ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей кафедрой

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
ЮУрГУ
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Замышляева А. А.
Пользователь: zamyshliaevaaa
Дата подписания: 16.10.2025

А. А. Замышляева

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.П0.05 Интеллектуальная обработка естественного языка
для направления 09.03.03 Прикладная информатика
уровень Бакалавриат
профиль подготовки Искусственный интеллект, глубокое обучение и анализ данных
форма обучения очная
кафедра-разработчик Центр ОП топ-уровня в сфере ИИ "ВиртУм"

Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.03.03 Прикладная информатика, утверждённым приказом Минобрнауки от 19.09.2017 № 922

Зав.кафедрой разработчика, д.физ.-мат.н., проф.

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
ЮУрГУ
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Замышляева А. А.
Пользователь: zamyshliaevaaa
Дата подписания: 06.10.2025

А. А. Замышляева

Разработчик программы, к.физ.-мат.н., доцент

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
ЮУрГУ
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Иванов С. А.
Пользователь: saivanov
Дата подписания: 30.09.2025

С. А. Иванов

Челябинск

1. Цели и задачи дисциплины

Целью дисциплины является формирование у студентов теоретических знаний и практических навыков в области интеллектуальной обработки естественного языка, включая базовые и современные методы предобработки, анализа, классификации, генерации и понимания текста. . Основные задачи дисциплины: ознакомить студента с основными определениями интеллектуальной обработки текстов на естественном языке, дать понимание базовых и современных подходов к решению задач обработки естественного языка, получить практический опыт в применении и оценке качества применения различных технологий для решения задач обработки естественного языка.

Краткое содержание дисциплины

Дисциплина посвящена изучению современных методов обработки естественного языка от базовых подходов до современных больших языковых моделях, их адаптацией под конкретные задачи и построением интеллектуальных систем. В курс входят следующие разделы: основы NLP и предобработка текста, нейросетевые модели и архитектура Transformer, большие языковые модели и промпт инжиниринг, fine-tuning и кастомизация больших языковых моделей, LLM-агенты. В рамках освоения дисциплины слушатели познакомятся с базовыми подходами к векторизации текста, архитектурами нейросетевых моделей, механизмом внимания и топологией трансформер, промпт-инжинирингом для работы с LLM, технологиями RAG (Retrieval-Augmented Generation), включая работу с векторными базами данных, подходами к дообучению и оптимизации больших языковых моделей, проектированием ИИ-агентов.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-14 [DL-4] Способен применять и (или) разрабатывать алгоритмы, методы и технологии обработки естественного языка	Знает: - [И-1, БУ] классические инструменты парсинга текстов: регулярные выражения, токенизация, морфологический анализ, синтаксичекий анализ [И-2, ПУ] основные архитектуры сетей, использующиеся для векторизации текстовых данных: Word2Vec, Doc2Vec, Glove, FastText, рекуррентные нейронные сети и сети-трансформеры (энкодеры) Умеет: - [И-2, ПУ] самостоятельно найти подходящую модель для векторизации текстовых данных в открытых источниках и применить её для конкретной задачи [И-3, ПУ] адаптировать и дорабатывать существующие архитектуры (например, fine-tuning BERT, GPT, T5) под конкретные задачи (классификация, генерация, NER); оптимизировать пайплайны обработки данных и обучения (ускорение через ONNX, Quantization, распределенные вычисления); строить CI/CD-процессы для NLP-моделей

