Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Сибирский государственный индустриальный университет»

Кафедра прикладных информационных технологий и программирования

УТВЕРЖДАЮ
Директор института
информационных технологий и
автоматизированных систем
_____ А.В. Зимин
подпись
«____» _____ 20___ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Обучение больших лингвистических моделей

09.03.03 «Прикладная информатика»
(направленность (профиль): «Искусственный интеллект и машинное обучение»)

Квалификация выпускника
Бакалавр

Форма обучения
Очная форма

Срок обучения: 4 года

Год начала подготовки 2025

Новокузнецк
2025

СибГИУ
Подлинник электронного документа хранится в системе электронного согласования документов СибГИУ
ДОКУМЕНТ ПОДПИСАН ЭП
Сертификат: e95695aa2c5fd4a2ad858d663abfeefbdf88f7d3
Владелец: Директор ИИТиАС Зимин Алексей Валерьевич
Действителен: с 2 апреля 2026 г. по 26 июня 2027 г.
Дата и время: 23-04-2026 13:21:40 ПРОВЕРИТЬ ПОДПИСЬ

1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

– формирование у обучающихся знаний и навыков осознанно настраивать, дообучать и адаптировать большие лингвистические модели под конкретные задачи.

Задачами учебной дисциплины являются:

– изучение понятий и принципов работы генеративного искусственного интеллекта;
– формирование практических навыков и умений по запуску и настройке локальных больших лингвистических моделей;
– изучение методов тонкой настройки, дообучения и обогащения больших лингвистических моделей;
– формирование практических навыков и умений по разработке ИИ-агентов.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

– Теория вероятностей и математическая статистика;
– Программирование;
– Численные методы решения инженерных задач;
– Анализ дискретных структур;
– Python для анализа данных;
– Обработка и анализ данных.

Учебная дисциплина дополняет знания и умения, получаемые по одновременно изучаемым и последующим учебным дисциплинам:

– Решение прикладных задач с использованием искусственного интеллекта;
– Фреймворки для работы с ML;
– Теория искусственных нейронных сетей и машинное обучение.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

Наименование категории	Код и наименование	Код и наименование индикатора	Планируемые результаты обучения

(группы) ПК	ПК	достижения ПК	
	ПК-2: Способен участвовать в проведении аналитического исследования с применением технологии больших данных	ПК-2.1 Участвует в выборе методов и инструментальных средств анализа данных для проведения аналитических работ, в том числе с применением технологий больших данных	– знать: принципы работы и архитектуру генеративного искусственного интеллекта, включая современные подходы к построению и функционированию больших языковых моделей; методы файн-тюнинга (fine-tuning), дообучения, обогащения моделей частными данными. – уметь: настраивать, дообучать и адаптировать LLM под конкретные прикладные задачи, используя методы тонкой настройки (fine-tuning), дообучения и обогащения моделей частными наборами данных; применять на практике техники Retrieval-Augmented Generation (RAG).
		ПК-2.2 Участвует в проведении аналитического исследования, в том числе с применением технологий больших данных	– знать: методы извлечения данных из текстов при помощи LLM; подходы к визуализации данных и генерации изображений большими лингвистическими моделями . – уметь: запускать и настраивать локальные экземпляры больших языковых моделей; разрабатывать и внедрять ИИ-агентов, способных решать различные задачи .

4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа включает в себя занятия лекционного типа (лекции), занятия семинарского типа (семинары, практические занятия, практикумы), промежуточную аттестацию обучающихся и иные формы взаимодействия обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации ООП на иных условиях, в том числе при проведении промежуточной аттестации обучающихся. Контактная работа может проводиться с применением электронного обучения, дистанционных образовательных технологий.

Объем учебной дисциплины

Семестр / курс		***ИТОГО***	***6 семестр***
Форма промежуточной аттестации			*экзамен, зачет с оценкой по КР*
Трудоёмкость	*академ. час.*	***180***	*180*
	зачетных единиц	***5***	*5*
Лекции, *академ. час.*		***16***	*16*
в форме практической подготовки		***0***	*0*
Лабораторные работы, *академ. час.*		***0***	*0*
в форме практической подготовки		***0***	*0*
Практические занятия, *академ. час.*		***48***	*48*
в форме практической подготовки		***0***	*0*
Курсовая работа, *академ. час.*		***36***	*36*
в форме практической подготовки		***0***	*0*
Консультации, *академ. час.*		***0***	*0*
в форме практической подготовки		***0***	*0*
Самостоятельная работа, *академ. час.*		***53***	*53*
в форме практической подготовки		***0***	*0*
Контроль, *академ. час.*		***27***	*27*
в форме практической подготовки		***0***	*0*

Содержание учебной дисциплины

Раздел 1 Архитектура и практическое применение больших лингвистических моделей;

Тема 1.1 Принципы и архитектура генеративного искусственного интеллекта (Принципы и понятия работы искусственного интеллекта. Основы глубокого обучения, нейронные сети и принципы их работы. Особенности обработки естественного языка.);

Тема 1.2 RAG и обогащение моделей LLM с помощью частных наборов данных (Эмбеддинг, векторная база данных, семантический поиск. Пример использования RAG и обогащения моделей LLM с помощью частных наборов данных.);

Тема 1.3 Файн-тюнинг (дообучение) больших лингвистических моделей (Техники дообучения LLM. Фреймворки файн-тюнинга. Практическое применение методов дообучения LLM.);

Тема 1.4 Обработка и генерация изображений при помощи LLM (Распознавание изображений при помощи LLM. Внедрение LLM в

приложение для анализа изображений. Обработка изображений при помощи LLM.);

Тема 1.5 Разработка ИИ-агентов (Архитектура ИИ-агентов. Отличие ИИ-агентов от ИИ-инструментов. Классификация ИИ-агентов в сфере генеративного ИИ. Фреймворки для разработки ИИ-агентов.).

