

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Таскаев Сергей Валерьевич
Должность: Ректор
Дата подписания: 15.09.2025 11:16:51
Уникальный программный ключ:
04c19ed8bf598f3b6c077a486b986788b8522523

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)

Фонд оценочных средств для промежуточной аттестации по дисциплине
«Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02
«Прикладная математика и информатика» направленности (профилю) Математическое
моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»

стр. 1

**Фонд оценочных средств
для промежуточной аттестации
по дисциплине**

Интеллектуальный анализ текстов

**Направление подготовки (специальность)
01.04.02 Прикладная математика и информатика**

Направленность (профиль)

Математическое моделирование и искусственный интеллект

**Присваиваемая квалификация
Магистр**

**Форма обучения
Очная**

Челябинск 2025

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 2

Содержание

1. Паспорт фонда оценочных средств
2. Перечень формируемых компетенций
 - 2.1. Компетенции, закреплённые за дисциплиной
3. Содержание оценочных средств по дисциплине
 - 3.1. Виды оценочных средств
 - 3.2. Содержание оценочных средств
4. Порядок проведения и критерии оценивания промежуточной аттестации
 - 4.1. Порядок проведения промежуточной аттестации
 - 4.2. Критерии оценивания промежуточной аттестации по видам оценочных средств.
 - 4.3. Результаты промежуточной аттестации и уровни сформированности компетенций

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 3

ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ

Направление подготовки: 01.04.02 «Прикладная математика и информатика»
 Направленность (профиль) Математическое моделирование и
 искусственный интеллект

Дисциплина: Интеллектуальный анализ текстов

Семестр (семестры) изучения: Семестр № 4

Форма (формы) промежуточной аттестации: Экзамен

2. ПЕРЕЧЕНЬ ФОРМИРУЕМЫХ КОМПЕТЕНЦИЙ

2.1. Компетенции, закреплённые за дисциплиной

Изучение дисциплины «Интеллектуальный анализ текстов»
 направлено на формирование следующих компетенций:

Коды компетенции согласно ФГОС (ОПОП ВО)	Содержание компетенций согласно ФГОС (ОПОП ВО)	Индикаторы достижения компетенции согласно ОПОП	Перечень планируемых результатов обучения по дисциплине
1	2	3	4
ПК-2	Способен исследовать и разрабатывать архитектуры систем искусственного интеллекта для различных предметных областей на основе комплексов методов и инструментальных средств систем искусственного интеллекта	ПК-2.1. Исследует и разрабатывает архитектуры систем искусственного интеллекта для различных предметных областей	Знать: архитектуры систем в области интеллектуального анализа текстов Уметь: Разрабатывать архитектуры искусственного интеллекта для интеллектуального анализа текстов

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 4

3. СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ

3.1 Виды оценочных средств

№ п/п	Код компетенции/ планируемые результаты обучения	Контролируемые темы/ разделы	Наименование оценочного средства для текущего контроля	Наименование оценочного средства на промежуточной аттестации/№ задания
1	ПК-2/Способен исследовать и разрабатывать архитектуры систем искусственного интеллекта для различных предметных областей на основе комплексов методов и инструментальных средств систем искусственного интеллекта	Задачи обработки естественного языка	Лабораторные работы	Экзамен
2		Предобработка текстов	Лабораторные работы	Экзамен
3		Статистические модели языка	Лабораторные работы	Экзамен
4		Нейросетевые языковые модели	Лабораторные работы	Экзамен
5		Классификация текстов. Перевод текстов	Лабораторные работы	Экзамен
6		Основные принципы работы трансформеров	Доклад	Экзамен
...				

Типовые задания, критерии и показатели оценивания в рамках текущего контроля представлены в рабочей программе дисциплины (модуля). Полные комплекты оценочных средств и контрольно-измерительных материалов хранятся на кафедре.

3.2 Содержание оценочных средств

Лабораторная работа 1

Введение в векторные представления текстовых данных

Каждый студент получает для выполнения работы фрагмент текста.

Задача: произвести обучение моделей используя библиотеку **Gensim**.

Этапы выполнения работы:

- 1) Установите библиотеку Gensim
- 2) Импортируйте набор данных
- 3) Выполните предобработку текстов
- 4) Проведите обучение модели Word2vec
- 5) Проведите обучение модели Doc2vec
- 6) Выполните проверку моделей
- 7) Сделайте выводы по работе, оформите отчёт.

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 5

Лабораторная работа 2

Реализация алгоритма word2vec, doc2vec. Классические алгоритмы

Каждый студент получает для выполнения работы фрагмент текста

Задача: произвести обучение моделей Word2vec и Doc2vec.

