

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность:

Документ подписан в:

Дата подписания: 20.06.2026 11:42:44

Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ

Начальник

учебно-методического управления

Т.К. Платонова

«25» мая 2026 г.

**Рабочая программа дисциплины
Обработка естественного языка**

Направление подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) программы магистратуры

01.04.02.04 Искусственный интеллект: математические модели и прикладные решения

Для набора 2026 года

Квалификация
Магистр

КАФЕДРА Прикладная математика и технологии искусственного интеллекта

Распределение часов дисциплины по семестрам / курсам

Семестр (<Курс>.<Семестр на курсе>)	3 (2.1)		Итого	
	13			
Неделя	уп	рп	уп	рп
Вид занятий				
Лекции	34	34	34	34
Лабораторные	34	34	34	34
Итого ауд.	68	68	68	68
Контактная работа	68	68	68	68
Сам. работа	40	40	40	40
Часы на контроль	36	36	36	36
Итого	144	144	144	144

ОСНОВАНИЕ

Учебный план утвержден учёным советом Университета (протокол № 9 от 03.03.2026 г.).

Программу составил(и): д.э.н., зав. кафедрой, Чернышева Юлия Гарьевна; доцент, Дмитриева Валерия Дмитриевна

Зав. кафедрой: д.э.н., доцент Ю.Г. Чернышева

Методический совет направления: д.э.н., доцент Ю.Г. Чернышева

Директор института магистратуры: д.э.н., профессор Е.А. Иванова

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	выработка у студентов компетенций, связанных с их теоретической и практической подготовкой к использованию средств и методов анализа и математического моделирования структур естественного языка
-----	---

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-6. способен руководить проектами по созданию, внедрению и использованию одной или нескольких сквозных цифровых субтехнологий искусственного интеллекта в прикладных областях

В результате освоения дисциплины обучающийся должен:

Знать:

принципы построения систем обработки естественного языка, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой субтехнологии «Обработка естественного языка» (соотнесено с индикатором ПК-6.1)

Уметь:

применять различные методы анализа и математического моделирования структур естественного языка (соотнесено с индикатором ПК-6.2)

Владеть:

использования средств и методов анализа и математического моделирования структур естественного языка (соотнесено с индикатором ПК-6.3)

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Раздел 1. Построение векторной модели текста и классификация длинных текстов

№	Наименование темы, краткое содержание	Вид занятия / работы / форма ПА	Семестр / Курс	Количество часов	Компетенции
1.1	Тема "Естественные языки, особенности обработки естественного языка" Естественный язык, его отличие от формальных языков. Многообразие языков и их лингвистические особенности. Проблемы автоматической обработки языка: неоднозначность, контекст, изменчивость, субъективность. Обзор основных задач обработки текстов на естественном языке. Обзор основных инструментов.	Лекционные занятия	3	4	ПК-6
1.2	Тема "Естественные языки, особенности обработки естественного языка" Выполнение лабораторных заданий на языке Python. Подготовка отчетов к лабораторным заданиям с использованием средств пакета LibreOffice	Лабораторные занятия	3	4	ПК-6
1.3	Тема "Методы построения векторной модели текста" Предобработка и векторизация текста. Методы представления текста в числовой форме. Мешок слов, TF-IDF, n-граммы. Эмбединги. Алгоритмы Word2Vec. Использование предобученных моделей. Программные продукты и библиотеки, используемые для построения векторной модели текста	Лекционные занятия	3	6	ПК-6
1.4	Тема "Методы построения векторной модели текста" Выполнение лабораторных заданий на языке Python. Подготовка отчетов к лабораторным заданиям с использованием средств пакета LibreOffice	Лабораторные занятия	3	6	ПК-6
1.5	Тема "Применение сверточных нейронных сетей для решения задач обработки текстов" Сверточные нейронные сети (CNN). Свертка, пуллинг, feature maps. Подготовка текстовых данных для CNN: векторизация текста, представление текста как "изображения", пайплайн. Архитектуры CNN для обработки языка	Лекционные занятия	3	6	ПК-6
1.6	Тема "Применение сверточных нейронных сетей для решения задач обработки текстов" Выполнение лабораторных заданий на языке Python. Подготовка отчетов к лабораторным заданиям с использованием средств пакета LibreOffice	Лабораторные занятия	3	6	ПК-6
1.7	Выполнение индивидуального задания по темам раздела	Самостоятельная работа	3	20	ПК-6

