

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Баламирзоев Назим Лиодинович  
Должность: Ректор  
Дата подписания: 25.03.2026 16:00:30  
Уникальный программный ключ:  
5cf0d6f89e80f49a334f6a4ba58e91f3326b9926



**МИНОБРНАУКИ РОССИИ**  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
**Институт кибербезопасности и цифровых технологий**  
**Региональный партнёр**  
**ФГБОУ ВО**  
**«Дагестанский государственный технический университет»**



**1. .09**

Направление подготовки 09.03.01 «Информатика и вычислительная техника»

Направленность (профиль подготовки): «Прикладной искусственный интеллект»

Квалификация выпускника бакалавр

Форма обучения очная

Махачкала 2023

**ПАСПОРТ**  
**фонда оценочных средств**  
**по дисциплине МЕТОДЫ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА**

1. Результаты обучения по дисциплине:

Код	Наименование компетенции	Индикатор достижения компетенции (закрепленный за дисциплиной)	В результате освоения дисциплины обучающийся должен:	Другая дисциплина (дисциплины)/практика, участвующая в формировании компетенции
ПК-2	Способен использовать системы искусственного интеллекта в решении задач анализа, прогнозирования, планирования, синтеза и принятия решений	ПК-2.2 (ПК-3.2 РЭУ) Решает задачи с использованием систем искусственного интеллекта	Уметь решать задачи с использованием систем искусственного интеллекта	Основы интеллектуальных систем Основы обучаемых алгоритмов Методы решения оптимизационных задач Нейронные сети в решении практических задач Основы компьютерного зрения Интеллектуальные методы обработки сигналов Производственная (проектно-технологическая) практика
ПК-4	Способен использовать одну или несколько сквозных цифровых субтехнологий искусственного интеллекта	ПК-4.2 (ПК-8.2 РЭУ) Решает прикладные задачи и участвует в реализации проектов в области сквозной цифровой субтехнологии «Обработка естественного языка»	Умеет решать прикладные задачи и участвовать в реализации проектов в области сквозной цифровой субтехнологии «Обработка естественного языка»	Основы обучаемых алгоритмов Основы компьютерного зрения Производственная (технологическая) практика

ПК-6	Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-6.2 (ПК-6.2 РЭУ) Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	Умеет подготавливать и размечать структурированные и неструктурированные данные для машинного обучения	Основы обучаемых алгоритмов Основы компьютерного зрения Обработка больших данных Производственная (технологическая) практика
------	---	--	--	---

## Программа оценивания контролируемой компетенции:

№ п/п	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции/ индикатора	Наименование оценочного средства
1	Тема 1.1. Задачи обработки естественного языка.	ПК-2,4,6	Защита лабораторных работ, Защита курсовой работы Вопросы экзамена
2	Тема 1.2. Работа с текстовыми и другими файлами. Импорт из различных форматов и источников.	ПК-2,4,6	Защита лабораторных работ, Защита курсовой работы Вопросы экзамена
3	Тема 1.3. Предварительная обработка текста. Извлечение и очистка текста.	ПК-2,4,6	Защита лабораторных работ, Защита курсовой работы Вопросы экзамена
4	Тема 1.4. Регулярные выражения. Стоп-слова.	ПК-2,4,6	Защита лабораторных работ, Защита курсовой работы Вопросы экзамена
5	Тема 1.5. Токенизация. Лемматизация и стемминг текста.	ПК-2,4,6	Вопросы экзамена
6	Тема 1.6. Векторные модели текста. Мешок слов. Частотный анализ текста, N-граммы.	ПК-2,4,6	Вопросы экзамена
7	Тема 1.7. Кластеризация текстов. Тематическое моделирование.	ПК-2,4,6	Вопросы экзамена
8	Тема 2.1. Фреймворки, реализующие методы обработки естественного языка.	ПК-2,4,6	Защита лабораторных работ, Вопросы экзамена
9	Тема 2.2. Принципы аккуратных данных для обработки естественного языка.	ПК-2,4,6	Защита лабораторных работ, Вопросы экзамена
10	Тема 2.3. Лексические базы данных.	ПК-2,4,6	Вопросы экзамена
11	Тема 2.4. Корпуса текстов.	ПК-2,4,6	Вопросы экзамена
12	Тема 2.5. Графовые методы анализа текстов.	ПК-2,4,6	Вопросы экзамена
13	Тема 2.6. Масштабирование анализа текста.	ПК-2,4,6	Вопросы экзамена
14	Тема 2.7. Нейросетевые методы работы с текстами.	ПК-2,4,6	Вопросы экзамена
15	Тема 2.8. Языковые модели и генерация текста.	ПК-2,4,6	Вопросы экзамена
Форма промежуточной аттестации в 7 семестре			экзамен, курс. работа