Этапы выполнения работы:

1) Импортируйте набор данных

- 1) Выполните предобработку текстов
- 2) Выполните токенизацию текстов
- 3) Реализуйте модель Word2vec
- 4) Проведите обучение модели Word2vec
- 5) Реализуйте модель Doc2vec
- 6) Проведите обучение модели Doc2vec
- 7) Выполните проверку моделей
- 8) Сделайте выводы по работе. Оформите отчёт.

Лабораторная работа 3

Основы нейронных сетей RNN, GRU, LSTM, bi-LSTM. Имплементация сетей на фреймворке Pytorch.

Каждый студент получает для выполнения работы фрагмент текста.

Задача: реализовать нейронные сети RNN, GRU, LSTM, bi-LSTM на фреймворке Pytorch.

Этапы выполнения работы:

- 1) Импортируйте набор данных
- 2) Выполните предобработку текстов
- 3) Выполните токенизацию текстов
- 4) Реализуйте модель RNN
- 5) Проведите обучение модели RNN
- 6) Реализуйте модель GRU
- 7) Проведите обучение модели GRU
- 8) Реализуйте модель LSTM

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 6

- 9) Проведите обучение модели LSTM
- 10) Реализуйте модель bi-LSTM
- 11) Проведите обучение модели bi-LSTM
- 12) Выполните проверку моделей
- 13) Сделайте выводы по работе. Оформите отчёт.

Лабораторная работа 4

Нейронный машинный перевод с последовательностью, вниманием и подсловами

Каждый студент получает для выполнения работы набор данных.

Задача: обучить рекуррентную модель глубокого обучения Sequence-to-sequence(seq2seq).

Этапы выполнения работы:

- 1) Импортируйте набор данных
- 2) Выполните предобработку текстов
- 3) Выполните токенизацию текстов
- 4) Реализуйте модель RNN
- 5) Проведите обучение модели RNN
- 6) Выполните проверку моделей

Сделайте выводы по работе. Оформите отчёт.

Лабораторная работа 5

Глубокие нейронные сети для анализа текстов. Self-supervised обучение (самообучение) и fine-tuning (дообучение под конкретную задачу) с помощью моделей трансформеров

Каждый студент получает для выполнения работы набор данных.

Задача: провести дообучение (fine-tuning) модели Bert на задаче извлечения именованных сущностей NER.

Этапы выполнения работы:

- 8) Импортируйте набор данных для задачи NER
- 9) Загрузите предобученную языковую модель Bert
- 10) Проведите дообучение модели Bert для задачи NER

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 7

11) Выполните проверку модели

12) Сделайте выводы по работе. Оформите отчет.

Примерные темы докладов Самостоятельное изучение принципов работы трансформеров

1. Анализ работы word2vec.
2. Анализ работы Bert.
3. Анализ работы CLIP.
4. Анализ работы Transformers.
5. Анализ работы GPT-2.
6. Анализ работы GPT-3.

Вопросы к экзамену по курсу «Интеллектуальный анализ текстов»

1. Интеллектуальный анализ данных, определение, сущность.
2. Задачи, решаемые Data Mining.
3. Методы Data Mining.
4. Сферы применения Data Mining.
5. Этапы процесса Data Mining. Процесс подготовки данных, понятия качества данных, грязных данных, этапы очистки данных.
6. Сущность и методы кластерного анализа.
7. Обобщенный алгоритм кластерного анализа.
8. Определение мер сходства (метрики).
9. Используемые расстояния между множествами объектов.
10. Виды критериев качества автоматического группирования.
11. Назначение нормировки данных. Способы нормировки исходных данных.
12. Алгоритм иерархического алгоритма кластерного анализа. Приемлемое количество кластеров.
13. Методы объединения объектов в кластеры при иерархическом алгоритме кластерного анализа.
14. Пример ручного расчета кластеризации четырех объектов методом одиночной связи (SingleLinkage).
15. Пример ручного расчета кластеризации четырех объектов методом полной связи (CompleteLinkage).
16. Метод K-средних.
17. Пример кластеризации методом K-средних.
18. Нейронные сети Кохонена, понятие, решаемы задачи.
19. Структура сети Кохонена.
20. Проблема «мертвых» нейронов. Обучение методом выпуклой комбинации
21. Word2vec

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 8

22. Doc2vec
23. GloVe
24. fastText
25. Задача классификации, понятие, виды.
26. Процесс классификации.
27. Конструирование и использование модели классификации.
28. Решение задачи классификации различными методами.
29. Деревья решений, понятие, назначение. Преимущества использования ДР.
30. Алгоритмы построения ДР.
31. Базовые понятия теории деревьев решений.
32. Алгоритм построения деревьев решений.
33. Выбор атрибута ветвления (критерия разбиения).
34. Рекуррентная нейросетевая языковая модель (RNN).
35. Рекуррентная нейронная сеть с долговременной и кратковременной памятью (LSTM).
39. Проблема, связанная с деревьями решений: переобучение (шум). Нормировка.
40. Проблема, связанная с деревьями решений: неизвестные значения атрибутов.
41. Управляемый рекуррентный блок (GRU).
42. Модели трансформеров.
43. Модель Bert.
44. Механизм внимания self-attention в трансформерах.
45. Метод главных компонент, содержание и назначение.
46. Достоинства и Недостатки МГК.
47. Алгоритм построения главных компонент.
48. Графическая интерпретация метода главных компонент.
49. Определение количества главных компонент. Критерий Кайзера. Критерий каменистой осыпи.
50. Определение названия для главных компонент.

