

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет компьютерных технологий и прикладной математики

УТВЕРЖДАЮ:

Проректор по учебной работе,
качеству образования – первый
проректор

подпись

Хагуров Т.А.

« 29 » августа 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1. В.06 Современные методы фронтальных исследований ИИ

Направление подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем

Профиль Искусственный интеллект и аналитика данных

Форма обучения очная

Квалификация бакалавр

Краснодар 2025

Рабочая программа дисциплины «Современные методы фронтальных исследований ИИ» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем.

Программу составил(и):

В.А. Акиншина, канд. пед. наук, доцент



ПОДПИСЬ

Рабочая программа дисциплины «Современные методы фронтальных исследований ИИ» на заседании центра искусственного интеллекта протокол № 01 «28» августа 2025 г.

Руководитель центра ИИ Коваленко А.В.



ПОДПИСЬ

Утверждена на заседании учебно-методической комиссии факультета компьютерных технологий и прикладной математики протокол № 01 «28» августа 2025 г.

Председатель УМК факультета Коваленко А.В.



ПОДПИСЬ

Рецензенты:

Мостовой Евгений Викторович, генеральный директор ООО «Портал-Юг»,
e-mail: mostovoy@portal-yug.ru

Луценко Евгений Вениаминович, доктор экономических наук, кандидат технических наук, профессор кафедры компьютерных технологий и систем Федерального государственного бюджетного образовательного учреждения высшего образования «Кубанский государственный аграрный университет имени И.Т. Трубилина», e-mail: prof.lutsenko@gmail.com

1 Цели и задачи изучения дисциплины (модуля)

1. **1.1. Целью** освоения учебной дисциплины «Современные методы фронтальных исследований ИИ» является формирование у студентов знаний по основам теории оптимизации и знаний об основных подходах к практическому решению оптимизационных задач, что позволит применять методы анализа прикладной области на концептуальном, логическом, математическом и алгоритмическом уровнях, а также применять системный подход и математические методы в формализации решения прикладных задач.

1.2. Задачи дисциплины. В ходе изучения дисциплины ставятся задачи научить студентов:

- Освоить современные архитектуры ИИ: трансформеры, диффузионные модели, нейро-символьные системы
- Сформировать исследовательские компетенции: формулировка гипотез, планирование экспериментов, анализ результатов
- Развить практические навыки работы с ИИ: реализация моделей, работа с данными, оптимизация и развертывание
- Освоить методологию научной работы: критический анализ литературы, оформление результатов, научная этика
- Научить применять фронтальные методы для решения прикладных задач: адаптация современных подходов для индустриальных кейсов

1.3. Место учебной дисциплины в структуре образовательной программы

Дисциплина «Современные методы фронтальных исследований ИИ» относится к вариативной части Блока 1 «Дисциплины (модули)» учебного плана.

Данная дисциплина («Современные методы фронтальных исследований ИИ») тесно связана с дисциплинами: «Алгебра и геометрия», «Фундаментальные дискретные модели», «Курс теории вероятностей». Знания, полученные при освоении дисциплины «Современные методы фронтальных исследований ИИ», используются при изучении дисциплины «Высоконагруженные приложения». В совокупности изучение этой дисциплины готовит обучаемых формулировать научно-обоснованные гипотезы, направленные на совершенствование и расширение границ искусственного интеллекта, разрабатывать и проверять новые алгоритмы и модели, способные преодолевать существующие ограничения ИИ, осваивать и применять современные исследовательские методы, включая методы статистического анализа, моделирования и машинного обучения.

1.4. Профессиональные роли в структуре образовательной программы

Роль 1: Data Analyst (Аналитик данных)

Задачи:

1. Статистический анализ, визуализация данных, предварительная обработка.
2. Создание прогнозных моделей
3. Построение аналитических моделей для поддержки бизнес-решений.

Роль 2: MLOps (Специалист по эксплуатации ИИ)

Задачи:

1. DevOps для ML.
2. Автоматизация, мониторинг ML-систем.
3. Операционное управление жизненным циклом ML-моделей.

Роль 3: AI PM (Менеджер проектов ИИ)

Задачи:

1. Управление ИИ-проектами от идеи до внедрения
2. Анализ бизнес-требований и постановка задач
3. Оценка эффективности и ROI ИИ-решений

1.5. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код, уровень и формулировка компетенции	Индикаторы	Уровни освоения индикаторов компетенции
<p>ML-1</p> <p>Способен применять знания об истории развития и трендах современного ИИ для формулирования корректных постановок задач и поиска перспективных способов решения проблем с помощью ИИ</p>	<p>ML-1.1 Позиционирует собственную задачу в заданной области знания с точки зрения трендов современного искусственного интеллекта</p>	<p>(П) Демонстрирует базовое, фрагментарное понимание трендов ИИ и их связи с предметной областью. Выбирает современные, проверенные на практике методы и фреймворки (например, для NLP-задачи предлагает fine-tuning BERT-like моделей). Не просто применяет готовое решение, а предлагает адаптацию или синтез методов из смежных трендов для решения конкретной проблемы. Предлагаемый способ является нестандартным, но обоснованным.</p>
<p>ML-1</p> <p>Способен применять знания об истории развития и трендах современного ИИ для формулирования корректных постановок задач и поиска перспективных способов решения проблем с помощью ИИ</p>	<p>ML-1.2 Определяет тенденции развития, оценивает новизну и практическую значимость своих решений с точки зрения современного искусственного интеллекта</p>	<p>(П) Критически анализирует тренды, выявляя их ограничения и возможности. Дает взвешенную оценку новизны и значимости своего решения, определяя его место в "ландшафте" развития ИИ. Проводит глубокий сравнительный анализ с SOTA-решениями. Оценивает долгосрочные последствия. Говорит о значимости в терминах формирования новых рынков, изменения практик в индустрии или переопределения границ возможного для ИИ.</p>
<p>FC-1</p> <p>Способен проводить фронтальные исследования в области архитектур, алгоритмов МО, оптимизации и матема-</p>	<p>FC-1.1 Разрабатывает фундаментальные основы и новые алгоритмы машинного обучения</p>	<p>(Б) Свободно оперирует глубоким математическим аппаратом для построения собственных доказательств и выводов. Разрабатывает прототипы новых архитектур нейронных сетей и алгорит-</p>

Код, уровень и формулировка компетенции	Индикаторы	Уровни освоения индикаторов компетенции
тики		мов обучения на языках программирования (Python, C++) с использованием низкоуровневых библиотек (NumPy, JAX, PyTorch/TensorFlow на уровне операций). Может планировать и проводить корректные вычислительные эксперименты для проверки гипотез на синтетических и реальных данных.
FC-2 Способен проводить фронтирные исследования в области фундаментальных и генеративных моделей	FC-2.1 Исследует и разрабатывает большие языковые модели (LLM) и другие модели для символьных данных	(Б) Знает различия между декодерными (GPT), энкодер-декодерными (T5) и энкодерными (BERT) архитектурами, их сильные и слабые стороны. Понимает и может реализовать с нуля ключевые механизмы (например, self-attention, rotary positional encodings). Способен проанализировать, почему та или иная архитектура лучше подходит для конкретной задачи.
FC-3 Способен проводить фронтирные исследования в области управления, решения, агентных и мультиагентных систем	FC-3.2 Исследует и создает агентные системы	Умеет создавать агентные системы с нуля, оптимизируя их производительность и надежность. Владеет продвинутыми техниками интеграции (оркестрация нескольких моделей, работа с гетерогенными инструментами). Способен реализовывать сложные системы памяти (иерархические, ассоциативные) и механизмы извлечения информации.