Раздел 2. Распознавание структуры коротких текстов и преобразование последовательностей

№	Наименование темы, краткое содержание	Вид занятия / работы / форма ПА	Семестр / Курс	Количество часов	Компетенции
2.1	Тема "Применение рекуррентных нейронных сетей для решения задачи генерации текстов" Проблема генерации текста. Особенности рекуррентных нейронных сетей (RNN). Архитектуры RNN для генерации текстов (Vanilla RNN, LSTM, GRU)	Лекционные занятия	3	6	ПК-6
2.2	Тема "Применение рекуррентных нейронных сетей для решения задачи генерации текстов" Выполнение лабораторных заданий на языке Python. Подготовка отчетов к лабораторным заданиям с использованием средств пакета LibreOffice	Лабораторные занятия	3	6	ПК-6
2.3	Тема "Методы решения задачи выделения фрагментов текста и их соотнесения с заданными классами" Выделение фрагментов текста: основные задачи (извлечение именованных сущностей, выделение ключевых слов, классификация текстовых сегментов), особенности применения. Основные методы выделения фрагментов: правила и регулярные выражения, статистические методы, машинное обучение. Методы классификации фрагментов	Лекционные занятия	3	6	ПК-6
2.4	Тема "Методы решения задачи выделения фрагментов текста и их соотнесения с заданными классами" Выполнение лабораторных заданий на языке Python. Подготовка отчетов к лабораторным заданиям с использованием средств пакета LibreOffice	Лабораторные занятия	3	6	ПК-6
2.5	Тема "Методы преобразования последовательностей" Задачи преобразования последовательностей: машинный перевод, перефразирование текста, генерация ответов в диалоговых системах. Архитектуры для преобразования последовательностей: Encoder-Decoder, Attention Mechanism, трансформеры	Лекционные занятия	3	6	ПК-6
2.6	Тема "Методы преобразования последовательностей" Выполнение лабораторных заданий на языке Python. Подготовка отчетов к лабораторным заданиям с использованием средств пакета LibreOffice	Лабораторные занятия	3	6	ПК-6
2.7	Выполнение индивидуального задания по темам раздела	Самостоятельная работа	3	20	ПК-6
2.8	Подготовка к промежуточной аттестации	Экзамен	3	36	ПК-6

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущего контроля и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Учебные, научные и методические издания

	Авторы, составители	Заглавие	Издательство, год	Библиотека / Количество
1	Седов, В. А., Седова, Н. А.	Введение в нейронные сети: методические указания к лабораторным работам по дисциплине «нейроинформатика» для студентов специальности 09.03.02 «информационные системы и технологии»	Саратов: Ай Пи Эр Медиа, 2018	ЭБС «IPR SMART»
2	Балджи А. С., Хрипунова М. Б., Александрова И. А.	Математика на Python: учебно-методическое пособие	Москва: Прометей, 2018	ЭБС «Университетская библиотека онлайн»
3		Студент. Аспирант. Исследователь: всероссийский научный журнал: журнал	Владивосток: Эксперт-Наука, 2020	ЭБС «Университетская библиотека онлайн»
4	Гуриков С.Р.	Основы алгоритмизации и программирования на Python: Учебное пособие	Москва: ООО "Научно-издательский центр ИНФРА-М", 2025	ЭБС «Znanium»
5	Гафаров Ф.М., Гилемзянов А.Ф.	Нейронные сети в PyTorch: Учебное пособие	Казань: Казанский (Приволжский) федеральный университет, 2024	ЭБС «Znanium»