Пример экзаменационного билета:

Дисциплина «Интеллектуальный анализ текстов»
Демонстрационный экзаменационный билет

1. Алгоритм word2vec.
2. Лемматизация и стемминг текста
3. Рекуррентная нейронная сеть RNN

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 9

4. ПОРЯДОК ПРОВЕДЕНИЯ И КРИТЕРИИ ОЦЕНИВАНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

4.1. Порядок проведения промежуточной аттестации

На экзамене происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля.

Студент может улучшить свой рейтинг, пройдя контрольное мероприятие промежуточной аттестации, которое не является обязательным. Контрольное мероприятие промежуточной аттестации проводится во время экзамена в виде устного опроса. Студенту выдаётся экзаменационный билет, содержащий 3 вопроса из разных тем курса. Студенту дается 60 минут на подготовку ответов. Затем студент озвучивает свои ответы.

4.2. Критерии оценивания текущей аттестации по видам оценочных средств.

Критерии оценивания лабораторной работы:

Максимальный балл -13

За каждый верно выполненный этап работы начисляется 1 балл. За этапы, выполненные не верно или не выполненные баллы не начисляются.

Критерии оценивания доклада:

Максимальный балл -5

Подготовлен доклад - 1 балл; Подготовлена презентация - 1 балл;

Оформление презентации соответствует ГОСТ - 1 балл; Тема раскрыта - 1 балл;

Доклад вызвал интерес у аудитории - 1 балл.

На каждом из 18 занятий студент может получить 2 балла: Студент задает вопросы по изучаемому материалу - 1 балл; Студент правильно отвечает на вопросы по изучаемому материалу - 1 балл. В противном случае баллы не начисляются.

4.3. Критерии оценивания промежуточной аттестации по видам оценочных средств.

Максимальный балл за ответ на вопрос 2 балла.

Отлично/ зачтено/ 2 балла	Хорошо/ зачтено/ 1 балл	Неудовлетворительно/ незачтено/ 0- баллов
Высокий уровень освоения проверяемых компетенций	Средний уровень освоения проверяемых компетенций	Недостаточный уровень освоения проверяемых компетенций
студент дал полный верный ответ на вопрос	студент дал не полный ответ на вопрос;	студент не ответил на вопрос или ответ был не верный.

МИНОБРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Фонд оценочных средств для промежуточной аттестации по дисциплине «Интеллектуальный анализ текстов» по направлению подготовки (специальности) 01.04.02 «Прикладная математика и информатика» направленности (профилю) Математическое моделирование и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 10

4.4. Результаты промежуточной аттестации и уровни сформированности компетенций

При подведении итогов учитываются результаты текущей аттестации. Полученные за текущую аттестацию баллы суммируются с баллами, полученными за каждый этап при прохождении промежуточной аттестации:

Особенности проведения процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья обозначены в рабочей программе дисциплины (модуля).

Уровни сформированности компетенций определяется следующим образом:

1. Высокий уровень сформированности компетенций соответствует оценке отлично:
 - предполагает формирование компетенций на высоком уровне, готовность к самостоятельной профессиональной деятельности: формируются навыки систематизации данных, необходимых для решения математических задач.
 - студент способен аргументировать собственную точку зрения по дискуссионным вопросам дисциплины, решать ситуационные задачи, критически оценивать информацию, формулировать собственные выводы.
2. Средний уровень соответствует оценке хорошо:
 - предполагает формирование компетенций на более высоком уровне: формируется комплексное знание особенностей применения и понимания математических основ представления текстовых данных, методов обработки текстов, методов классификации и кластеризации текстов, реализации алгоритмов обработки и анализа текстов с помощью различных библиотек, методы обработки текстов с помощью глубоких нейронных сетей.
 - студент способен давать развернутые ответы на теоретические вопросы дисциплины на уровне не ниже оценки «удовлетворительно».
3. Базовый уровень соответствует оценке удовлетворительно:
 - предполагает формирование компетенций на начальном уровне: знание основных математических основ представления текстовых данных.
 - студент способен отвечать на вопросы в форме закрытого теста. Количество правильных ответов – не менее 50%.
4. Низкий уровень соответствует оценке неудовлетворительно.