Результаты обучения по дисциплине «Современные методы фронтирных исследований ИИ» достигаются в рамках осуществления всех видов контактной и самостоятельной работы обучающихся в соответствии с утвержденным учебным планом.

Индикаторы достижения компетенций считаются сформированными при достижении соответствующих им результатов обучения.

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 3 зач.ед. (108 часов), их распределение по видам работ представлено в таблице.

Вид учебной работы	Всего часов	Трудоёмкость, часов	
		7 семестр	
Контактная работа, в том числе:	52,2	52,2	
Аудиторная работа:	50	50	
Лекции (Л)	16	16	
Практические занятия (ПЗ)			
Лабораторные работы (ЛР)	34	34	
Иная контактная работа:			
Контроль самостоятельной работы (КСР)	2	2	
Промежуточная аттестация (ИКР)	0,2	0,2	
Самостоятельная работа (СР):	55,8	55,8	
Курсовой проект (КП), курсовая работа (КР)	-	-	
Самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам и т.д.)	30	30	
Выполнение индивидуальных заданий (подготовка сообщений, презентаций)	15,8	15,8	
Реферат	-	-	
Подготовка к текущему контролю	10	10	
Контроль:			
Подготовка и сдача экзамена ¹			
Общая трудоёмкость	час.	108	108
	в том числе контактная работа	56,2	56,2
	зач. ед	3	3
Вид итогового контроля		зачет	

2.2 Структура дисциплины:

Распределение видов учебной работы и их трудоёмкости по разделам дисциплины. Разделы (темы) дисциплины, изучаемые в 7 семестре

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Введение в фронтальные исследования ИИ	8,8	2		2	4,8
2.	Продвинутое проектирование глубокого обучения	16,5	2		8	6,5

¹ При наличии экзамена по дисциплине

3.	Обучение с подкреплением нового поколения	15,2	2		5	8,2
4.	Генеративные AI и творческие системы	18,5	2		7	9,5
5.	Интеллектуальные системы и нейро-символьный AI	14	2		4	8
6.	Эффективный и устойчивый AI	11,2	2		4	5,2
7.	Мультимодальный и воплощенный AI	11,2	2		2	7,2
8.	Будущее исследований и подготовка научной публикации	10,4	2		2	6,4
	<i>ИТОГО по разделам дисциплины</i>	105,8	16	0	34	55,8
	Контроль самостоятельной работы (КСР)	2				
	Промежуточная аттестация (ИКР)	0,2				
	Подготовка к текущему контролю					
	Общая трудоемкость по дисциплине	108				

Примечание: Л – лекции, ЛР – лабораторные занятия, СРС – самостоятельная работа студента, КСР -контролируемая самостоятельная работа студентов

2.3 Содержание разделов (тем) дисциплины

2.3.1 Занятия лекционного типа

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
1.	Введение в фронтальные исследования ИИ	Что такое "фронтальные исследования". Тренды и вызовы современного ИИ. Обзор ключевых направлений: от больших языковых моделей до нейробиологически инспирированных алгоритмов. Научная методология в ИИ: как формулировать исследовательский вопрос, проводить литературный обзор и оценивать вклад.	УО, ЛР
2.	Продвинутое архитектуры глубокого обучения	Эволюция архитектур нейронных сетей. Принципы Transformer: внимание, самовнимание, механизмы позиционного кодирования. Современные вариации Transformer (Sparse, Linformer, Perceiver). Диффузионные модели: принцип работы, тренировка, варианты архитектур (U-Net, U-Net с условием)..	ЛР
3.	Обучение с подкреплением нового поколения	Обучение с подкреплением (RL) для сложных сред. Meta-RL: обучение алгоритмов обучения. Иерархическое RL. RL	ЛР

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
		от человеческих предпочтений (RLHF) – ключевой метод для (alignment) больших моделей.	
4.	Генеративные AI и творческие системы	Пейзаж генеративных моделей: GAN, VAE, Normalizing Flows, Diffusion Models. Контролируемая генерация: техники guidance (Classifier-Free, Classifier). Мультиязычные генеративные модели (DALL-E, Stable Diffusion). Оценка качества генеративных моделей.	ЛР
5.	Интеллектуальные системы и нейро-символьный AI	Интеграция глубокого обучения и символьных рассуждений. Нейро-символьные модели. Обучение с подкреплением на графах знаний. AI для научных открытий: предсказание свойств материалов, ускорение разработки лекарств.	ЛР
6.	Эффективный и устойчивый AI	Проблема вычислительной стоимости больших моделей. Методы эффективного инференса: квантование, праунинг, дистилляция. Алгоритмы для устройств с ограниченными ресурсами (TinyML). Учет этических аспектов и снижение смещений в моделях.	ЛР
	Мультиязычный и воплощенный AI	Модели, работающие с текстом, изображением, звуком и другими модальностями одновременно. Архитектуры для мультиязычного обучения. "Воплощенный" интеллект (Embodied AI): AI в робототехнике и виртуальных мирах, обучение в симуляции.	ЛР
	Будущее исследований и подготовка научной публикации	Обзор самых свежих и спекулятивных направлений (искусственное обобщение, AGI). Как читать и понимать научные статьи по ИИ. Структура и написание собственной исследовательской статьи. Процесс рецензирования и подачи на конференции (NeurIPS, ICML, ICLR).	ЛР

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач, УО – Устный опрос.