5.2. Профессиональные базы данных и информационные справочные системы

ИСС "Гарант" <https://internet.garant.ru>

Федеральная государственная служба статистики <https://rosstat.gov.ru> (свободный доступ)

База данных Центрального банка РФ http://cbr.ru/hd_base (свободный доступ)

5.3. Перечень программного обеспечения

Операционная система РЕД ОС

Свободно распространяемый офисный пакет LibreOffice

Python

5.4. Учебно-методические материалы для обучающихся с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Помещения для всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения:

- столы, стулья;
- персональный компьютер / ноутбук (переносной);
- проектор;
- экран / интерактивная доска.

Лабораторные занятия проводятся в компьютерных классах, рабочие места в которых оборудованы необходимыми лицензионными и/или свободно распространяемыми программными средствами и выходом в Интернет, и/или в специализированных лабораториях, предусмотренных образовательной программой.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа к электронной информационно-образовательной среде.

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

1.1. Показатели и критерии оценивания компетенций:

ЗУН, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
ПК-6: способен руководить проектами по созданию, внедрению и использованию одной или нескольких сквозных цифровых субтехнологий искусственного интеллекта в прикладных областях			
Знать принципы построения систем обработки естественного языка, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой субтехнологии «Обработка естественного языка»	формулирует и знает основные понятия в области технологий искусственного интеллекта	полнота и содержательность ответа умение приводить примеры	Вопросы к экзамену (1-16), лабораторные задания 1-5
Уметь применять различные методы анализа и математического моделирования структур естественного языка	отвечает на вопросы, применяет методы и знания в области технологий искусственного интеллекта для выполнения лабораторных заданий	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	Вопросы к экзамену (1-16), лабораторные задания 1-5
Владеть навыками использования средств и методов анализа и математического моделирования структур естественного языка	выполняет лабораторные задания, проводит анализ данных и их обработку с использованием методов и технологий искусственного интеллекта	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	Вопросы к экзамену (1-16), лабораторные задания 1-5

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале:

- 84-100 баллов (оценка «отлично»);
- 67-83 баллов (оценка «хорошо»);
- 50-66 баллов (оценка «удовлетворительно»);
- 0-49 баллов (оценка «неудовлетворительно»).

2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Вопросы к экзамену

1. Естественные языки, особенности обработки естественного языка.
2. Основные задачи обработки текстов на естественном языке: лингвистический анализ, извлечение признаков из текстов, прикладные задачи обработки текстов, генерация текста.
3. Метод построения векторной модели текста: мешок слов, n-граммы.
4. Метод T-IDF.
5. Модель Word2Vec.
6. Метод Skip-Gram.
7. Метод Continuous Bag of Words.
8. Метод GloVe.
9. Модель FastText
10. Дистрибутивно-семантические модели.
11. Программные продукты и библиотеки, используемые для построения векторной модели текста.
12. Примеры применения сверточных нейронных сетей для решения задач обработки текстов.
13. Моделирование языка.
14. Применение рекуррентных нейронных сетей для решения задачи генерации текстов.

15. Методы решения задачи выделения фрагментов текста и их соотнесения с заданными классам.

16. Методы преобразования последовательностей.

Экзаменационное задание включает три вопроса – два теоретических вопроса и одно задание из числа приведенных ниже лабораторных заданий.

Критерии оценивания:

- 84-100 баллов (оценка «отлично») – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой;

- 67-83 баллов (оценка «хорошо») – наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, обучающийся усвоил основную литературу, рекомендованную в рабочей программе дисциплины;

- 50-66 баллов (оценка «удовлетворительно») – наличие твердых знаний в объеме пройденного курса в соответствии с целями обучения, изложение ответов с отдельными ошибками, уверенно исправленными после дополнительных вопросов; правильные в целом действия по применению знаний на практике;

- 0-49 баллов (оценка «неудовлетворительно») – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Лабораторные задания

Все лабораторные работы выполняются с применением Python

Задание к лабораторной работе 1

Выберите текст (корпус), проведите векторизацию текста. Обучите модель Word2Vec на заданном корпусе. Оцените качество работы нейронной сети. Оформите отчет.