	(тестирование, мониторинг дрейф данных) Имеет практический опыт: - [И-3, ПУ] разворачивания сервисов в продакшн-среде (Docker, Kubernetes, облачные NLP-API)
ПК-16 [LLM-2] Способен дообучать, адаптировать и оптимизировать генеративные модели под специфические задачи и условия применения	Знает: - [И-2, БУ] базовые адаптивные методы дообучения (prefix, adapter) [И-7, ПУ] способы настройки пайплайнов с кастомными компонентами Умеет: - [И-1, ПУ] адаптировать и валидировать датасеты под задачи обработки естественного языка [И-5, ПУ] анализировать прирост метрик моделей в задачах обработки естественного языка в зависимости от этапов обучения [И-7, ПУ] оптимизировать векторные базы данных; настраивать механизмы RAG; применять техники ускорения и повышения точности (reranking, rephrasing) Имеет практический опыт: - [И-4, ПУ] подбора параметров моделей под задачи обработки естественного языка с помощью grid и random search [И-6, ПУ] применения fine-tune к предобученным моделям на новых датасетах [И-7, ПУ] настройки retriever и reader под разные типы запросов
ПК-17 [LLM-4] Проектирует, разрабатывает и интегрирует интеллектуальных агентов на базе генеративных моделей	Умеет: - [И-2, ПУ] реализовать рассуждение на основе цепочек (ReAct, Plan&Solve)
ПК-18 [LLM-5] Организует взаимодействие с генеративными моделями через проектирование, анализ и применение промптов	Знает: - [И-2, БУ] способы создания простых последовательностей промптов Умеет: - [И-3, ПУ] управлять параметрами генерации для контроля результата настройки API при работе с LLM Имеет практический опыт: - [И-5, ПУ] настройки system prompts и ввода ограничений

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Нет	Разработка цифровых сервисов с использованием LLM, Экспертные системы, Генеративные нейронные сети

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 5 з.е., 180 ч., 92,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		6
Общая трудоёмкость дисциплины	180	180
Аудиторные занятия:	80	80
Лекции (Л)	32	32
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	0	0
Лабораторные работы (ЛР)	48	48
Самостоятельная работа (СРС)	87,5	87,5
Подготовка к тестам	10	10
Дообучить модель на бенчмарке RuCoLA	37,5	37.5
Подготовка к экзамену	20	20
Подготовка к практическим занятиям	20	20
Консультации и промежуточная аттестация	12,5	12,5
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Основы NLP и предобработка текста	12	6	0	6
2	Нейросетевые модели и архитектура Transformer	14	8	0	6
3	Большие языковые модели (LLM) и промпт инжиниринг	30	10	0	20
4	Fine-tuning и кастомизация LLM	12	4	0	8
5	LLM-агенты, протоколы MCP и A2A	12	4	0	8

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Введение в NLP: задачи, история, современные тренды	2
2	1	Токенизация, стемминг, лемматизация, нормализация текста	2
3	1	Векторизация текста: Bag-of-Words, TF-IDF, Word2Vec, GloVe, FastText	2
4	2	Введение в нейросети для NLP: RNN, LSTM, GRU	2
5	2	Трансформеры и механизм внимания.	2
6	2	BERT и его вариации: RoBERTa, DistilBERT, EuroBERT	2
7	2	Разбор архитектур: Encoder (BERT-like) vs. Decoder (GPT, Llama, Phi) vs Encoder-Decoder (T5, BART).	2
8	3	Что такое LLM? От GPT-2 до современных открытых моделей. Основные методики обучения LLM: SFT, RLHF, DPO, ORPO, PPO, RLAIF	2
9	3	Основы промпт инжиниринга: zero-shot, few-shot, chain-of-thought. Token limits, temperature, top_p, context.	2

10	3	Оптимизация запросов к GigaChat, DeepSeek, Qwen.	2
11	3	Концепция RAG. Продвинутые техники RAG.	2
12	3	Оценка качества LLM: перплексия, BLEU, ROUGE	2
13	4	Fine-tuning LLM: LoRA, QLoRA, PEFT. Квантизация.	2
14	4	Изучение кейсов партнёров (ЧКПЗ): 1. Ответы на вопросы по документации (NLP). 2. Краткое изложение и тестирование загруженных документов (NLP). 3. Работа с базой данных (Database & NLP)	2
15	5	Что такое ИИ-агенты? От простого бота к автономным системам. Протоколы MCP и A2A	2
16	5	Проектирование агентов: цели, действия, память, планирование. API, базы данных, внешние инструменты. Reasoning loop, tool use, planner-executor	2