2.3.2 Занятия семинарского типа

Семинарские занятия – не предусмотрены

2.3.3 Лабораторные занятия

№ п/п	Темы лабораторных занятий	Наименование лабораторных работ	Форма текущего контроля
1	Введение в фронтирные исследования ИИ	Лабораторная работа №1: Формулировка исследовательской гипотезы: анализ литературных источников: сравнение методов и результатов в табличной форме, выявление research gaps и нерешенных проблем, определение независимых и зависимых переменных, планирование экспериментальной проверки	ЛР
2	Продвинутое архитектуры глубокого обучения	Лабораторная работа №2: Трансформеры: реализовать и проанализировать механизм самовнимания, реализация параллельных голов внимания, конкатенация и линейная проекция, полный трансформер-блок, добавление residual connections и layer normalization, позиционные кодировки, Feed-forward network.	ЛР
3	Обучение с подкреплением нового поколения	Лабораторная работа №3: Модификация архитектуры: Linear attention через kernel approximation, Perceiver - cross-attention с латентными векторами, расчет метрик perplexity, Memory usage и inference time. Лабораторная работа №4: PPO и основы RLHF: методы обучения с подкреплением. GAE (Generalized Advantage Estimation), Value function baseline, Ratio clipping для стабильности, Entropy regularization, Reward modeling на парных сравнениях, Fine-tuning с KL-regularization	ЛР
4	Генеративные AI и творческие системы	Лабораторная работа №5: Диффузионные модели. Forward process, U-Net архитектура, Training и sampling. Лабораторная работа №6: методы управляемой генерации. Joint	ЛР

		training conditional/unconditional , Guidance scale при sampling, Маскирование внимания по условию, Cross-attention для текстовых условий, CLIP score для текст-изображение, Accuracy co-ответствия промпту	
5	Интеллектуальные системы и нейросимвольный AI	Лабораторная работа №7: Нейросимвольная интеграция. Создание гибридной системы. Named Entity Recognition, Relation extraction нейросетью, построение графа знаний, Rule-based reasoning на графе, Path finding для цепочек рассуждений.	ЛР
6	Эффективный и устойчивый AI	Лабораторная работа №8: Оптимизация моделей, освоение методов эффективного инференса. Post-training quantization, Quantization-aware training, Magnitude pruning, Structured pruning, Knowledge distillation	ЛР
	Мультимодальный и воплощенный AI	Лабораторная работа №9: Построить систему кросс-модального поиска. Image encoder (CNN/ViT), Text encoder (Transformer), InfoNCE loss, Hard negative mining, Text-to-image retrieval, Image-to-text retrieval	ЛР
	Будущее исследований и подготовка научной публикации	Лабораторная работа №10: Оформление научной статьи. Формат LaTeX (шаблон конференции), 4-6 страниц, включая схемы и таблицы с результатами	ЛР

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.4 Примерная тематика курсовых работ (проектов)

Курсовые работы (КР) – не предусмотрены.

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающегося по дисциплине

Целью самостоятельной работы студента является углубление знаний, полученных в результате аудиторных занятий. Вырабатываются навыки самостоятельной работы. Закрепляются опыт и знания полученные во время лабораторных занятий.

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1	Изучение теоретического материала	Методические указания по организации самостоятельной работы студентов, утвержденные УСФ, протокол №1 от

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
		30.06.2025
2	Решение задач	Методические указания по организации самостоятельной работы студентов, утвержденные УСФ, протокол №1 от 30.06.2025

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии

В соответствии с требованиями ФГОС в программа дисциплины предусматривает использование в учебном процессе следующих образовательные технологии: чтение лекций с использованием мультимедийных технологий; метод малых групп, разбор практических задач и кейсов.

При обучении используются следующие образовательные технологии:

- Технология коммуникативного обучения – направлена на формирование коммуникативной компетентности студентов, которая является базовой, необходимой для адаптации к современным условиям межкультурной коммуникации.

- Технология разноуровневого (дифференцированного) обучения – предполагает осуществление познавательной деятельности студентов с учётом их индивидуальных способностей, возможностей и интересов, поощряя их реализовывать свой творческий потенциал. Создание и использование диагностических тестов является неотъемлемой частью данной технологии.

- Технология модульного обучения – предусматривает деление содержания дисциплины на достаточно автономные разделы (модули), интегрированные в общий курс.

- Информационно-коммуникационные технологии (ИКТ) - расширяют рамки образовательного процесса, повышая его практическую направленность, способствуют интенсификации самостоятельной работы учащихся и повышению познавательной активности. В рамках ИКТ выделяются 2 вида технологий:

- Технология использования компьютерных программ – позволяет эффективно дополнить процесс обучения языку на всех уровнях.

- Интернет-технологии – предоставляют широкие возможности для поиска информации, разработки научных проектов, ведения научных исследований.

– Технология индивидуализации обучения – помогает реализовывать личностно-ориентированный подход, учитывая индивидуальные особенности и потребности учащихся.

– Проектная технология – ориентирована на моделирование социального взаимодействия учащихся с целью решения задачи, которая определяется в рамках профессиональной подготовки, выделяя ту или иную предметную область.

– Технология обучения в сотрудничестве – реализует идею взаимного обучения, осуществляя как индивидуальную, так и коллективную ответственность за решение учебных задач.

– Игровая технология – позволяет развивать навыки рассмотрения ряда возможных способов решения проблем, активизируя мышление студентов и раскрывая личностный потенциал каждого учащегося.

– Технология развития критического мышления – способствует формированию разносторонней личности, способной критически относиться к информации, умению отбирать информацию для решения поставленной задачи.

Комплексное использование в учебном процессе всех вышеназванных технологий стимулируют личностную, интеллектуальную активность, развивают познавательные процессы, способствуют формированию компетенций, которыми должен обладать будущий специалист.

Основные виды интерактивных образовательных технологий включают в себя:

– работа в малых группах (команде) - совместная деятельность студентов в группе под руководством лидера, направленная на решение общей задачи путём творческого сложения результатов индивидуальной работы членов команды с делением полномочий и ответственности;

– проектная технология - индивидуальная или коллективная деятельность по отбору, распределению и систематизации материала по определенной теме, в результате которой составляется проект;

– анализ конкретных ситуаций - анализ реальных проблемных ситуаций, имевших место в соответствующей области профессиональной деятельности, и поиск вариантов лучших решений;

– развитие критического мышления – образовательная деятельность, направленная на развитие у студентов разумного, рефлексивного мышления, способного выдвинуть новые идеи и увидеть новые возможности.

Подход разбора конкретных задач и ситуаций широко используется как преподавателем, так и студентами во время лекций, лабораторных занятий и анализа результатов самостоятельной работы. Это обусловлено тем, что при исследовании и решении каждой конкретной задачи имеется, как правило, несколько методов, а это требует разбора и оценки целой совокупности конкретных ситуаций.

Семестр	Вид занятия	Используемые интерактивные образовательные технологии	количество интерактивных часов
	ЛР	Практические занятия в режимах взаимодействия «преподаватель – студент» и «студент – студент»	12
Итого			12

Примечание: Л – лекции, ПЗ – практические занятия/семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

Темы, задания и вопросы для самостоятельной работы призваны сформировать навыки поиска информации, умения самостоятельно расширять и углублять знания, полученные в ходе лекционных и практических занятий.

Подход разбора конкретных ситуаций широко используется как преподавателем, так и студентами при проведении анализа результатов самостоятельной работы.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

4. Оценочные и методические материалы

4.1 Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Современные методы фронтальных исследований ИИ».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме индивидуальных самостоятельных заданий и **промежуточной аттестации** в форме вопросов к экзамену.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями.