## Вопросы для экзамена

по дисциплине: Б1.В.09 Методы обработки естественного языка

1. Архитектура LSTM для задач обработки естественного языка.
2. Векторные модели текста.
3. Генерация текста.
4. Лексические базы данных.
5. Лемматизация и стемминг текста.
6. Методы компьютерной лингвистики.
7. Методы поиска информации по коллекциям документов, фильтрация и визуализация на уровне корпуса, документа и предложения.
8. Методы разметки при работе с корпусом языка.
9. Моделирование графа, извлечение графа, структурный анализ, разрешение сущностей.
10. Нейросетевые векторные представления текста.
11. Нейросетевые методы работы с текстами.
12. Организация параллельного выполнения задач по обработке естественного языка.
13. Основные задачи обработки текстов?
14. Основные понятия обработки естественного языка.
15. Основные принципы компьютерной лингвистики.
16. Особенности извлечения информации для задач обработки естественного языка.
17. Отличия классификации и кластеризации текстов
18. Параметры регулярных выражений.
19. Перечислить методы классификации и кластеризации текстовой информации. Сформулировать основные принципы.
20. Перечислить направления компьютерной лингвистики.
21. Подходы к исследованию семантики текстов.
22. Предварительная обработка текста.
23. Привести примеры искусственных языков и нотаций.
24. Применение частотных методов в компьютерной лингвистике. Перечислить, описать, привести примеры.
25. Принципы аккуратных данных для обработки естественного языка.
26. Разъяснить принципы работы автоматических систем извлечения информации.
27. Создание тезауруса на основе графа.
28. Общие принципы построения автоматизированных систем обработки текстов.
29. Сформулировать принципы обработки неструктурированной и плохо структурированной информации. Индексация текстов.
30. Тематическое моделирование текста.
31. Токенизация.

32. Форматы текстовых данных. Достоинства и недостатки различных форматов и источников текстовых данных.
33. Фреймворки, реализующие методы обработки естественного языка.
34. Частотный анализ текста.
35. Что понимается под сентимент-анализом текста?
36. Что такое корпусная лингвистика?
37. Языковое моделирование для задач обработки и анализа естественного языка.
38. Языковые модели на базе архитектуры Трансформер.

**Описание показателей и критериев оценивания с указанием шкалы оценивания для очной и других форм обучения (с применением балльно-рейтинговой системы):**

Каждый экзаменационный билет состоит из двух вопросов.

Максимальная оценка за каждый вопрос – 20 баллов; минимальная – 12 баллов; максимальная сумма баллов – 40. Минимальный балл, свидетельствующий об успешной сдаче экзамена – 24.