5 Перечень тем лекций

№ раздела / темы дисциплины	Темы лекций	Трудоемкость, *академ.час*	
		всего	**в форме практической подготовки**
Раздел 1.	Архитектура и практическое применение больших лингвистических моделей		
Тема 1.1.	Принципы и архитектура генеративного искусственного интеллекта	4	
Тема 1.2.	RAG и обогащение моделей LLM с помощью частных наборов данных	2	
Тема 1.3.	Файн-тюнинг (дообучение) больших лингвистических моделей	4	
Тема 1.4.	Обработка и генерация изображений при помощи LLM	2	
Тема 1.5.	Разработка ИИ-агентов	4	
Итого:		**16**	**0**

6 Перечень тем практических занятий (семинаров)

№ раздела / темы дисциплины	Темы практических занятий (семинаров)	Трудоемкость, *академ.час*	
		всего	**в форме практической подготовки**
Раздел 1.	Архитектура и практическое применение больших лингвистических моделей		
Тема 1.1.	Подготовка среды для работы с большими лингвистическими моделями	4	
Тема 1.2.	Работа с локальной большой лингвистической моделью	8	
Тема 1.2.	Улучшение ответов большой лингвистической	8	

	модели за счет интеграции с базой данных		
Тема 1.3.	Дообучение LLM при помощи различных техник	8	
Тема 1.4.	Генерация и обработка изображений при помощи больших лингвистических моделей	10	
Тема 1.5.	Разработка ИИ-агентов	10	
Итого:		**48**	**0**

7 Перечень тем лабораторных работ

№ раздела / темы дисциплины	**Темы лабораторных работ**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
	Отсутствуют		
Итого:		**0**	**0**

8 Перечень тем курсовых работ (проектов)

№ раздела / темы дисциплины	**Темы курсовых работ (проектов)**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
Раздел 1.	Разработка приложения с локальной большой лингвистической моделью для решения различных задач	36	
Итого:		**36**	**0**

9 Виды самостоятельной работы

№ раздела / темы дисциплины	**Виды самостоятельной работы**	**Трудоемкость, *академ.час***	
		всего	**в форме практической подготовки**
Раздел 1.	1. Изучение лекционного материала; 2. Оформление отчета о практической работе; 3. Подготовка к практическому занятию; 4. Прохождение тестирования.	53	
Курсовая работа	*Выполнение курсовой работы*	36	0
Контроль	*Подготовка к экзамену*	27	
Итого:		**116**	**0**

10 Матрица компетентностных задач по учебной дисциплине

Контролируемые разделы (темы) дисциплины	Контролируемые компетенции	Оценочные средства
Раздел 1.	ПК-2	1. Наблюдение за работой на практическом занятии; 2. Отчет о практической работе; 3. Оценивание этапов выполнения курсовой работы; 4. Тестовые задания
Курсовая работа.	ПК-2	1. Вопросы к защите курсовой работы
Экзамен.	ПК-2	1. Экзамен на основе результатов текущего контроля

11 Входной контроль

11.1 Цель входного контроля − определить начальный уровень подготовленности обучающихся и выстроить индивидуальную траекторию обучения. В условиях личностно-ориентированной образовательной среды результаты входного оценивания обучающегося используются как начальные значения в индивидуальном профиле академической успешности обучающегося.

11.2 Описание оценочных средств

Форма проведения входного контроля – тестирование.

Длительность тестирования – 20 минут.

Тестирование проводится в течение первых двух недель изучения учебной дисциплины в СУО СибГИУ.

Для проведения тестирования используются тестовые задания для входного контроля (для ответа на 1 вопрос, как правило, выделяется 1 минута).

Шкала оценивания результатов тестирования

За каждый правильный ответ выставляется 1 балл. Оценка выставляется в зависимости от процента правильных ответов:

Оценка формируется в соответствии с критериями таблицы:

Процент результативности (правильных ответов)	Оценка
86 до 100	«отлично»
от 67до 85	«хорошо»
от 50 до 66	«удовлетворительно»
менее 50	«неудовлетворительно»

Основываясь на результатах тестирования, формируется индивидуальная траектория обучающегося (группы обучающихся), исходя из следующих соображений:

– 67 – 100 % – обучающийся готов к изучению учебной дисциплины без дополнительного повторения изученного ранее материала;

– 50 – 66 % – обучающийся в целом готов к изучению учебной дисциплины, требуется повторение отдельных вопросов(тем) изученного ранее материала, необходимого для успешного освоения учебной дисциплины;

– 49% и менее – обучающийся имеет слабую подготовку к изучению учебной дисциплины, необходимы корректирующие занятия и консультации.

Тестовые задания для входного контроля знаний обучающихся приведены в приложении.

12 Текущий контроль

12.1 Текущий контроль знаний используется для оперативного и регулярного управления учебной деятельностью (в том числе самостоятельной) обучающихся. Данный вид контроля стимулирует у обучающихся стремление к систематической самостоятельной работе по изучению учебной дисциплины и позволяет оценить степень освоения изученного материала.

Текущий контроль успеваемости осуществляется в течение семестра, в ходе повседневной учебной работы по индивидуальной инициативе педагогического работника, в том числе посредством использования СУО СибГИУ.

Формы проведения текущего контроля – наблюдение за работой на практическом занятии; отчет о практической работе; оценивание этапов выполнения курсовой работы; тестовые задания.

12.2 Описание оценочных средств

Шкалы оценивания (методики оценки)

12.2.1 Рекомендации по оцениванию этапов выполнения курсовой работы

Курсовая работа – конечный продукт, получаемый в результате планирования и выполнения комплекса учебных и исследовательских заданий. Позволяет оценить умения обучающихся самостоятельно конструировать свои знания в процессе решения практических задач и проблем, ориентироваться в информационном пространстве, а также уровень сформированности аналитических, исследовательских умений, умений практического и творческого мышления.

Оценивание этапов выполнения курсовой работы проводится в

рамках текущей аттестации (три раза в семестр) в соответствии с графиком выполнения курсовой работы по учебной дисциплине. Результаты аттестации обучающихся по выполнению ими курсовой работы оцениваются руководителями в процентах от общего объема выполненной работы. Сроки выполнения курсовой работы устанавливаются кафедрой в графике выполнения курсовой работы. График выполнения курсовой работы выдается старостам групп в начале каждого семестра.