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Структура оценочных средств для текущей и промежуточной аттестации

№ п/п	Контролируемые разделы (темы) дисциплины*	Код контролируемой компетенции (или ее части)	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация
1	Введение в фронтальные исследования ИИ	ML-1.1; ML-1.2; FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №1</i>	<i>Вопросы к зачету</i>
2	Продвинутое архитектуры глубокого обучения	FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №2</i>	<i>Вопросы к зачету</i>
3	Обучение с подкреплением нового поколения	FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №3, 4</i>	<i>Вопросы к зачету</i>
4	Генеративные AI и творческие системы	FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №5,6</i>	<i>Вопросы к зачету</i>
5	Интеллектуальные системы и нейросимвольный AI	FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №7</i>	<i>Вопросы к зачету</i>
6	Эффективный и устойчивый AI	FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №8</i>	<i>Вопросы к зачету</i>
7	Мультимодальный и воплощенный AI	FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №9</i>	<i>Вопросы к зачету</i>
8	Будущее исследований и подготовка научной публикации	ML-1.1; ML-1.2; FC-1.1; FC-2.1; FC-3.2	<i>Тестирование, Лабораторные работы №10</i>	<i>Вопросы к зачету</i>

Показатели, критерии и шкала оценки сформированных компетенций

Соответствие **пороговому уровню** освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **удовлетворительно /зачтено**):

ML-1 (II) Способен применять знания об истории развития и трендах современного ИИ для формулирования корректных постановок задач и поиска перспективных способов решения проблем с помощью ИИ

ML-1.1 Позиционирует собственную задачу в заданной области знания с точки зрения трендов современного искусственного интеллекта

Знать Основные направления и актуальные тренды развития искусственного интеллекта. Ключевые задачи и проблемы, решаемые современными методами ИИ. Примеры успешного применения ИИ в близких областях знаний.

Уметь Определять принадлежность собственной задачи к той или иной группе проблем, характерных для ИИ.

Обнаруживать сходства и различия между собственными задачами и примерами аналогичных задач, решаемых с помощью ИИ. Применять базовые приёмы и инструменты позиционирования задачи среди трендовых

	направлений ИИ. Начальными навыками анализа научной литературы и новостных статей, связанных с искусственным интеллектом.
Владеть	Способностью аргументировано обосновывать свою позицию в дискуссиях о месте собственной задачи в контексте развития ИИ. Пониманием, как представить свою задачу с позиций современных трендов и перспектив ИИ.
ML-1.2	Определяет тенденции развития, оценивает новизну и практическую значимость своих решений с точки зрения современного искусственного интеллекта
Знать	Основные направления и тенденции развития искусственного интеллекта. Общие черты прорывных решений и нововведений в области ИИ. Признаки новизны и уникальности решения в контексте искусственного интеллекта. Оценивать новизну своего решения, сравнивая его с существующими аналогами.
Уметь	Формулировать основную идею и ценность собственного решения с точки зрения современных трендов ИИ. Аргументировать полезность и потенциальную выгоду предложенного решения для практики. Начальными навыками анализа научной и отраслевой литературы по искусственному интеллекту.
Владеть	Пониманием и аргументацией преимуществ и недостатков предложенного решения. Способностью объяснить, почему данное решение актуально именно сейчас и соответствует последним достижениям в области ИИ.
FC-1	(Б) Способен проводить фронтальные исследования в области архитектур, алгоритмов МО, оптимизации и математики
FC-1.1	Разрабатывает фундаментальные основы и новые алгоритмы машинного обучения
Знать	Основы математического аппарата: линейная алгебра, теория вероятностей, математическая статистика, основы оптимизации. Фундаментальные алгоритмы МО: линейные модели, методы кластеризации, деревья решений, SVM, базовые нейросетевые архитектуры. Критерии оценки алгоритмов: понятия переобучения, регуляризации, методы валидации (кросс-валидация, train/test split). Принципы построения исследований: структура научной статьи, формулировка гипотезы, планирование эксперимента. Анализировать существующие алгоритмы: сравнивать методы, выявлять их ограничения и области применения. Модифицировать алгоритмы: вносить изменения в существующие методы (например, добавлять регуляризацию, менять функцию потерь).
Уметь	Реализовывать базовые алгоритмы: программировать с нуля классические методы (k-Means, линейная регрессия, наивный байесовский классификатор). Проводить эксперименты: оценивать качество моделей, интерпретировать метрики (accuracy, precision, recall, F1-score). Работать с данными: выполнять предобработку, feature engineering, анализ распределений.
Владеть	Навыками программирования: Python, библиотеки (NumPy, Pandas, Scikit-learn), основы работы с Jupyter Notebook, методами отладки моделей: анализ ошибок, диагностика переобучения, подбор гиперпараметров, инструментами визуализации: построение графиков (Matplotlib, Seaborn) для анализа данных и результатов, базовыми практиками исследований: поиск научных статей, оформление результатов в виде отчета.
FC-2	Способен проводить фронтальные исследования в области фундаментальных и генеративных моделей
FC-2.1	Исследует и разрабатывает большие языковые модели (LLM) и другие модели для символьных данных
Знать	Архитектурные основы: принципы работы Transformer, механизм внимания (self-attention), структура encoder-decoder Токенизация: методы Byte Pair Encoding (BPE), WordPiece, SentencePiece Области применения LLM: классификация текста, генерация, суммаризация, вопросно-ответные системы Типы символьных данных: тексты, последовательности, графы, деревья Базовые подходы к символьным данным: графовые нейронные сети, рекуррентные архитек-

	туры
	Этические аспекты: смещения в данных, конфиденциальность, безопасность генерации
	Работать с предобученными моделями: загружать и использовать модели из Hugging Face Hub
Уметь	Fine-tuning LLM: адаптировать существующие модели под конкретные задачи Подготавливать данные: токенизация, создание датасетов для обучения, паддинг и маскирование Оценивать качество моделей: вычислять перплексию, BLEU, ROUGE метрики Визуализировать внимание: анализировать матрицы внимания для интерпретации решений модели
Владеть	Применять графовые модели: строить простые GNN для анализа символьных структур Инструментарием NLP: Hugging Face Transformers, Tokenizers, Datasets Фреймворками глубокого обучения: PyTorch или TensorFlow Библиотеками для работы с графами: NetworkX, PyTorch Geometric (базовый уровень) Методами предобработки текста: очистка, нормализация, лемматизация Навыками отладки моделей: анализ потерь, диагностика переобучения, мониторинг обучения
ФС-3	(II) Способен проводить фронтальные исследования в области управления, решения, агентных и мультиагентных систем
ФС-3.2	Исследует и создает агентные системы
Знать	Базовые концепции агентов: архитектура "восприятие-мышление-действие", типы агентов (реактивные, основанные на модели, на целях) Методы принятия решений: деревья решений, конечные автоматы, базовые правила поведения Основы обучения с подкреплением: марковский процесс принятия решений (MDP), ценность состояния и действия, политика агента Инструменты и среды: OpenAI Gym, Unity ML-Agents, PettingZoo для многоагентных систем Архитектурные паттерны: системы с черной доской, многоагентное взаимодействие (кооперация, конкуренция) Проектировать простых агентов: создавать реактивные системы на основе правил и конечных автоматов
Уметь	Реализовывать базовые алгоритмы RL: Q-learning, Policy Gradients для простых сред Интегрировать ИИ-компоненты: соединять планировщики, системы восприятия и модули действий Настраивать среды обучения: создавать custom среды в Gym, определять пространства действий и наблюдений Визуализировать поведение: отслеживать и анализировать траектории агентов, строить графики обучения
Владеть	Тестировать агентов: проводить A/B тестирование разных стратегий, оценивать Robustness Библиотеками RL: Stable-Baselines3, Ray RLLib, Tianshou Инструментами симуляции: OpenAI Gym API, Unity ML-Agents Toolkit Методами оценки: расчет cumulative reward, анализ сходимости, сравнение политик Техниками отладки: визуализация value functions, трассировка решений, логирование действий Паттернами проектирования: State Pattern для управления поведением, Strategy Pattern для смены политик

