

Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»

Уфимский филиал Финуниверситета

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине «Обработка текстов на естественных языках»

Разработчик: кафедра «Математика и информатика»

Направления подготовки: 09.03.03 Прикладная информатика

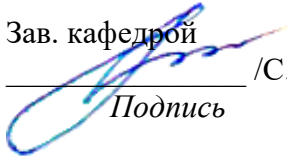
Образовательная программа: Прикладная информатика

Профиль: ИТ-сервисы и технологии обработки данных в экономике и финансах

Форма образования: заочная

РАССМОТРЕН
На заседании кафедры
«Математика и информатика»

Протокол № 12
от « 30 » июня 2023 г.

Зав. кафедрой

_____/С.А. Фархиева
Подпись

Разработан на основе

*ОС ФГОС ВО по направлению подготовки
09.03.03 Прикладная информатика
(уровень бакалавриата)
№ 922 от 19.09.2017 г.*

Паспорт фонда оценочных средств

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины Обработка текстов на естественных языках.

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля и промежуточной аттестации.

1 Описание показателей и критериев оценивания компетенций

| Планируемые результаты освоения компетенции (индикаторы достижения компетенции) | Уровень освоения | | | | Оценочное средство ¹ |
|--|---|---|--|--|--|
| | «неудовлетворительно» (не зачтено) минимальный не достигнут | «удовлетворительно» (зачтено) минимальный пороговый | «хорошо» (зачтено) средний | «отлично» (зачтено) высокий | |
| ОПК-7 Способен разрабатывать алгоритмы и программы, пригодные для практического применения | | | | | |
| Индикатор 1. Разрабатывает алгоритмы решения прикладных задач с использованием математических и современных аналитических методов. | | | | | |
| <u>Знать:</u> методы разработки алгоритмов решения прикладных задач с использованием математических и современных аналитических методов обработки текстов на естественных языках. | не демонстрирует достаточных знаний основ и методов обработки текстов на естественных языках. | основные понятия и простые методы обработки текстов на естественных языках. | методы обработки текстов на естественных языках, включая статистический и базовый уровень машинного обучения. | основы и передовые методы обработки текстов на естественных языках, включая морфологический, синтаксический, семантический анализ и методы машинного обучения. | Тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач |
| <u>Уметь:</u> разрабатывать алгоритмы решения прикладных задач с использованием математических и современных аналитических методов обработки текстов на естественных языках. | не способен эффективно применять алгоритмы для решения прикладных задач. | следовать готовым алгоритмам для выполнения базовых операций обработки текстов и применять их в практических задачах. | разрабатывать алгоритмы для анализа текстов и применять стандартные инструменты для решения типовых задач обработки данных в экономике и финансах. | разрабатывать и оптимизировать сложные алгоритмы для обработки и анализа текстов, а также эффективно применять их для решения конкретных задач в области экономики и финансов. | Тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач |
| Индикатор 2. Реализует алгоритмы с использованием современных средств разработки прикладного программного обеспечения. | | | | | |

¹ Виды оценочных средств: *тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач, мини-кейсы, ситуационные задачи, практико-ориентированные задания.*

| Планируемые результаты освоения компетенции (индикаторы достижения компетенции) | Уровень освоения | | | | Оценочное средство ¹ |
|---|---|--|--|---|--|
| | «неудовлетворительно» (не зачтено) минимальный не достигнут | «удовлетворительно» (зачтено) минимальный пороговый | «хорошо» (зачтено) средний | «отлично» (зачтено) высокий | |
| <u>Знать:</u> алгоритмы с использованием современных средств разработки прикладного программного обеспечения для обработки текстов на естественных языках. | не обладает необходимыми знаниями для реализации алгоритмов обработки текстов с использованием современных программных средств и их практического применения. | базовые концепции и простейшие методы обработки текстов на естественных языках. | существующие методы и инструменты для обработки текстов на естественных языках. | передовые техники обработки текстов на естественных языках включая глубокое обучение и высокоуровневые фреймворки. | Тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач |
| <u>Уметь:</u> реализовывать алгоритмы с использованием современных средств разработки прикладного программного обеспечения для обработки текстов на естественных языках. | не обладает необходимыми навыками для реализации алгоритмов обработки текстов с использованием современных программных средств и их практического применения. | использовать готовые программные решения для выполнения стандартных задач обработки текстов. | реализовывать алгоритмы обработки текстов с использованием стандартных средств программирования и библиотек. | самостоятельно реализовывать и интегрировать сложные алгоритмы обработки естественного языка, используя современные среды разработки. | Тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач |
| Индикатор 3. Владеет навыками тестирования программного обеспечения, необходимыми для создания программных продуктов промышленного качества. | | | | | |
| <u>Знать:</u> методы тестирования программного обеспечения, необходимыми для создания программных продуктов обработки текстов на естественных языках промышленного качества. | не имеет достаточных знаний о методах тестирования. | базовые понятия тестирования программного обеспечения. | основные принципы и методы тестирования программного обеспечения. | современные методы и инструменты тестирования программного обеспечения, включая автоматическое и ручное тестирование. | Тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач |
| <u>Уметь:</u> тестировать программное обеспечение, необходимое для создания программных продуктов обработки текстов на естественных языках промышленного качества. | не имеет достаточных знаний о методах тестирования | выполнять простые виды тестирования и исправлять наиболее очевидные ошибки в программном коде. | применять стандартные подходы и инструменты для обнаружения и устранения ошибок в программных продуктах. | проводить всестороннее тестирование и отладку программных продуктов, обеспечивая их надёжность и соответствие требованиям качества. | Тестовые задания, вопросы для устного/письменного опроса, задания в виде расчетных задач |

2. Оценочные средства для оценки сформированности компетенций (контроль остаточных знаний)

