разработки предлагаемого в блоге решения - ограничение на память и вычислительные ресурсы,

б) проблема размерности данных: необходимость уменьшения размерности

в) краткое описание алгоритма.

2) собственные комментарии:

а) нужна конструктивная оценка алгоритма (достоинства и недостатки),

б) предложения по улучшению,

в) применение в других задачах.

В правильном ответе достаточно было указать б) или в), но в данных пунктах должны быть указаны какие-то конкретные предложения

Не учитывался как положительный ответ, в котором содержалась только эмоциональная оценка