Критерии оценки % выполнения курсовой работы

Этап работы	% выполнения	Форма отчетности
Выбор темы курсовой работы и объекта исследования		Утвержденный перечень тем курсовых работ
Постановка цели и задач курсового проектирования	10	Утвержденное задание на курсовую работу
Развертка локальной большой лингвистической модели	30	Раздел 1 пояснительной записки к курсовому проекту.
Дообучение и тонкая настройка большой лингвистической модели в соответствии с поставленной задачей	60	Раздел 2 пояснительной записки к курсовой работе.
Результаты испытаний LLM для решения поставленных задач	100	Раздел 3 пояснительной записки к курсовой работе.

Критерии оценки курсовой работы

Критерии оценки выполнения курсовой работы	Оценка
Задание выполнено полностью: цель достигнута; основные понятия выделены; имеются в наличии схемы, графическое выделение особо значимой информации; все выводы и рекомендации обоснованы, работа тщательно вычитана, отсутствуют грамматические и стилистические ошибки, работа выполнена в полном объёме; процент выполнения задания соответствует плановому.	отлично
Задание выполнено: цель выполнения достигнута; наличие правильных эталонных ответов; однако работа выполнена не в полном объёме, имеются единичные опечатки, орфографические, пунктуационные, грамматические, стилистические ошибки; график выполнения задания нарушен на 10 – 20 %.	хорошо
Задание выполнено частично: цель выполнения достигнута не полностью; многочисленные ошибки снижают качество выполненной работы; работа представлена с отклонением от запланированного срока на 50 %.	удовлетворительно
Задание не выполнено, цель работы не достигнута.	неудовлетворительно

12.2.2 Рекомендации по оцениванию результатов тестирования обучающихся

Тестирование проводится по изученным разделам (темам) учебной дисциплины для выявления уровня знаний и умений обучающегося, мотивирования обучающегося к активизации самостоятельной работы по усвоению учебного материала, выявления и устранения пробелов в знаниях, повышения дисциплины и организации учебной деятельности обучающегося.

Тестирование проводится в СУО СибГИУ в рамках часов, отведенных на самостоятельную работу обучающихся. Для тестирования используется 20 вопросов из базы типовых тестовых заданий по изученным разделам (темам) учебной дисциплины. Обучающийся выбирает один ответ из предложенных вариантов ответа. В случае выбора неправильного ответа обучающийся получает за данный вопрос 0 баллов; в случае правильного ответа – 1 балл.

Оценка формируется в соответствии с критериями таблицы:

Максимальный балл	Проходной балл	Оценка
20	не менее 18	отлично
17	не менее 14	хорошо
13	не менее 10	удовлетворительно
9	–	неудовлетворительно

12.2.3 Рекомендации по оцениванию в процессе наблюдения за работой на практическом занятии

Наблюдение за работой на практическом занятии – важнейший компонент контроля деятельности обучающихся и качества их обучения.

Критерии оценки в процессе наблюдения за работой на практическом занятии:

– самостоятельность выполнения задания и работы с методическими указаниями (учитывается работа в течение занятия, быстрота и способность нахождения в методических указаниях нужной информации);

– правильность выполнения работы (учитывается последовательность выполняемых действий, умение работать с инструментами и/или средствами вычислительной техники);

– рациональность использования времени, отведенного на задание (не одобряется затянутость выполнения работы с учетом индивидуальных особенностей обучающихся);

– активность обучающегося при выполнении практической работы (учитывается индивидуальная работа и работа в группе).

Критерии оценки по результатам наблюдения за работой на практическом занятии

Наименование параметра оценки	Максимальное количество баллов
Самостоятельно выполняет задания; умеет работать с методическими указаниями.	2,5
Указанные в методических указаниях действия выполняет в правильной последовательности; умеет работать с инструментами и/или средствами вычислительной техники.	1,25
Рационально использует время, отведенное на задание; проявляет активность при выполнении практической работы.	1,25
Итого	**5**

Максимальное количество – 5 баллов соответствует оценке «отлично», 4 балла – «хорошо», 3 балла – «удовлетворительно», 2 балла – «неудовлетворительно».

12.2.4 Рекомендации по оцениванию отчета о практической работе

Оформление отчета о проделанной практической работе обучающиеся осуществляют в рамках самостоятельной работы.

Критерии оценки отчета о практической работе:

– правильность выполнения работы (учитывается полнота отчета о практической работе, наличие всех необходимых данных, наличие выводов, наличие ошибок и неточностей);

– аккуратность оформления отчета о практической работе.

Критерии оценки отчета о практической работе

Наименование параметра оценки	Максимальное количество баллов
Отчет о практической работе содержит все необходимые данные и выводы.	2,5
Отчет о практической работе не содержит ошибок и неточностей.	1,25
Отчет о практической работе оформлен аккуратно.	1,25
Итого	**5**

Максимальное количество – 5 баллов соответствует оценке «отлично», 4 балла – «хорошо», 3 балла – «удовлетворительно», 2 балла – «неудовлетворительно».

Задания и вопросы для текущего контроля знаний обучающихся по разделам (темам) учебной дисциплины приведены в приложении.

13 Промежуточная аттестация

13.1 Оценочные материалы для промежуточной аттестации обучающихся по учебной дисциплине предназначены для оценки степени достижения запланированных результатов обучения по

завершению изучения учебной дисциплины в установленной учебным планом форме и позволяют определить качество усвоения изученного материала.

Итоговой формой контроля сформированности компетенций у обучающихся по учебной дисциплине является: курсовая работа; экзамен.

Оценочные материалы для промежуточной аттестации обучающихся включают: вопросы к защите курсовой работы; экзамен на основе результатов текущего контроля.

13.2 Оценивание обучающегося на защите курсовой работы

Оценка	Требования к знаниям
отлично	Оценка «отлично» выставляется обучающемуся, если выполнены все требования к выполнению и защите курсовой работы: соблюдены требования к внешнему оформлению, последовательно изложен материал; при защите обучающийся демонстрирует умения применять теоретические знания основной и дополнительной литературы, может показать и объяснить применение технологических решений, использованных в курсовой работе; даны правильные ответы на вопросы, заданные в процессе защиты курсовой работы.
хорошо	Оценка «хорошо» выставляется обучающемуся, если основные требования к выполнению и защите курсовой работы выполнены, но при этом допущены недочеты. В частности, имеются неточности в изложении материала; отсутствует логическая последовательность в решениях, имеются упущения в оформлении; на дополнительные вопросы при защите курсовой работы даны неполные ответы.
удовлетворительно	Оценка «удовлетворительно» выставляется обучающемуся, если имеются существенные отступления от требований к выполнению и защите курсовой работы. В частности: задача решена лишь частично; допущены фактические ошибки в содержании и оформлении; при ответе на дополнительные вопросы на защите курсовой работы обучающийся показывает знания только основного материала, испытывает затруднения при объяснении выбранных решений.
неудовлетворительно	Оценка «неудовлетворительно» выставляется обучающемуся, если задача решена лишь частично; допущены грубые ошибки в содержании и при оформлении или при ответе на дополнительные вопросы; во время защиты курсовой работы обучающийся обнаруживает существенное непонимание проблемы.