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

Практические кейсы по тематике лабораторных работ

ЛАБОРАТОРНАЯ РАБОТА №1: Формулировка исследовательской гипотезы

Кейс 1: Оптимизация процесса научных исследований в фармацевтике

Постановка задачи: Фармацевтическая компания тратит значительные ресурсы на анализ научных публикаций для поиска перспективных молекул. Необходимо автоматизировать процесс выявления research gaps и формулировки исследовательских гипотез.

Требования:

- Проанализируйте 50+ статей по целевой терапевтической области
- Сформулируйте 3 перспективные исследовательские гипотезы
- Разработайте план экспериментальной проверки для каждой гипотезы
- Подготовьте аннотации проектов в формате научного предложения
- Предложите метрики оценки успешности исследований

ЛАБОРАТОРНАЯ РАБОТА №2: Трансформеры**Кейс 2: Улучшение семантического поиска в юридической базе**

Постановка задачи: Юридическая фирма нуждается в системе точного поиска прецедентов по смысловому сходству, а не только по ключевым словам.

Требования:

- Реализуйте механизм самовнимания для векторного представления документов
- Настройте multi-head attention для анализа юридических текстов
- Визуализируйте матрицы внимания для интерпретации результатов
- Протестируйте систему на базе судебных решений
- Сравните точность с традиционными методами поиска

ЛАБОРАТОРНАЯ РАБОТА №3: Модификация архитектуры

Кейс 3: Оптимизация архитектуры для обработки длинных последовательностей

Постановка задачи: Fintech компания обрабатывает длинные временные ряды котировок. Стандартные трансформеры не справляются с вычислительной сложностью.

Требования:

- Реализуйте sparse attention для оконного анализа временных рядов
- Настройте linear attention через kernel approximation
- Сравните качество предсказаний на финансовых данных
- Проанализируйте выигрыш в памяти и скорости обработки
- Предложите оптимальную архитектуру для production

ЛАБОРАТОРНАЯ РАБОТА №4: PPO и основы RLHF**Кейс 4: Обучение торгового агента для фондового рынка**

Постановка задачи: Хедж-фонд хочет автоматизировать торговые стратегии с учетом рыночных условий.

Требования:

- Реализуйте PPO для обучения торгового агента
- Настройте reward function с учетом риска и доходности
- Интегрируйте различные источники данных (котировки, новости)
- Протестируйте стратегию на исторических данных
- Проведите backtesting с сравнением с бенчмарком

ЛАБОРАТОРНАЯ РАБОТА №5: Диффузионные модели**Кейс 5: Генерация молекулярных структур для drug discovery**

Постановка задачи: Биотех компания ищет новые подходы к генерации перспективных молекулярных структур.

Требования:

- Реализуйте диффузионную модель для генерации молекул

- Настройке контроль свойств генерируемых соединений
- Интегрируйте химические знания в процесс генерации
- Оцените валидность и новизну сгенерированных структур
- Сравните с существующими методами генерации

ЛАБОРАТОРНАЯ РАБОТА №6: Контроль генерации

Кейс 6: Персонализированная генерация образовательного контента

Постановка задачи: Edtech стартап разрабатывает систему адаптивного обучения с генерацией индивидуальных материалов.

Требования:

- Реализуйте classifier-free guidance для контроля сложности
- Настройте attention control для тематической фокусировки
- Разработайте метрики оценки образовательной ценности
- Протестируйте на различных предметных областях
- Оцените персональную релевантность контента

ЛАБОРАТОРНАЯ РАБОТА №7: Нейро-символьная интеграция

Кейс 7: Построение медицинской онтологии для диагностики

Постановка задачи: Медицинский исследовательский центр создает систему поддержки диагностики на основе графа знаний.

Требования:

- Постройте граф знаний медицинских понятий и симптомов
- Настройте извлечение отношений из клинических текстов
- Реализуйте логический вывод для дифференциальной диагностики
- Визуализируйте цепочки диагностических рассуждений
- Оцените точность на тестовых клинических случаях

ЛАБОРАТОРНАЯ РАБОТА №8: Оптимизация моделей

Кейс 8: Оптимизация мобильного приложения компьютерного зрения

Постановка задачи: Разработчик мобильных приложений нуждается в эффективной модели для распознавания объектов на устройстве.

Требования:

- Примените квантование для уменьшения размера модели
- Проведите прунинг для ускорения инференса
- Реализуйте дистилляцию знаний в компактную модель
- Сравните производительность на мобильном устройстве
- Добейтесь баланса между точностью и скоростью

ЛАБОРАТОРНАЯ РАБОТА №9: Мультимодальное обучение

Кейс 9: Кросс-модальный поиск для арт-галереи

Постановка задачи: Виртуальный музей хочет улучшить поиск произведений искусства по текстовым описаниям и изображениям.

Требования:

- Настройте dual-encoder для совместного embedding пространства
- Реализуйте contrastive learning для пар изображение-текст
- Оцените качество поиска на каталоге произведений искусства
- Сравните с одомодальными методами поиска
- Протестируйте на различных художественных стилях

ЛАБОРАТОРНАЯ РАБОТА №10: Оформление научной статьи

Кейс 10: Подготовка публикации по результатам исследований

Постановка задачи: Научная группа завершила серию экспериментов и готовит статью для топовой конференции.

Требования:

- Структурируйте результаты всех проведенных экспериментов
- Подготовьте сравнительные таблицы и графики
- Сформулируйте научный вклад и ограничения
- Оформите статью в стиле конференции NeurIPS
- Подготовьте ответы на потенциальные вопросы рецензентов

Пример лабораторной работы

Лабораторная работа №7: Нейро-символьная интеграция

Тема: Построение медицинской онтологии для диагностики

Цель:

Разработать гибридную нейро-символьную систему для поддержки медицинской диагностики на основе анализа клинических текстов и логического вывода.

Задачи:

1. Изучить принципы нейро-символьного интегрирования и построения онтологий.
2. Реализовать извлечение медицинских сущностей из текста с помощью трансформеров.
3. Построить граф знаний на основе извлеченных данных.
4. Реализовать логический вывод для проверки диагностических гипотез.

