УТВЕРЖДАЮ

Зав. кафедрой СП ЮУрГУ

\_\_\_\_\_\_\_\_\_Л.Б. Соколинский

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2022 г.

Фонд оценочных средств

ООП «Инженерия информационных и интеллектуальных систем»

по направлению 09.03.04 – Программная инженерия

Дисциплина «Введение в обработку естественного языка»

|  |  |  |  |
| --- | --- | --- | --- |
| № КМ | Вид контроля | Название контрольного мероприятия | Оценочные средства |
| 1 | Текущий контроль | Проверка  практического задания 1 | Вопросы для подготовки к устному собеседованию:  1. Что такое POS tagging?  2. Когда и зачем использовать теги PoS?  2. Что такое процесс лемматизации?  3. Что такое стеминг?  4. Что такое морфологический анализ? |
| 2 | Текущий контроль | Проверка  практического задания 2 | Вопросы для подготовки к устному собеседованию:  1. Что такое классификация? Чем она отличается от кластеризации?  2. Назовите основные метрики оценки качества классификации.  3. Опишите модель TF-IDF?  4. Опишите модель BoW?  5. Что такое Деревья решений  6. Что такое наивный байесовский классификатор?  7. Логистическая регрессия в задаче классификации текстов. |
| 3 | Текущий контроль | Проверка  практического задания 3 | Вопросы для подготовки к устному собеседованию:  1. Опишите модель word2vec  2. Опишите модель fasttext  3. Как определяется семантической близости текстов в нейросетевых моделях  4. Что такое дестрибутивная семантика.  5. Опишите процесс классификация текстов на основе нейросетевых моделей языка. |
| 4 | Текущий контроль | Проверка  практического задания 4 | Вопросы для подготовки к устному собеседованию:  1. Что такое кластеризация? Чем она отличается от классификации.  2. Основные метрики оценки качества кластеризации.  3. Что такое тематическое моделирование? Какие методы ТМ вы знаете.  4. Опишите метод LDA  5. Опишите метод pLSA  6. Опишите библиотеку BigARTM |
| 5 | Текущий контроль | Проверка  практического задания 5 | Вопросы для подготовки к устному собеседованию:  1. Опишите архитектуру CNN.  2. Опишите архитектуру LSTM.  3. Опишите различие архитектур CNN и LSTM |
| 6 | Текущий контроль | Проверка  практического задания 6 | Вопросы для подготовки к устному собеседованию:  1. В чем заключается задача автоматического реферирования.  2. Что такое encoder?  3. Что такое decoder?  4. Что такое механиз внимания (attention)?  5. В чем заключается задача автоматического реферирования (text summarization). |
| 7 | Текущий контроль | Проверка  практического задания 7 | Вопросы для подготовки к устному собеседованию:  1. Опишите архитектуру BERT?  2. Опишите архитектуру GPT?  3. Что такое Transfer learning?  4. Что такое Self-Attention?  5. Опишите процесс Fine-tuning моделей на архитектуре трансформеров. |
| 9 | Проме- жуточная аттестация | Письменный экзамен | Вопросы для подготовки к экзамену:  1. Основные задачи NLP.  2. Представления текстовых данных.  3. Предобработка текста, лемматизация, стемминг.  4. Методы машинного обучения для классификации текстовых документов на основе  частотных мер (TF-IDF).  5. Деревья решений  6. наивный байесовский классификатор  7. логистическая регрессия в задаче классификации текстов.  8. Языковые модели. Word embeddings.  9. Нейросетевые модели языка: word2vec, fasttext. Мера семантической близости.  10. Классификация текстов на основе нейросетевых моделей языка.  11. Кластеризация текстовых документов.  12. Тематическое моделирование Методы LSA, pLSA.  13. Аддитивная регуляризация тематических моделей в BigARTM  14. Классификация текстов с помощью глубоких нейронных сетей: CNN, LSTM.  15. Задачи обработки последовательностей: машинной перевод, автоматическое  реферирование (summarization), вопросно-ответные системы.  16. Механизм внимания (attention).  17. Архитектуры encoder-decoder-attention.  18. Transfer learning в задачах анализа текстов. Self-Attention.  19. Архитектуры трансформеров: BERT, GPT в задачах классификации текстов, предсказания  пропущенных слов, генерации текстов. Fine-tuning трансформеров.  20. Построение диалоговых систем. Архитектура диалоговых систем.  21. Модули понимания естественного языка (NLU) и диалоговый менеджер (DM).  22. Сложности построения диалоговых систем.  23. Проектирование UX/UI диалоговых ассистентов в чатах и голосе.  24. Обзор современных фреймворков для построения диалоговых систем: DeepPavlov, Rasa,  Just AI Conversational Platform |

Паспорт фонда оценочных средств приведен в п. 6.3 РПД.

Разработчик И.Е. Николаев

|  |
| --- |
| ФГАОУ ВО «Южно-Уральский государственный университет  (национальный исследовательский университет)»  Кафедра системного программирования  Дисциплина «Введение в обработку естественного языка»  БИЛЕТ № 1   1. Кластеризация текстовых документов. 2. Transfer learning в задачах анализа текстов. Self-Attention.   Ст.преподаватель И.Е. Николаев |
| ФГАОУ ВО «Южно-Уральский государственный университет  (национальный исследовательский университет)»  Кафедра системного программирования  Дисциплина «Введение в обработку естественного языка»  БИЛЕТ № 2   1. Деревья решений 2. Тематическое моделирование Методы LSA, pLSA.   Ст.преподаватель И.Е. Николаев |