Если допущены существенные недостатки в оформлении курсовой работы, опущен или не написан какой-либо раздел, имеются отступления от задания на курсовую работу, то такая курсовая работа возвращается обучающемуся на доработку.

13.3 Оценивание обучающегося на экзамене

Оценка	Требования к знаниям
отлично	Оценка «отлично» выставляется обучающемуся, если он

Оценка	Требования к знаниям
	получил среднюю оценку «отлично» по 3-м результатам тестирования в период текущих аттестаций в семестре, представил полное и верное решение задачи, поставленной в курсовой работе, выполнил правильно и полно все практические задания и самостоятельную работу курса.
хорошо	Оценка «хорошо» выставляется обучающемуся, если он получил среднюю оценку «хорошо» по 3-м результатам тестирования в период текущих аттестаций в семестре, решил задачу, поставленную в курсовой работе, выполнил правильно и полно все практические задания и самостоятельную работу.
удовлетворительно	Оценка «удовлетворительно» выставляется обучающемуся, если он получил среднюю оценку «удовлетворительно» по 3-м результатам тестирования в период текущих аттестаций в семестре, выполнил все практические задания и самостоятельную работу, решил задачу, поставленную в курсовой работе,.
неудовлетворительно	Оценка «неудовлетворительно» выставляется обучающемуся, если он получил среднюю оценку «неудовлетворительно» по 3-м результатам тестирования в период текущих аттестаций в семестре или не выполнил все практические задания и самостоятельную работу, не участвовал в выполнении курсовой работы.

Задания и вопросы для промежуточной аттестации обучающихся приведены в приложении.

14 Учебно-методическое и информационное обеспечение учебной дисциплины

а) литература:

1 Мохов, В. А. Системы искусственного интеллекта: современные методы программной инженерии : учебное пособие / В. А. Мохов, А. В. Кузнецова. — Новочеркасск : ЮРГПУ (НПИ), 2021. — 150 с. — ISBN 978-5-9997-0756-7. – URL: https://e.lanbook.com/book/292217 (дата обращения: 19.04.2026);

2 Жаткина, К. Н. Системы искусственного интеллекта : учебное пособие / К. Н. Жаткина, Т. О. Махалкина. — Дубна : Государственный университет «Дубна», 2023. — 73 с. — ISBN 978-5-89847-682-3. – URL: https://e.lanbook.com/book/369356 (дата обращения: 19.04.2026);

3 Воронов, М. В. Системы искусственного интеллекта : учебник и практикум для вузов / М. В. Воронов, В. И. Пименов, И. А. Небаев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 268 с. — (Высшее образование). — ISBN 978-5-534-17032-0. – URL: https://urait.ru/bcode/544161 (дата обращения: 19.04.2026).

б) ресурсы информационно-телекоммуникационной сети «Интернет»:

1 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 –]. – URL: http://libr.sibsiu.ru. – URL: https://libr.sibsiu.ru;

2 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 –]. – URL: http://library.sibsiu.ru/LibrELibraryFullText.asp. – Режим доступа: для авторизир. пользователей. – URL: https://library.sibsiu.ru/LibrELibraryFullText.asp;

3 Консультант студента : электронно-библиотечная система / ООО «КОНСУЛЬТАНТ СТУДЕНТА». – Москва, [200 –]. – URL: http://www.studentlibrary.ru. – Режим доступа: для авторизир. пользователей;

4 ЛАНЬ : электронно-библиотечная система : [коллекция «Единая профессиональная база знаний для технических вузов - Издательство Лань" ЭБС»] / ООО «Издательство ЛАНЬ». – Санкт-Петербург, [200 –]. – URL: http://e.lanbook.com. – Режим доступа: для авторизир. пользователей;

5 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство ЮРАЙТ». – Москва, [200 –]. – URL: https://urait.ru. – Режим доступа: для авторизир. пользователей;

6 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 –]. – URL: https://biblioclub.ru. – Режим доступа: для авторизир. пользователей. – URL: http://www.biblioclub.ru;

7 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 –]. – URL: http://elibrary.ru. – Режим доступа: по подписке;

8 Электронные периодические издания ИВИС : универсальная база данных / ООО «ИВИС». – Москва, [200 –]. – URL: http://eivis.ru. – Режим доступа: по подписке.

в) лицензионное и свободно распространяемое программное обеспечение:

– 7-Zip;
– Adobe Acrobat Reader;
– Astra Linux Special Edition;
– Jupyter Notebook;
– Kaspersky Endpoint Security;
– Microsoft Office;
– Microsoft Windows;
– OnlyOffice;
– Р7-Офис.

г) базы данных и информационно-справочные системы:

1 ГАРАНТ : справочно-правовая система / ООО «Правовой центр «Гарант». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

2 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.

15 Материально-техническое обеспечение учебной дисциплины

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа, оборудованную учебной доской, экраном и мультимедийным проектором;
- учебную аудиторию для проведения занятий семинарского типа (практических занятий), оснащенную персональными компьютерами и выходом в сеть «Интернет»;
- учебную аудиторию для выполнения курсовых работ;
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ.

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Составитель(и):

доцент Бабичева Надежда Борисовна (кафедра прикладных информационных технологий и программирования).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

Согласована:

Системный аналитик группы
системной архитектуры департамента
информационных технологий
ООО "СИ СОФТ"  Е.А. Шапошникова

Приложение

Тестовые задания для входного контроля знаний обучающихся

Выберите один правильный ответ.