Ожидаемые результаты:

Знает:

- Основы нейро-символьного искусственного интеллекта.
- Принципы построения онтологий и графов знаний.

Умеет:

- Извлекать медицинские сущности и отношения из текста.
- Строить графы знаний и выполнять на них логический вывод.

Владеет:

- Инструментами для обработки естественного языка (spaCy, Transformers).
- Библиотеками для работы с графами (NetworkX).

Инструменты и библиотеки:

- Язык программирования Python.
- Jupyter Notebook / Google Colab.
- Библиотеки: spaCy, transformers, torch, networkx, matplotlib.

Исходные данные:

- Датасет клинических текстов с описанием симптомов, диагнозов и методов лечения.
- Медицинские онтологии (например, SNOMED CT или UMLS).

Ход работы:

1. Установка библиотек

```
python
!pip install spacy transformers torch networkx matplotlib
!python -m spacy download en_core_web_sm
```

2. Загрузка и предобработка данных

```
python
import pandas as pd
import spacy
```

```
from transformers import AutoTokenizer, AutoModelForTokenClassification
import torch
```

```
# Загрузка датасета
```

```
data = pd.read_csv("clinical_texts.csv")
texts = data["text"].tolist()
```

```
# Загрузка моделей для обработки текста
```

```
nlp = spacy.load("en_core_web_sm")
tokenizer = AutoTokenizer.from_pretrained("dbmdz/bert-large-cased-finetuned-conll03-english")
model = AutoModelForTokenClassification.from_pretrained("dbmdz/bert-large-cased-finetuned-conll03-english")
```

3. Извлечение медицинских сущностей

```
python
```

```
def extract_entities(text):
    doc = nlp(text)
    entities = []
    for ent in doc.ents:
        if ent.label_ in ["DISEASE", "SYMPTOM", "TREATMENT"]:
            entities.append((ent.text, ent.label_))
    return entities
```

```
# Пример извлечения сущностей
```

```
sample_text = "Patient presents with fever and cough. Diagnosed with pneumonia."
entities = extract_entities(sample_text)
print(entities) # [('fever', 'SYMPTOM'), ('cough', 'SYMPTOM'), ('pneumonia', 'DISEASE')]
```

4. Построение графа знаний

```
python
```

```
import networkx as nx
import matplotlib.pyplot as plt
```

```
# Создание графа
```

```
G = nx.DiGraph()
```

```
# Добавление узлов и рёбер на основе извлеченных сущностей
```

```
for text in texts:
    entities = extract_entities(text)
    for i, (entity1, label1) in enumerate(entities):
        G.add_node(entity1, label=label1)
    for j, (entity2, label2) in enumerate(entities):
        if i != j:
            G.add_edge(entity1, entity2, relationship="associated_with")
```

```
# Визуализация графа
```

```
plt.figure(figsize=(12, 8))
pos = nx.spring_layout(G)
nx.draw(G, pos, with_labels=True, node_color="lightblue", font_size=10)
plt.title("Медицинский граф знаний")
plt.show()
```

5. Логический вывод на графе

```
python
```

```
def diagnose_symptoms(symptoms, graph):
    possible_diseases = []
    for symptom in symptoms:
        for node in graph.nodes:
            if graph.nodes[node]["label"] == "DISEASE":
                if nx.has_path(graph, symptom, node):
                    possible_diseases.append(node)
    return list(set(possible_diseases))
```

Пример диагностики

```
symptoms = ["fever", "cough"]
diagnoses = diagnose_symptoms(symptoms, G)
print("Возможные диагнозы:", diagnoses) # ['pneumonia']
```

6. Интеграция с трансформерами для уточнения диагнозов

python

```
from transformers import pipeline
```

Использование модели для классификации текста

```
classifier = pipeline("text-classification", model="bvanaken/clinical-assertion-negation-bert")
statement = "Patient has no fever."
result = classifier(statement)
print(result) # [{'label': 'negated', 'score': 0.98}]
```

Требования к отчету

1. **Титульный лист:** Название работы, ФИО студента, группа, дата.
2. **Введение:** Цель, задачи, описание датасета и медицинской области.
3. **Математическая модель:**
 - Описание архитектуры трансформеров для извлечения сущностей.
 - Принципы построения графов знаний и логического вывода.
4. **Реализация:**
 - Код с комментариями.
 - Примеры извлеченных сущностей и визуализация графа.
5. **Результаты:**
 - Метрики извлечения сущностей (точность, полнота).
 - Примеры диагностических выводов.
6. **Выводы:**
 - Проблемы при обработке текстов (например, негативные утверждения).
 - Предложения по улучшению системы.
7. **Приложения:** Исходный код, примеры текстов и графов.

Критерии оценки:

Отлично: Полная реализация системы, анализ метрик, корректный логический вывод.

Хорошо: Корректное извлечение сущностей и построение графа без глубокого анализа.

Удовлетворительно: Частичное выполнение с ошибками в логическом выводе.

Неудовлетворительно: Невыполнение ключевых этапов (извлечение сущностей, построение графа).

Проверяемые компетенции комплексом практических заданий: FC-1.1; FC-2.1; FC-3.2

Зачетно-экзаменационные материалы для промежуточной аттестации (зачет)

Вопросы для подготовки к зачету

Теоретические вопросы

Тема 1: Введение в фронтирные исследования ИИ

1. Дайте определение фронтирным исследованиям в ИИ. Чем они отличаются от прикладных разработок?
2. Назовите 3 наиболее перспективных направления фронтирных исследований ИИ на текущий момент.
3. Опишите структуру научной статьи в области ИИ. Какие разделы являются обязательными?
4. В чем заключаются основные проблемы воспроизводимости в исследованиях ИИ?

Тема 2: Продвинутое архитектуры глубокого обучения

5. Объясните принцип работы механизма самовнимания (self-attention) в трансформерах.
6. Сравните архитектуры Transformer, Perceiver и Linear Transformer.
7. Для чего используется позиционное кодирование в трансформерах?
8. Опишите преимущества и ограничения sparse attention.

Тема 3: Обучение с подкреплением нового поколения

9. В чем отличие PPO от традиционных методов обучения с подкреплением?
10. Опишите процесс RLHF (Reinforcement Learning from Human Feedback).
11. Какие проблемы решает метод GAE (Generalized Advantage Estimation)?
12. Приведите пример применения многоагентного RL в реальных задачах.

Тема 4: Генеративные AI и творческие системы

13. Опишите прямой и обратный процессы в диффузионных моделях.
14. В чем отличие DDPM от DDIM?
15. Как работает Classifier-Free Guidance для управления генерацией?
16. Сравните диффузионные модели с GAN и VAE.

Тема 5: Интеллектуальные системы и нейро-символьный AI

17. Что такое нейро-символьный искусственный интеллект?
18. Опишите процесс построения графа знаний из текстовых данных.
19. Какие методы логического вывода используются на графах знаний?
20. Приведите пример практического применения нейро-символьных систем.