**Критерии оценивания:**

**20 баллов** – полный и правильный ответ, содержащий развернутую аргументацию и примеры применения сущностей, обозначенных в вопросах;

**17-19 баллов** – неполный, но правильный ответ без ошибок и неточностей с примерами применения сущностей, обозначенных в вопросах;

**14-16 баллов** – неполный, но правильный ответ, содержащий неточности;

**12-13 баллов** – неполный и неточный ответ без достаточной аргументации, либо правильный ответ с достаточной аргументацией, но без примеров применения сущностей, обозначенных в вопросах;

**8-11 баллов** – неполный и неточный ответ, свидетельствующий лишь об общем представлении о сути вопроса;

**0-7 баллов** – неверный ответ, либо наличие хотя бы 1 грубой ошибки, свидетельствующей о непонимании сути вопроса.

Текущий контроль знаний студентов:

Максимальная оценка - 60 баллов, минимальная - 36 баллов.

Оценка по дисциплине складывается из текущего рейтинга и экзаменационного рейтинга

Конечная оценка по дисциплине оценивается по 4-балльной шкале по следующему правилу:

Число баллов	Оценка по 4-балльной шкале
87 - 100	Отлично
73 - 86	Хорошо
60 - 72	Удовлетворительно
0-59	Неудовлетворительно

## **Вопросы для защиты лабораторных работ и контроль выполнения лабораторных работ**

по дисциплине: Б1.В.09 Методы обработки естественного языка

### **Лабораторная работа 1. Задачи обработки естественного языка.**

#### **Вопросы для защиты лабораторной работы:**

1. Охарактеризуйте основные задачи обработки текстов?
2. Опишите основные принципы компьютерной лингвистики.
3. Назовите основные способы обработки и анализа естественного языка.
4. Перечислите методы компьютерной лингвистики.
5. Дайте характеристику основных задач по работе с текстовыми и другими файлами.
6. Определите выражение на конкретном языке программирования для импорта данных.
7. Охарактеризуйте форматы текстовых данных.
8. Укажите достоинства и недостатки различных форматов и источников текстовых данных.

### **Лабораторная работа 2. Предварительная обработка текста.**

#### **Вопросы для защиты лабораторной работы:**

1. Перечислите основные составляющие предварительной обработки текста.
2. Назовите виды извлекаемой информации.
3. Опишите особенности извлечения информации для задач обработки естественного языка.
4. Охарактеризуйте подходы к извлечению информации.
5. Опишите способы оценки качества извлечения.
6. Перечислите основные особенности и классы символов.
7. Дайте характеристику параметров регулярных выражений.
8. Опишите способы использования привязок, конструкций чередования, подстановок.
9. Дайте характеристику задачи при использовании стоп-слов.

### **Лабораторная работа 3. Лемматизация и стемминг.**

#### **Вопросы для защиты лабораторной работы:**

1. Перечислите особенности токенизации.
2. Дайте характеристику задачам лемматизации и стемминга текста.
3. Опишите инструменты лемматизации и стемминга.
4. Дайте характеристику задаче сегментации.

### **Лабораторная работа 4. Векторные модели текста.**

#### **Вопросы для защиты лабораторной работы:**

1. Перечислите особенности векторных моделей текста.
2. Дайте характеристику задаче частотного анализа текста.
3. Опишите варианты интерпретации векторных представлений текстов.

4. Дайте характеристику нейросетевым векторным представлениям.

#### **Лабораторная работа 5. Тематическое моделирование.**

##### **Вопросы для защиты лабораторной работы:**

1. Перечислите особенности задачи кластеризация для задач обработки и анализа естественного языка.
2. Дайте характеристику задаче тематического моделирования.
3. Назовите преимущества и недостатки вероятностного тематического моделирования.
4. Дайте характеристику тематическим векторным представлениям слов.

#### **Лабораторная работа 6. Фреймворки, реализующие методы обработки естественного языка.**

##### **Вопросы для защиты лабораторной работы:**

1. Перечислите особенности фреймворков, реализующих методы обработки естественного языка.
2. Дайте характеристику задач обработки естественного языка.
3. Опишите варианты использования фреймворков для обработки текстов.