1 Выберите определение термина «Данные».

a. Сведения об окружающем мире, определенные только в рамках какой-либо предметной области;

b. Информация, фиксированная в определенной форме, пригодной для последующей обработки, хранения и передачи;

c. Сведения об окружающем мире, протекающих процессах и явлениях, рассматриваемые с учетом некоторой их семантической сущности;

d. Сведения об окружающем мире, определенным образом истолкованные, обеспечивающие достижение некоторой цели и структуры.

2 Система управления базой данных (СУБД) -это:

a. служебная информация, содержащая упорядоченные сведения о ключевых значениях;

b. комплекс программных и языковых средств, необходимых для создания и модификации базы данных;

c. поименованная совокупность структурированных данных, относящихся к определенной предметной области.

3 Перечислите все известные типы аномалий:

a. Модификации;

b. Перемещения;

c. Вставки;

d. Удаления.

4 Укажите требование к первой нормальной форме:

a. Каждый неключевой атрибут функционально зависит от первичного ключа;

b. Отсутствие транзитивных зависимостей неключевых атрибутов от первичного ключа;

c. Каждый детерминант является ключом-кандидатом;

d. Каждая таблица является отношением.

5 Транзитивная зависимость – это:

a. Прямая зависимость неключевых атрибутов от первичного ключа;

b. Зависимость неключевых атрибутов от первичного ключа через другой атрибут;

c. Прямая зависимость неключевых атрибутов от уникального ключа;

d. Зависимость неключевых атрибутов от уникального ключа через другой атрибут;

e. Зависимость внешнего ключа от первичного ключа.

6 Многозначная зависимость – это:

a. Зависимость, когда, нескольким значениям кортежа соответствует много значений атрибутов;

b. Зависимость, когда, одному значению кортежа соответствует много значений атрибутов;

c. Зависимость, когда, одному значению атрибута соответствует много значений кортежа;

d. Зависимость, когда, одному значению одного атрибута может соответствовать много значений другого атрибута.

7 Отличительной особенностью каких языков программирования является их ориентация не на систему команд той или иной ЭВМ, а на систему операторов, характерных для записи определенного класса алгоритмов?

a. языков программирования низкого уровня;

b. языков программирования высокого уровня;

c. языков программирования сверхвысокого уровня;

d. правильных ответов нет.

8 Что понимают под языком программирования (ЯП)?

a. язык, предназначенный для решения определенного класса задач (проблем);

b. правила представления данных и записи алгоритмов их обработки, которые автоматически выполняются ЭВМ;

c. язык, предназначенный для создания пакетов прикладных программ, в том числе для современных операционных систем;

d. правильных ответов нет.

9 Как называется именованная спецификация одного или более столбцов (для каждого столбца указывается имя, а также его тип или домен)?

a. строчный тип данных;

b. объектный тип данных;

c. комбинированный тип данных;

d. правильных ответов нет.

10 Что понимается под наследованием типов?

a. однородная масса разрядов, имеющая какую-либо структуру;

b. возможность дисциплинированного создания новых типов на основе уже определенных;

c. многовходовой программный модуль, точки входа которого соответствуют набору операций реализуемого типа;

d. правильных ответов нет.

11 Что из перечисленного не относится к наиболее распространенным конструируемым типам данных?

a. тип записи;

b. тип множества;

c. тип массива;

d. тип распределения;

e. правильных ответов нет.

12 Что понимается под классом в объектно-ориентированном программировании (ООП)?

a. процедуры и функции любого языка программирования;

b. структурный тип данных, который включает описание полей данных, процедур и функций, работающих с этими полями данных;

c. списки передаваемых параметров, типы функции;

d. правильных ответов нет.

13 Какие типы в языке C# относятся к значимым:

a. логический;

b. арифметический;

c. структуры;

d. перечисление;

e. массивы;

f. строки;

g. классы;

i. правильных ответов нет.

14 Определение класса в C# и других объектных языках содержит:

a. вершины;

b. методы;

c. приключения;

d. равнины;

e. поля;

f. события;

g. правильных ответов нет.

15 Какие типы в языке C# относятся к ссылочным:

a. логический;

b. арифметический;

c. структуры;

d. перечисление;

e. массивы;

f. строки;

g. классы;

i. правильных ответов нет.

16 Что называют статистической гипотезой?

a. предположение о теоретических значениях параметров распределения;

b. утверждение о виде неизвестного распределения;

c. предположение о виде неизвестного распределения или о параметрах известного распределения;

d. утверждение о параметрах известного распределения.

17 Статистические характеристики – это

a. закон распределения случайной величины;

b. числовые параметры, которые описывают случайную величину полностью;

c. числовые параметры, которые характеризуют степень разбросанности значений случайной величины;

d. числовые параметры, которые отражают существенные черты случайной величины;

e. ничего из перечисленного;

f. все перечисленное.

18 Какие из приведенных высказываний являются свойствами дисперсии?

a. дисперсия произведения независимой случайной величины на постоянную равна произведению дисперсии случайных величины на квадрат постоянной;

b. дисперсия суммы независимых случайных величин равно сумме дисперсий этих случайных величин;

c. дисперсия постоянной равна нулю;

d. дисперсия произведения случайных величин равно произведению дисперсий этих случайных величин;

e. дисперсия суммы случайных величин равно сумме дисперсий этих случайных величин.

19 Коэффициент взаимной корреляции r_x=0,9 – это означает, что

a. между двумя отсчетами одной случайно величины есть сильная линейная связь;

b. между двумя случайными величинами линейная связь отсутствует;

c. между двумя случайными величинами есть сильная линейная связь;

d. между двумя отсчетами одной случайно величины линейная связь отсутствует;

e. между двумя случайными величинами есть сильная нелинейная связь;

f. случайные величины независимы.

20 Обучающая выборка должна

a. быть репрезентативной;

b. удовлетворять условиям качества;

c. быть разбита на две части – обучающую и контрольную;

d. отражать правила и закономерности исследуемого процесса;

e. быть закодирована на машинном языке.

Задания и вопросы для текущего контроля знаний обучающихся по разделам (темам) учебной дисциплины

1 Тестирование по результатам освоения лекционного материала и самостоятельной работы

Тестирование проводится для определения степени усвоения материалов лекций, а также выполнения самостоятельной работы в рамках часов, отведенных на самостоятельную работу обучающихся, в

СУО
СибГИУ.