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1	1	Подготовка данных, токенизация, балансировка классов.	2
2	1	Построение fasttext-модели.	2
3	1	Классификация текстов (Naive Bayes, Logistic Regression). Определение тональности текстов.	2
4	2	Классификация текста с помощью LSTM. Слой Embedding.	2
5	2	Fine-tuning BERT для классификации. Настройка параметров (batch size, epochs, learning rate, warmup)	2
6	2	Fine-tuning T5 для генерации саммари.	2
7	3	Генерация текста с помощью LLM (GigaChat, DeepSeek, Qwen).	2
8	3	Создание эффективных промтов для выполнения различных задач zero-shot, few-shot, CoT, prompt chaining, flow, conditionals	2
9	3	Разметка, декомпозиция задач, промптинг-грамматика.	2
10	3	Векторные БД (FAISS, Chroma)	2
11-12	3	Подготовка данных для векторизации. Инструменты и подходы к разбиению на чанки, использованию мета информации.	4
13-14	3	Поиск и ранжирование релевантных документов. Реранкинг. DPR, BM25, hybrid search	4
15-16	3	Построение RAG-системы с векторной БД. LangChain, LlamaIndex	4
17-18	4	Fine-tuning Gemma 3 на собственном датасете.	4
19-20	4	Обучение LoRA адаптера.	4
21-22	5	Создание агента для автоматического анализа отзывов.	4
23-24	5	Автоматизация поиска информации с помощью агента.	4

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к тестам	Основная литература 1-2, материалы	6	10

	презентаций		
Дообучить модель на бенчмарке RuCoLA	Методические указания по СРС	6	37,5
Подготовка к экзамену	Все материалы основной литературы, материалы презентаций	6	20
Подготовка к практическим занятиям	Основная литература 1-2, материалы презентаций	6	20

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учи-тыва-ется в ПА
1	6	Текущий контроль	Подготовка данных, токенизация, балансировка классов.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
2	6	Текущий контроль	Построение fasttext-модели.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
3	6	Текущий контроль	Классификация текстов (Naive Bayes, Logistic Regression). Определение тональности текстов.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
4	6	Текущий контроль	Классификация текста с помощью LSTM. Слой Embedding.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
5	6	Текущий контроль	Fine-tuning BERT для классификации. Настройка параметров (batch size, epochs, learning rate, warmup)	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
6	6	Текущий контроль	Fine-tuning T5 для генерации саммари.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3	экзамен

						задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	
7	6	Текущий контроль	Генерация текста с помощью LLM (GigaChat, DeepSeek, Qwen).	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
8	6	Текущий контроль	Создание эффективных промтов для выполнения различных задач zero-shot, few-shot, CoT, prompt chaining, flow, conditionals	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
9	6	Текущий контроль	Разметка, декомпозиция задач, промптинг-грамматика.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
10	6	Текущий контроль	Векторные БД (FAISS, Chroma)	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
11	6	Текущий контроль	Подготовка данных для векторизации. Инструменты и подходы к разбиению на чанки, использованию мета информации.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
12	6	Текущий контроль	Поиск и ранжирование релевантных документов. Реранкинг. DPR, BM25, hybrid search	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
13	6	Текущий контроль	Построение RAG-системы с векторной БД. LangChain, LlamaIndex	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
14	6	Текущий контроль	Fine-tuning Gemma 3 на собственном датасете.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и	экзамен

						показывает преподавателю. Каждое верно решенное задание 1 балл.	
15	6	Текущий контроль	Обучение LoRA адаптера.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
16	6	Текущий контроль	Создание агента для автоматического анализа отзывов.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
17	6	Текущий контроль	Автоматизация поиска информации с помощью агента.	3	3	На лабораторной работе выдается базовый вариант python-кода и 3 задания к нему. Студент выполняет задания, загружает отчет на портал и показывает преподавателю. Каждое верно решенное задание 1 балл.	экзамен
18	6	Текущий контроль	Тест 1	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
19	6	Текущий контроль	Тест 2	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
20	6	Текущий контроль	Тест 3	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
21	6	Текущий контроль	Тест 4	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
22	6	Текущий контроль	Тест 5	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест	экзамен

						доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	
23	6	Текущий контроль	Тест 6	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
24	6	Текущий контроль	Тест 7	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
25	6	Текущий контроль	Тест 8	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
26	6	Текущий контроль	Тест 9	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
27	6	Текущий контроль	Тест 10	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
28	6	Текущий контроль	Тест 11	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
29	6	Текущий контроль	Тест 12	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение	экзамен

						теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	
30	6	Текущий контроль	Тест 13	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
31	6	Текущий контроль	Тест 14	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
32	6	Текущий контроль	Тест 15	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
33	6	Текущий контроль	Тест 16	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	экзамен
34	6	Проме-жуточная аттестация	Итоговое тестирование	-	25	В финальном тесте 25 вопросов. Каждый вопрос оценивается 1 баллом. Ограничение по времени на прохождение теста - 50 минут. Вопросы выбираются случайным образом из всех разделов дисциплины, по 4-5 вопросов из каждой темы.	экзамен

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	При оценивании результатов учебной деятельности обучающегося по дисциплине используется балльно-рейтинговая система оценивания результатов учебной	В соответствии с пп. 2.5, 2.6 Положения

	деятельности обучающихся (Положение о БРС утверждено приказом ректора от 24.05.2019 г. № 179, в редакции приказа ректора от 10.03.2022 г. No 25-13/09). Процедура прохождения промежуточной аттестации осуществляется согласно Положению о текущем контроле успеваемости и промежуточной аттестации (приказ ректора от 27.02.2024 № 33-13/09). Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля следующим образом: • Отлично: Величина рейтинга обучающегося по дисциплине 85…100 %. • Хорошо: Величина рейтинга обучающегося по дисциплине 75…84 %. • Удовлетворительно: Величина рейтинга обучающегося по дисциплине 60…74 %. • Неудовлетворительно: Величина рейтинга обучающегося по дисциплине 0…59 %. Если студент согласен с оценкой, полученной по результатам текущего контроля, то он может в день, предшествующий промежуточной аттестации дать свое согласие на автомат в личном кабинете. В случае явки студента на промежуточную аттестацию, давшего свое согласие на автомат в личном кабинете, студент имеет право пройти мероприятия текущего контроля по дисциплине на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Снижение оценки в этом случае запрещено. Если студент не дал согласия в личном кабинете, то он может согласиться с оценкой лично на промежуточной аттестации в день ее проведения. Если студент не согласен с оценкой, то он имеет право пройти контрольно-рейтинговые мероприятия на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Фиксация результатов учебной деятельности по дисциплине проводится в день промежуточной аттестации на основе согласия студента, данного им в личном кабинете. При отсутствии согласия в журнале дисциплины фиксация результатов происходит при личном присутствии студента. Если студент не дал согласие в личном кабинете и не явился на промежуточную аттестацию – ему выставляется «неявка». Промежуточная аттестация проводится в форме тестирования в системе edu.susu.ru, тест содержит 25 вопросов, на выполнение теста дается 50 минут. В этом случае оценка за дисциплину рассчитывается на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации.	

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ																												
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ПК-14	Знает: - [И-1, БУ] классические инструменты парсинга текстов: регулярные выражения, токенизация, морфологический анализ,	+	+	+	+	+	+												+	+	+	+	+	+	+						

	синтаксичекий анализ [И-2, ПУ] основные архитектуры сетей, использующиеся для векторизации текстовых данных: Word2Vec, Doc2Vec, Glove, FastText, рекуррентные нейронные сети и сети-трансформеры (энкодеры)																															
ПК-14	Умеет: - [И-2, ПУ] самостоятельно найти подходящую модель для векторизации текстовых данных в открытых источниках и применить её для конкретной задачи [И-3, ПУ] адаптировать и дорабатывать существующие архитектуры (например, fine-tuning BERT, GPT, T5) под конкретные задачи (классификация, генерация, NER); оптимизировать пайплайны обработки данных и обучения (ускорение через ONNX, Quantization, распределенные вычисления); строить CI/CD-процессы для NLP-моделей (тестирование, мониторинг дрейф данных)		+	+	+	+	+								+	+	+	+					+	+	+							
ПК-14	Имеет практический опыт: - [И-3, ПУ] разворачивания сервисов в продакшн-среде (Docker, Kubernetes, облачные NLP-API)												+	+	+	+	+														+	+
ПК-16	Знает: - [И-2, БУ]					+	+				+	+	+	+	+	+		+						+						+		+