Тема 6: Эффективный и устойчивый AI

21. Опишите методы квантования нейронных сетей и их влияние на производительность.
22. В чем разница между magnitude pruning и structured pruning?
23. Как работает knowledge distillation и в каких случаях она наиболее эффективна?
24. Какие этические аспекты необходимо учитывать при разработке больших моделей ИИ?

Тема 7: Мультимодальный и воплощенный AI

25. Дайте определение мультимодального обучения. Приведите примеры задач.
26. Как работает архитектура dual-encoder для кросс-модального поиска?
27. Что такое Embodied AI и каковы его основные challenges?
28. Опишите применение трансформеров для обработки мультимодальных данных.

Тема 8: Будущее исследований и подготовка научной публикации

29. Каковы критерии оценки научного вклада в области ИИ?
30. Опишите процесс рецензирования на конференциях NeurIPS/ICML.
31. Какие emerging trends в фронтирных исследованиях ИИ вы считаете наиболее перспективными?
32. Как обеспечить этичность и безопасность исследований в области ИИ?

Практические вопросы

Перечень компетенций (части компетенции), проверяемых оценочным средством ML-1.1; ML-1.2; FC-1.1; FC-2.1; FC-3.2

Вопросы по практической части

ЛР1: Формулировка исследовательской гипотезы

33. Опишите этапы формулировки исследовательской гипотезы в области ИИ.
34. Какие источники информации вы использовали бы для анализа state-of-the-art?
35. Как оценить практическую значимость исследовательской гипотезы?

ЛР2: Трансформеры

36. Опишите реализацию механизма многоголового внимания.
37. Как визуализировать матрицы внимания и какую информацию они предоставляют?
38. Какие проблемы возникают при обучении трансформеров с нуля?

ЛР3: Модификация архитектуры

39. Сравните эффективность sparse attention и linear attention.
40. Какие метрики используются для оценки эффективности модифицированных архитектур?
41. Как выбрать подходящую модификацию архитектуры для конкретной задачи?

ЛР4: PPO и основы RLHF

42. Опишите процесс сбора rollout в PPO.
43. Как рассчитывается advantage function в GAE?
44. Какие особенности настройки reward model в RLHF?

ЛР5: Диффузионные модели

45. Опишите процесс обучения диффузионной модели.
46. Как работает U-Net архитектура в диффузионных моделях?
47. Сравните качество генерации DDPM и DDIM.

ЛР6: Контроль генерации

48. Как реализовать контроль стиля в генеративных моделях?
49. Какие метрики используются для оценки управляемости генерации?
50. Опишите работу Classifier-Free Guidance.

ЛР7: Нейро-символьная интеграция

51. Опишите процесс извлечения сущностей и отношений из текста.
52. Как построить граф знаний на основе извлеченных данных?
53. Какие методы логического вывода можно реализовать на графах?

ЛР8: Оптимизация моделей

54. Сравните эффективность различных методов квантования.
55. Как прунинг влияет на производительность модели?
56. Опишите процесс knowledge distillation.

ЛР9: Мультиязычное обучение

57. Как работает contrastive learning в мультиязычных системах?
58. Опишите процесс оценки качества кросс-модального поиска.
59. Какие особенности обработки разных модальностей в одной модели?

ЛР10: Оформление научной статьи

60. Какие разделы должны быть включены в научную статью по ИИ?
61. Как правильно представить экспериментальные результаты?
62. Какие требования к воспроизводимости результатов исследования?

Перечень компетенций (части компетенции), проверяемых оценочным средством ML-1.1; ML-1.2; FC-1.1; FC-2.1; FC-3.2

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

Методические рекомендации, определяющие процедуры оценивания на зачете:

Процедура промежуточной аттестации проходит в соответствии с Положением о текущем контроле и промежуточной аттестации обучающихся ФГБОУ ВО «КубГУ».

Итоговой формой контроля сформированности компетенций у обучающихся по дисциплине является зачет. Студенты обязаны сдать зачет в соответствии с расписанием и учебным планом.

ФОС промежуточной аттестации состоит из вопросов к зачету и результатов текущего контроля.

Зачет по дисциплине преследует цель оценить работу студента за курс, получение теоретических знаний, их прочность, развитие творческого мышления, приобретение навыков самостоятельной работы, умение применять полученные знания для решения практических задач.

Форма проведения зачета: устно.

Результат сдачи зачета заносится преподавателем в зачетную ведомость и зачетную книжку.

Оценивание уровня освоения дисциплины основывается на качестве выполнения студентом заданий текущего контроля и ответов на вопросы зачета.

Критерии оценки:

1. Оценка ответов на вопросы к зачету (50% итоговой оценки)

Зачет

Полные, развернутые ответы с демонстрацией глубокого понимания темы.

Ответы содержат основные идеи, но без углубленного анализа.

Использование примеров, формул, корректных терминов.

Возможны небольшие ошибки в деталях или формулировках.

Умение анализировать и сравнивать методы.

% выполнения: 60–100% (допускаются незначительные неточности).

Незачет

Отсутствие понимания ключевых концепций.

Грубые ошибки или неспособность ответить на большую часть вопросов.

% выполнения: <60%.

2. Оценка выполнения практических кейсов и лабораторных работ (50% итоговой оценки)

Зачет

Полное выполнение всех этапов кейса с инновационными решениями.

Достижение целевых метрик (например, $F1 > 0.9$).

Четкая документация кода и анализ результатов.

% выполнения: 60–100%.

или

Выполнены основные задачи, но без дополнительной оптимизации.

Незначительные отклонения от целевых метрик (например, $F1 = 0.85$).

или

Решены базовые задачи, но с критическими ошибками.

Низкое качество кода или отсутствие анализа.

Незачет

Невыполнение ключевых этапов.
Код нерабочий или отсутствует.
% выполнения: <60%.

Методические рекомендации, определяющие процедуры оценивания лабораторных работ:

Процедура оценивания лабораторных работ проходит в соответствии с Положением о текущем контроле и промежуточной аттестации обучающихся ФГБОУ ВО «КубГУ».

По каждой лабораторной работе оформляется отчет. Отчеты сдаются на проверку руководителю в течение курса по мере их выполнения, и защищаются студентами в установленном порядке.

При защите отчета студенту могут быть заданы вопросы и дополнительные задания по сути лабораторной работы, в том числе из списка контрольных вопросов к данной лабораторной работе. При неудовлетворительной оценке знаний студента по теме данного отчета, студент возвращается к повторному изучению соответствующих материалов, после чего допускается к повторной защите. Неудовлетворительно выполненный отчет также возвращается на доработку.

Отчет должен содержать заголовок, тему лабораторной работы, цель, задание, индивидуальную тему, описание хода выполнения работы, необходимые прикладные материалы (схемы, макеты документов и т.п.), в соответствии с требованиями к содержанию, и выводы по работе.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на зачете;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

4.3. Методические указания по организации вычислительной инфраструктуры **Условия применения:**

- Курс рассчитан на студентов 3-го года обучения.