Перечень вопросов для тестирования:

Раздел 1. Архитектура и практическое применение больших лингвистических моделей

1. Агент RAG - Retrieval-Augmented Generation Agent в мультиагентной системе (MAS) является агентом

1) измерения физикальных параметров;

2) первичного ввода данных пациентом;

3) проведения лабораторных тестов;

4) контекстного анализа и генерации выводов.

2. Агент причинно-следственных связей (CRA - Causal Reasoning Agent) отвечает за

1) маршрутизацию данных;

2) выявление возможных диагнозов на основе агрегированных данных;

3) генерацию финального заключения;

4) контроль качества выводов.

3. В каких случаях рекомендуется выбирать SLM?

1) задача представляет собой классификацию или извлечение данных;

2) требуется генерация сложных художественных текстов;

3) требуется сертификация регулирующими органами;

4) данные являются чувствительными - предпочтительна конфиденциальная (on-premise) обработка.

4. В чем заключается основной недостаток монолитных больших языков моделей (LLM)?

1) чрезмерная простота архитектуры;

2) отсутствие специализации и слабая интерпретируемость;

3) слишком низкая стоимость;

4) неспособность работать с текстом.

5. На чем обучаются современные модели искусственного интеллекта?

1) исключительно на синтетических данных;

2) на данных из библиотек;

3) на значительном проценте всех данных из интернета;

4) только на научных статьях.

6. Какая из перечисленных моделей является примером большой языковой модели (LLM)?

1) TinyLLaMA-Med;

2) GPT-4;

3) MedAlpaca-1.3B;

4) BioBERT.

7. Какая из перечисленных моделей является примером малых языковых моделей (Small Language Model, SLM)?

1) BioBERT;
2) GPT-4;
3) Llama-3-70B;
4) Med-PaLM 2.

8. Какая из проблем относится к техническим вызовам мультиагентной системы (MAS)?

1) зависимость от качества данных;
2) ответственность;
3) требования контролирующих органов;
4) защита персональных данных.

9. Какая из проблем относится к этическим вызовам мультиагентной системы (MAS)?

1) выбор цвета интерфейса;
2) стоимость разработки;
3) прозрачность решений (возможность понять логику модели);
4) скорость работы системы.

10. Какая из стратегий совмещения большой языковой модели (LLM) и малой языковой модели (SLM) предполагает первичный анализ данных локально?

1) ансамбли из SLM;
2) использование только облачных LLM;
3) только Retrieval-Augmented Generation (RAG) для обоих типов;
4) on-device SLM + cloud LLM по запросу.

11. Что представляет собой "Окно контекста"?

1) максимальный размер файла, который пользователь может загрузить в систему искусственного интеллекта;
2) ограничение на размер обучающей выборки, которое используется для первоначального формирования знаний нейронной сети перед её выпуском;
3) период времени, в течение которого искусственный интеллект имеет доступ к интернет-поиску и может получать актуальную информацию из внешней базы данных;
4) сколько текста искусственный интеллект может "удерживать в голове" за раз - его кратковременная память; если окно = 4000 токенов, искусственный интеллект забудет начало, если ваш текст длиннее.

12. Какие из следующих утверждений о малых языковых моделях (SLM) верны?

1) проще поддаются контролю и сертификации;
2) могут работать на локальных серверах;
3) имеют, как правило, более 7B параметров;
4) требуют для работы мощные GPU и облако.

13. Каков приблизительный порог количества параметров для отнесения модели к классу большие языковые модели (Large Language Models, LLM)?

1) >7 миллиардов (7B);

2) >1 миллиарда (1B);
3) >10 миллиардов (10B);
4) <3 миллиардов (3B).
14. Каков приблизительный порог количества параметров для отнесения модели к классу малых языковых моделей (Small Language Models, SLM)?
1) <3 миллиардов (3B);
2) >10 миллиардов (10B);
3) >5-7 миллиардов (5-7B);
4) >4 миллиардов (4B).
15. Какую модель, как правило, проще контролировать?
1) малую языковую модель (SLM);
2) большую языковую модель (LLM);
3) GPT-4;
4) Llama-3-70B.
16. Ключевой критерий различия между большими языковыми моделями (LLM) и малыми языковыми моделями (SLM) - это
1) используемый фреймворк для обучения;
2) практическая применимость и масштаб;
3) только количество параметров;
4) язык, на котором обучена модель.
17. Малая языковая модель (SLM) предназначена для
1) для ограниченного использования под закрытой лицензией;
2) для исследований, прототипов, сложных диагностических ассистентов (при наличии ресурсов);
3) для обучения и дообучения сверхкрупных языковых моделей в распределённых дата-центрах;
4) для реального внедрения в ресурсо-ограниченные, регулируемые среды.
18. Модель не может быть сертифицирована, если
1) она не использует Retrieval-Augmented Generation (RAG);
2) она работает слишком медленно;
3) нельзя объяснить, как она работает;
4) она имеет мало параметров.
19. Мультиагентная система (MAS) имитирует
1) одного универсального эксперта;
2) базу данных пациентов;
3) работу одного мощного алгоритма;
4) консилиум экспертов разных специальностей.
20. Основное предназначение большой языковой модели (LLM)
1) широкое внедрение в различных сферах;
2) реальное внедрение в регулируемых средах;
3) для исследований, прототипов, сложных диагностических ассистентов;
4) работа на edge-устройствах.

21. Роль агента "Критик" - это

1) агрегация данных;

2) разбор текста;

3) проверка логической согласованности, полноты и достоверности выводов, сформированных агентами Causal Reasoning Agent (CRA) и Retrieval-Augmented Generation Agent (RAG);

4) формирование рекомендаций по запросу.

22. Функция агента агрегации данных агрегатора заключается в

1) извлечении сущностей из текста;

2) визуализации данных для врача;

3) принятии окончательного клинического решения;

4) сборе, нормализации и объединении данных от всех специализированных агентов.

23. Что из перечисленного является техническим преимуществом мультиагентной системы (MAS)?

1) соблюдение протоколов;

2) модульность;

3) повышение точности диагностики;

4) снижение времени на анализ.

24. Какая архитектура лежит в основе современных больших языковых моделей?