	базовые адаптивные методы дообучения (prefix, adapter) [И-7, ПУ] способы настройки пайплайнов с кастомными компонентами																															
ПК-16	Умеет: - [И-1, ПУ] адаптировать и валидировать датасеты под задачи обработки естественного языка [И-5, ПУ] анализировать прирост метрик моделей в задачах обработки естественного языка в зависимости от этапов обучения [И-7, ПУ] оптимизировать векторные базы данных; настраивать механизмы RAG; применять техники ускорения и повышения точности (reranking, rephrasing)					+	+				+	+	+	+	+	+		+												+		+
ПК-16	Имеет практический опыт: - [И-4, ПУ] подбора параметров моделей под задачи обработки естественного языка с помощью grid и random search [И-6, ПУ] применения fine-tune к предобученным моделям на новых датасетах [И-7, ПУ] настройки retriever и reader под разные типы запросов					+	+				+	+	+	+	+	+		+														
ПК-17	Умеет: - [И-2, ПУ] реализовать рассуждение на основе цепочек (ReAct, Plan&Solve)																+	+														
ПК-18	Знает: - [И-2, БУ] способы создания простых							+	+	+							+	+							+	+	+	+				

	последовательностей промптов																														
ПК-18	Умеет: - [И-3, ПУ] управлять параметрами генерации для контроля результата настройки API при работе с LLM						+	+	+							+	+									+	+	+	+		
ПК-18	Имеет практический опыт: - [И-5, ПУ] настройки system prompts и ввода ограничений						+	+	+							+	+														

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

Не предусмотрена

б) дополнительная литература:

Не предусмотрена

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Дообучение моделей на RuCoLA

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Дообучение моделей на RuCoLA

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	ЭБС издательства Лань	Гольдберг Й. Нейросетевые методы в обработке естественного языка. Москва : ДМК Пресс, 2019 https://e.lanbook.com/book/131704
2	Основная литература	ЭБС издательства Лань	Йылдырым, С. Осваиваем архитектуру Transformer / С. Йылдырым, М. Асгари-Ченаглу ; перевод с английского В. С. Яценкова. — Москва : ДМК Пресс, 2022. — 320 с. — ISBN 978-5-93700-106-1. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/241148
3	Основная литература	ЭБС издательства	Ганегедара Т. Обработка естественного языка с TensorFlow. Москва : ДМК Пресс, 2020 https://e.lanbook.com/book/140584

		Лань	
4	Дополнительная литература	ЭБС издательства Лань	Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. Москва : ДМК Пресс, 2018 https://e.lanbook.com/book/107901
5	Дополнительная литература	ЭБС издательства Лань	Паттерсон Д., Гибсон А. Глубокое обучение с точки зрения практика. Москва : ДМК Пресс, 2018 https://e.lanbook.com/book/116122
6	Дополнительная литература	eLIBRARY.RU	Shliazhko O., Fenogenova A., Tikhonova M. et al. mGPT: Few-Shot Learners Go Multilingual // Transactions of the Association for Computational Linguistics. — 2024. — Vol. 12. — P. 58–79. https://www.elibrary.ru/item.asp?id=66026100
7	Дополнительная литература	eLIBRARY.RU	Vladislav Mikhailov, Tatiana Shamardina, Max Ryabinin, Alena Pestova, Ivan Smurov, and Ekaterina Artemova. 2022. RuCoLA: Russian Corpus of Linguistic Acceptability. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, pages 5207–5227, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://elibrary.ru/item.asp?id=67939381

Перечень используемого программного обеспечения:

Нет

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Не предусмотрено