Примеры тестовых заданий

Тесты (ОПК-7)

1. Вопрос: Что такое NLP (Natural Language Processing)? (ОПК-7)
 - 1) Обработка символьных данных
 - 2) Обработка естественного языка
 - 3) Нейронное программирование
 - 4) Работа с числовыми данными
2. Вопрос: Как называется процесс преобразования слов в числовые значения для обработки компьютером? (ОПК-7)
 - 1) Токенизация
 - 2) Лемматизация
 - 3) Векторизация
 - 4) Аннотирование
3. Вопрос: Какой метод используется для определения корня слова? (ОПК-7)
 - 1) Токенизация
 - 2) Стемминг
 - 3) Векторизация
 - 4) Классификация
4. Вопрос: Что из перечисленного является задачей классификации текста? (ОПК-7)
 - 1) Машинный перевод
 - 2) Распознавание речи
 - 3) Сентимент-анализ
 - 4) Обучение без учителя
5. Вопрос: Какой компонент языковой модели отвечает за проверку правильности грамматических конструкций? (ОПК-7)
 - 1) Лексический анализатор
 - 2) Синтаксический анализатор
 - 3) Семантический анализатор
 - 4) POS-теггер
6. Вопрос: Что такое "корпус" в обработке естественных языков? (ОПК-7)
 - 1) Программа для анализа текста
 - 2) Коллекция документов для машинного обучения
 - 3) Словарь синонимов и антонимов

4) Встроенный модуль в язык программирования

7. Вопрос: Какой метод часто используется для извлечения основной информации из текста? (ОПК-7)

- 1) Кластеризация
- 2) Извлечение сущностей
- 3) Глубинное обучение
- 4) Стемминг

8. Вопрос: Какая модель обработки языка основана на двунаправленных трансформерах? (ОПК-7)

- 1) GloVe
- 2) FastText
- 3) BERT
- 4) Word2Vec

9. Вопрос: Что такое токенизация в контексте обработки естественных языков? (ОПК-7)

- 1) Процесс идентификации валюты
- 2) Разбиение текста на абзацы
- 3) Разбиение текста на предложения или слова
- 4) Шифрование данных

10. Вопрос: Что такое семантическое выравнивание в области NLP? (ОПК-7)

- 1) Выравнивание текста в документе
- 2) Сопоставление значений слов в разных языках
- 3) Изменение структуры текста
- 4) Выделение ключевых слов

11. Вопрос: Как называется процесс извлечения информации из неструктурированного текста и преобразования ее в структурированную форму? (ОПК-7)

12. Вопрос: Как называется модель, используемая для предсказания следующего слова в последовательности на основе предыдущих слов? (ОПК-7)

13. Вопрос: Какой алгоритм применяется для машинного перевода и оценки сходства предложений путем преобразования слов в векторные пространственные представления? (ОПК-7)

14. Вопрос: Как называется технология, позволяющая компьютерам интерпретировать и обрабатывать естественные языки аналогично тому, как это делает человек? (ОПК-7)

15. Вопрос: Какой инструмент используется для разметки частей речи в тексте? (ОПК-7)

Ключ к тесту

| Вопрос | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--------|---|---|---|---|---|---|---|---|---|----|--------------------------------|----------------------------------|----------|--|------------------------------------|
| Ответ | 2 | 3 | 2 | 3 | 2 | 2 | 2 | 3 | 3 | 2 | Текстовый анализ (Text Mining) | N-граммная модель (N-gram Model) | Word2Vec | Обработка естественного языка (Natural Language Processing, NLP) | POS-тэггер (Part-of-Speech Tagger) |
| Баллы | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |

3 Методические материалы, определяющие процедуры оценивания знаний и умений, характеризующих степень сформированности компетенций

Критерии оценки знаний при проведении устного/письменного опроса

Оценка **«отлично»** (зачтено) – выставляется обучающемуся, показавшему всесторонние, систематизированные, глубокие знания вопросов дисциплины.

Оценка **«хорошо»** (зачтено) – выставляется обучающемуся, если он твердо знает материал, грамотно и по существу излагает его, но допускает в ответе некоторые неточности, которые может устранить с помощью дополнительных вопросов преподавателя.

Оценка **«удовлетворительно»** (зачтено) – выставляется обучающемуся, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными понятиями, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации.

Оценка **«неудовлетворительно»** (не зачтено) – выставляется обучающемуся, который не знает большей части основного содержания вопросов тем дисциплины, допускает грубые ошибки в формулировках основных понятий.

Критерии оценки знаний при решении задач

Оценка **«отлично»** (зачтено) – выставляется обучающемуся, показавшему всесторонние, систематизированные, глубокие знания вопросов дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка **«хорошо»** (зачтено) – выставляется обучающемуся, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе некоторые неточности, которые может устранить с помощью дополнительных вопросов преподавателя.

Оценка **«удовлетворительно»** (зачтено) – выставляется обучающемуся, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные форму-

лировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными понятиями, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации.

Оценка **«неудовлетворительно»** (не зачтено) – выставляется обучающемуся, который не знает большей части основного содержания вопросов тем дисциплины, допускает грубые ошибки в формулировках основных понятий, не умеет использовать полученные знания при решении типовых практических задач.

Критерии оценки знаний при проведении тестирования

Оценка **«отлично»** (зачтено) выставляется при условии правильного ответа студента не менее чем на 85 % тестовых заданий;

Оценка **«хорошо»** (зачтено) выставляется при условии правильного ответа студента не менее чем на 70 % тестовых заданий;

Оценка **«удовлетворительно»** (зачтено) выставляется при условии правильного ответа студента не менее чем на 51 %;

Оценка **«неудовлетворительно»** (не зачтено) выставляется при условии правильного ответа студента менее чем на 50 % тестовых заданий.