1) перцептроны;

2) рекуррентные сети;

3) сверточные сети;

4) трансформеры.

25. Какие из перечисленных вариантов ниже являются методами обучения нейронных сетей?

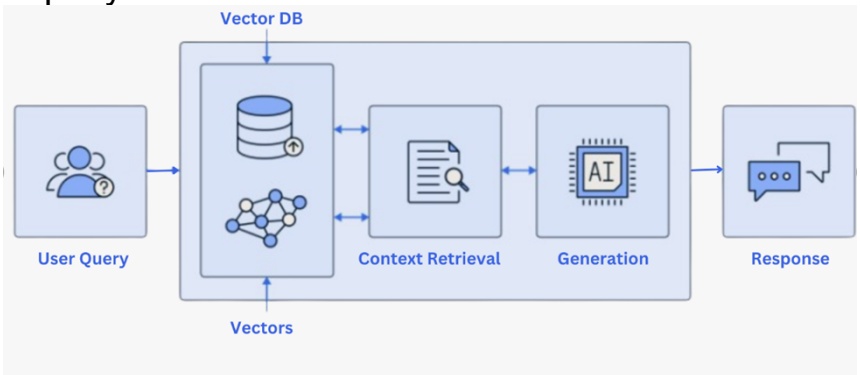
1) генетические алгоритмы;

2) деревья решений;

3) с учителем, без учителя, с подкреплением;

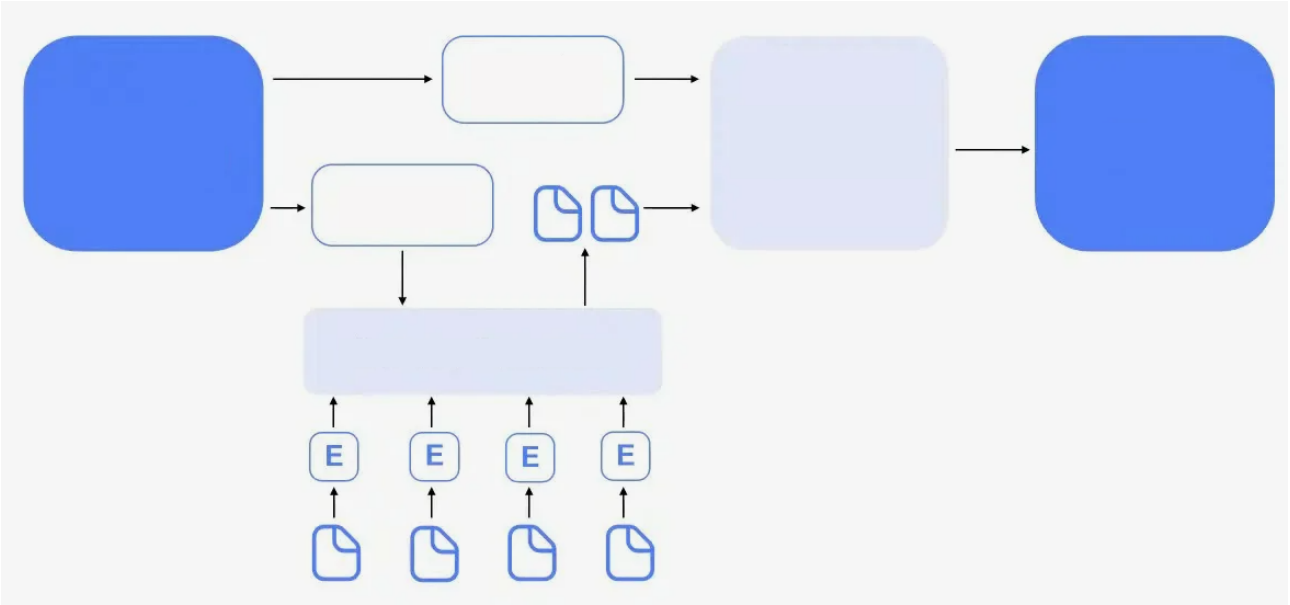
4) самоорганизующиеся карты.

26. Укажите с какой архитектурной особенностью показана схема LLM на рисунке.



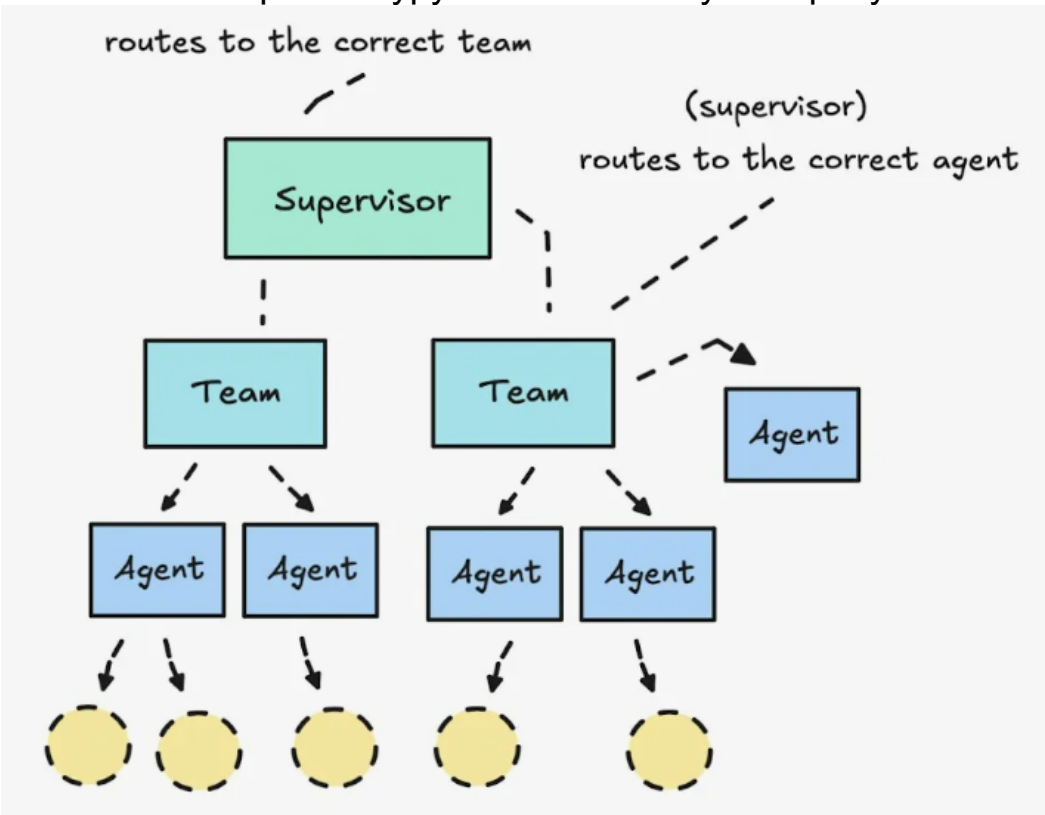
Введите ответ:_____.