Задание к лабораторной работе 2

Выберите текст и подготовьте данные. Подберите соответствующую архитектуру сверточной сети и реализуйте ее. Обучите модель. Оцените качество работы нейронной сети. Оформите отчет.

Задание к лабораторной работе 3

Выберите текст (корпус), подготовьте данные. Постройте модель: Embedding → LSTM → Dense с softmax. Реализуйте генерацию текста. Оцените качество работы нейронной сети. Оформите отчет.

Задание к лабораторной работе 4

Используйте готовую модель spaCy для извлечения сущностей. Продемонстрируйте примеры извлечения сущностей. Оцените качество работы нейронной сети. Оформите отчет.

Задание к лабораторной работе 5

Подготовьте данные (англо-русские пары предложений). Реализуйте модель Encoder-Decoder с LSTM и механизмом внимания. Оцените качество работы нейронной сети. Оформите отчет.

Критерии оценивания:

Всего за выполнение заданий к лабораторным работам обучающийся может получить **100 баллов**. Каждое задание к лабораторной работе оценивается максимум в **20 баллов**

16-20 баллов	Обучающийся выполнил лабораторную работу в полном объеме, самостоятельно, с соблюдением необходимой последовательности; грамотно оформил представленный от-
---------------------	---

	чет; дана содержательная интерпретация полученных при решении задач результатов; материал изложен четко
11-15 баллов	Обучающийся выполнил лабораторную работу в полном объеме, самостоятельно, с соблюдением необходимой последовательности; грамотно оформил представленный отчет; дана содержательная интерпретация полученных при решении задач результатов; материал изложен четко; но при этом допускаются отдельные логические и стилистические погрешности, уверенно исправленные обучающимся после дополнительных вопросов
6-10 баллов	Обучающийся выполнил работу в полном объеме с соблюдением необходимой последовательности; грамотно оформил представленный отчет; дана содержательная интерпретация полученных результатов; допускаются отдельные логические и стилистические погрешности; обучающийся может испытывать некоторые затруднения в формулировке суждений
0-5 баллов	Лабораторная работа не выполнена или выполнена не в полном объеме, чрезмерно фрагментарно; обучающийся практически не владеет теоретическим материалом, допускает грубые ошибки, испытывает затруднения в формулировке собственных суждений, не способен ответить на дополнительные вопросы.

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

Текущий контроль успеваемости проводится с использованием оценочных средств, представленных в п. 2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

Промежуточная аттестация проводится в форме экзамена.

Экзамен проводится по расписанию промежуточной аттестации в письменном виде. Количество вопросов в экзаменационном задании – 3 (два теоретических вопроса и одно задание из числа приведенных ниже лабораторных заданий). Объявление результатов производится в день экзамена. Результаты аттестации заносятся в ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Учебным планом предусмотрены следующие виды занятий:

- лекции;
- лабораторные занятия.

В ходе лекционных занятий рассматриваются основные теоретические вопросы, даются рекомендации для самостоятельной работы и подготовке к практическим занятиям.

В ходе лабораторных занятий углубляются и закрепляются знания студентов по ряду рассмотренных на лекциях вопросов, развиваются навыки практической работы.

При подготовке к лабораторным занятиям каждый студент должен:

- изучить рекомендованную учебную литературу;
- изучить конспекты лекций;
- подготовить ответы на все вопросы по изучаемой теме.

В процессе подготовки к лабораторным занятиям студенты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях и лабораторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом опроса и выполнения лабораторных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, дополнить конспекты лекций недостающим материалом, выписками из рекомендованных первоисточников, выделить непонятные термины, найти их значение в энциклопедических словарях.

Студент должен готовиться к предстоящему лабораторному занятию по всем обозначенным в рабочей программе дисциплины вопросам.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации студенты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.