#### **Лабораторная работа 7. Принципы аккуратных данных для обработки естественного языка.**

##### **Вопросы для защиты лабораторной работы:**

1. Перечислите основные принципы аккуратных данных.
2. Назовите атрибуты структуры аккуратных данных.
3. Опишите особенности работы с аккуратными данными для задач обработки естественного языка.
4. Охарактеризуйте способ связи структуры набора данных с его значением.

#### **Лабораторная работа 8. Лексические базы данных.**

##### **Вопросы для защиты лабораторной работы:**

1. Перечислите основные виды лексических баз данных.
2. Дайте характеристику основным задачам работы с лексическими базами данных.
3. Опишите особенности инструментов работы с лексическими базами данных.

#### **Лабораторная работа 9. Корпуса текстов.**

##### **Вопросы для защиты лабораторной работы:**

1. Перечислите особенности корпуса текстов.
2. Дайте характеристику задачам работы с корпусами текстов.
3. Опишите инструменты работы с корпусом языка.
4. Дайте характеристику задачам разметки, поиска информации по коллекциям документов.

**Лабораторная работа 10. Графовые методы анализа текстов.****Вопросы для защиты лабораторной работы:**

1. Перечислите особенности графовых методов анализа текстов.
2. Дайте характеристику задачам вычисления и анализа графов.
3. Опишите инструменты создания тезауруса на основе графа.
4. Дайте характеристику терминам моделирование графа, извлечение графа, структурный анализ, разрешение сущностей.

**Лабораторная работа 11. Масштабирование анализа текста.****Вопросы для защиты лабораторной работы:**

1. Перечислите основные составляющие задачи масштабирования анализа текста.
2. Опишите особенности организации параллельного выполнения задач по обработке естественного языка.
3. Дайте характеристику пулам процессов и очередей.
4. Опишите способы кластерных вычислений для обработки естественного языка.

**Лабораторная работа 12. Нейросетевые методы работы с текстами.****Вопросы для защиты лабораторной работы:**

1. Перечислите особенности применения нейросетевых методов работы с текстами.
2. Опишите области применения нейронных сетей для работы с текстами.
3. Назовите преимущества и недостатки нейронных сетей для задач обработки и анализа естественного языка.
4. Дайте характеристику архитектуре LSTM.

**Лабораторная работа 13. Языковые модели и генерация текста.****Вопросы для защиты лабораторной работы:**

1. Перечислите особенности применения языковых моделей для работы с текстами.
2. Опишите задачи генерации текста.
3. Назовите преимущества и недостатки языкового моделирования для задач обработки и анализа естественного языка.
4. Дайте характеристику языковым моделям на базе архитектуры Трансформер.

**Описание показателей и критериев оценивания с указанием шкалы оценивания для очной и других форм обучения (с применением балльно-рейтинговой системы)**

Оцениваются следующие показатели: знание теоретических основ лабораторной работы, умение применить их на практике, обосновать используемое решение, выполнение в установленные сроки. В рамках защиты по каждой лабораторной работе задается несколько вопросов.

*4-5 баллов выставляется, если* студент выполнил программную реализацию работы, правильно и полно отвечает на вопросы по каждой лабораторной работе, объясняет их на примерах, связывает с программной реализацией.

*2-3 бала выставляется, если* студент выполнил программную реализацию работы, отвечает на вопросы недостаточно полно или с неточностями, или не отвечает на часть заданных вопросов, не может объяснить их на примере, есть недочеты в лабораторной работе.

*0-3 балла выставляется, если* студент выполнил программную реализацию работы, не отвечает на вопросы, не может объяснить их на примере, лабораторная работа выполнена некорректно.

**Оформление сведений о дополнениях и изменениях, внесенных в ФОС дисциплины**

---

**Сведения о дополнениях и изменениях, внесенных в ФОС дисциплины**

Учебный год	Решение кафедры (№ протокола, дата)	Внесенные в ФОС дополнения изменения	Подпись заведующей кафедрой