27. Подпишите блоки на изображении, которое иллюстрирует схему применения RAG при обучении большой языковой модели.



Варианты ответов:
a. Ответ.
b. Запрос пользователя.
c. LLM.
d. Embedding.
e. Prompt Engineering
f. Векторная база данных.

28. Назовите архитектуру LLM показанную на рисунке.



Введите ответ:_____.

29. Как называется фреймворк тонкой настройки, который позволяет дообучать большие языковые модели, изменяя лишь малую часть их параметров?

Введите ответ:_____.

30. Назовите ключевую особенность архитектуры мультиагентной системы, которая обеспечивает эффективное взаимодействие между агентами.

Введите ответ:_____.

2 Выполнение практических работ

Задания для практических работ, теоретический и методический материал для самостоятельного изучения и вопросы для собеседования представлены в методических указаниях к практическим занятиям.

Примеры заданий на практические работы.

Тема 1. Подготовка среды для работы с большими лингвистическими моделями

Установка и настройка LLM с локальным инференсом. Установка клиента с графическим интерфейсом пользователя для работы с локальной большой лингвистической моделью. Настройка виртуальной среды Python для разработки генеративного ИИ. Обеспечение аппаратного ускорения вычислений для LLM.

Тема 2. Работа с локальной большой лингвистической моделью

Разработка простого приложения с локальной LLM. Основы промптинга для локальной LLM, формирование различных ролей для LLM. Подключение эмбеддингов.

Тема 2. Улучшение ответов большой лингвистической модели за счет интеграции с базой данных

Основы преобразования текста в SQL (Text-to-SQL) при помощи LLM. Изучение шаблонов проектирования систем преобразования текста в SQL.

Тема 3. Дообучение LLM при помощи различных техник

Применение техник файн-тюнинга предварительно обученной модели (полный файн-тюнинг, параметрически эффективный файн-тюнинг (PEFT), дистилляция знаний). Работа с фреймворками для файн-тюнинга LLM. Применение файн-тюнинга при построении LLM для решения задач исследования рынка.

Тема 4. Генерация и обработка изображений при помощи больших лингвистических моделей

Распознавание изображений при помощи локальной LLM. Внедрение LLM в приложение для анализа изображений. Обработка изображений при помощи LLM.

Тема 5. Разработка ИИ-агентов

Архитектура и принципы работы ИИ-агентов. Формулирование цели и задачи для агента. Разработка прототипа ИИ-агента, способного выполнять поставленные задачи.

3 Выполнение курсовой работы

В рамках учебной дисциплины предусмотрена самостоятельная работа обучающихся, результатом выполнения которой является оформленная в соответствии с требованиями пояснительная записка курсовой работы по заданной теме.

Примерный перечень тем курсовых работ представлен в рабочей программе дисциплины. Задание и методические указания к выполнению курсовой работы размещены в СУО СибГИУ.

Перед выдачей темы курсовой работы обучающийся знакомится с графиком его выполнения и критериями оценки. Выполненная пояснительная записка курсовой работы загружается обучающимся в СУО СибГИУ.

4 Вопросы к защите курсовой работы

1 Почему для вашего проекта было выбрано именно локальное развертывание LLM, а не использование облачных сервисов?

2 Какие аппаратные требования предъявляются для запуска выбранной вами модели?

3 Почему вы выбрали именно эту модель? Какие альтернативы рассматривались?

4 Как вы оценивали производительность модели на вашем оборудовании?

5 Какие сложности возникли при установке и настройке окружения?

6 Как вы обеспечивали безопасность и конфиденциальность данных при локальном запуске модели?

7 Какие инструменты мониторинга использовались для контроля загрузки CPU/GPU и памяти?

8 Как организован процесс обновления модели и окружения?

9 Почему был выбран именно такой способ взаимодействия с моделью?

10 Как вы проверяли корректность работы развернутой модели?

11 В чем отличие дообучения (fine-tuning) от использования промптов или RAG?

12 Почему возникла необходимость в дообучении модели для вашей задачи?

13 Как формировался и подготавливался обучающий датасет?

14 Какие критерии использовались для оценки качества обучающих примеров?

15 Как вы боролись с переобучением и сохраняли обобщающую способность модели?

16 Какие параметры тонкой настройки (гиперпараметры) вы изменяли и почему?

17 Как оценивалась эффективность дообучения? Какие метрики использовались?

18 Как изменился стиль и структура ответов модели после дообучения?

19 Какие ограничения есть у вашего подхода к дообучению?

20 Как вы документировали процесс и результаты дообучения?

21 Какие конкретные задачи были поставлены перед приложением?

22 Как вы формировали тестовые сценарии для оценки работы модели?

23 Каковы основные количественные и качественные результаты испытаний?

24 С какими типами ошибок чаще всего сталкивалась модель (галлюцинации, логические ошибки, стилистические несоответствия)?

25 Как вы анализировали и интерпретировали полученные ответы модели?

26 Как изменились результаты после дообучения по сравнению с базовой моделью?

27 Какова была скорость генерации ответов и как она зависит от сложности запроса?

28 Какие задачи оказались наиболее сложными для вашей реализации?

29 Как вы оценивали пользовательский опыт при работе с приложением?

30 Какие направления для дальнейшего развития и улучшения приложения вы видите по итогам испытаний?

5 Оценивание обучающегося на экзамене

Промежуточная аттестация производится на основе текущего контроля успеваемости при условии выполнения всех заданий в семестре, предусмотренных фондом оценочных средств по дисциплине.