- Наличие доступа к вычислительным ресурсам (GitLab, Google Colab или Yandex DataSphere, JupyterHub, Hugging Face).
- Разработаны лабораторные работы;
- Инфраструктура для приёма задач (gitlab, CI/CD) согласована с лабораторными работами.

Цели, задачи и ожидаемые результаты

Цели организации вычислительной инфраструктуры:

- дать начальное представление о работе в IT инфраструктуре (приучить пользоваться гитом, jupyter-ноутбуками).

Задачи преподавателя:

- Организация регистрации студентов в Google Colab и Yandex DataSphere
- Создание учетных записей студентов в gitlab вуза;
- Настройка GitLab Runner для автоматического тестирования кода.
- Разработка шаблонного репозитория для лабораторных работ с предустановленными зависимостями (PyTorch, Keras, TensorFlow, Hugging Face Transformers).
- Написание автотестов для проверки корректности выполнения заданий (например, fine-tuning моделей).
- Визуализация результатов тестирования через HTML-отчеты.
- Подготовка инструкций по работе с Git и облачными ресурсами.

Ожидаемые результаты студентов:

- начальное представление о работе в IT инфраструктуре (гит, нейминг).
- Навыки запуска и тестирования глубоких нейросетей в облачных средах.
- Понимание CI/CD-процессов в контексте разработки глубоких нейросетей.

Порядок реализации

Задача №1: Организация регистрации студентов в Google Colab и Yandex DataSphere

Задача №2: Создание учетных записей студентов в gitlab вуза

Задача №3: Настройка GitLab Runner:

Для автоматического тестирования кода используется Docker-образ с предустановленными библиотеками (PyTorch, Keras, TensorFlow, Hugging Face Transformers).

Для выполнения CI/CD пайплайна был настроен GitLab Runner на удалённой виртуальной машине с ОС Ubuntu 24.04.

Последовательность настройки включала следующие шаги:

- Настройка системы – установка необходимых компонентов, таких, как Docker.
- Установка GitLab Runner по официальной инструкции.
- Регистрация Runner для частного сервера GitLab.

Задача №4: Шаблонный репозиторий:

Включает:

.gitlab-ci.yml для CI/CD.

Скрипты для предобработки текста и обучения моделей.

Примеры кода для работы с BERT, GPT и другими архитектурами.

Задача №5: Автотесты:

Проверяют корректность fine-tuning моделей (например, ассурасу на тестовом датасете).

Задача №6: Визуализация результатов:

Генерация HTML-отчетов с результатами тестирования, включая метрики качества моделей.

Порядок проверки корректности:

- Наличие Git-репозитория у всех студентов.
- Шаблонный репозиторий с подключенными автотестами.
- Инструкция по работе с Git и CI/CD в формате README.md.

Вся структура максимально адаптирована для копирования студентами и минимизации порога входа при выполнении лабораторных

4.4. Методические указания по организации лабораторных работ

Условия применения:

- Курс рассчитан на студентов 4-го года обучения.
- Наличие доступа к вычислительным ресурсам (GitLab) и к GPU/CPU (Kaggle, локальные серверы).
- Разработана инфраструктура для приёма задач (Gitlab, CI/CD) и согласована с лабораторными работами и настроена на всех студентов образовательной программы;
- Использование открытых датасетов и библиотек.

Цели, задачи и ожидаемые результаты

- Закрепление теоретических знаний в области современных методов искусственного интеллекта
- Развитие навыков проведения независимых научных исследований в области ИИ
- Подготовка к решению сложных исследовательских задач в академической и индустриальной среде
- Формирование компетенций в области разработки и оптимизации передовых архитектур нейронных сетей

Задачи преподавателя:

- Обеспечить студентов структурированным циклом лабораторных работ от формулировки гипотезы до публикации результатов
- Предоставить доступ к необходимым вычислительным ресурсам (GPU-кластеры, облачные платформы)
- Организовать регулярную проверку и содержательную обратную связь по выполненным работам
- Создать условия для развития научного мышления и исследовательских навыков

Ожидаемые результаты студентов:

- Умение формулировать и проверять научные гипотезы в области ИИ
- Владение современными архитектурами глубокого обучения (Transformer, Diffusion Models, нейро-символьные системы)
- Опыт реализации и модификации передовых алгоритмов машинного обучения
- Навыки оптимизации и сжатия больших моделей для production-среды
- Умение оформлять результаты исследований в формате научных статей
- Опыт работы с мультимодальными данными и системами

Порядок реализации

Задача №1: Подготовка лабораторных работ(в соответствии с рабочей программой дисциплины)

Определение тем лабораторных работ:

- 1. Формулировка исследовательской гипотезы**
 - Анализ state-of-the-art и выявление research gaps
 - Планирование эксперимента и метрик оценки
- 2. Реализация архитектуры Transformer**
 - Построение механизма самовнимания с нуля
 - Визуализация и анализ матриц внимания
- 3. Модификация архитектур глубокого обучения**
 - Реализация sparse attention и linear transformer
 - Сравнительный анализ эффективности архитектур
- 4. Обучение с подкреплением нового поколения**
 - Реализация PPO и основ RLHF
 - Fine-tuning моделей с учетом человеческих предпочтений
- 5. Диффузионные модели для генерации**

- Построение и обучение диффузионных процессов

- Сравнение методов DDPM и DDIM

6. Контроль генерации в творческих системах

- Реализация classifier-free guidance

- Управление стилем и содержанием генерации

7. Нейро-символьная интеграция

- Построение графов знаний и логический вывод

- Создание гибридных интеллектуальных систем

8. Оптимизация и сжатие моделей

- Применение квантования, прунинга и дистилляции

- Баланс между эффективностью и качеством

9. Мультимодальное обучение

- Реализация кросс-модального поиска

- Работа с совместными embedding-пространствами

10. Оформление научной публикации

- Структурирование результатов исследований

- Подготовка материалов в формате научной статьи

Методическое обеспечение:

- Вычислительные ресурсы: GPU-серверы, Google Colab Pro, облачные платфор-

мы

- Программное обеспечение: PyTorch, TensorFlow, Hugging Face, NetworkX

- Данные: открытые датасеты (MNIST, CIFAR, COCO, Flickr), специализированные исследовательские наборы данных

- Документация: подробные методические указания к каждой лабораторной работе, шаблоны отчетов, примеры кода

Наименование Лабораторной работы	Содержание Лабораторной работы	Распределение часов
2	3	
Формулировка исследовательской гипотезы	<p>Анализ современных тенденций в исследованиях ИИ</p> <p>Поиск и изучение научных статей на arXiv и в рецензируемых журналах</p> <p>Методология формулировки проверяемых научных гипотез</p> <p>Планирование эксперимента и определение метрик оценки</p> <p>Оформление исследовательского предложения</p>	6,8 часа (2 часа ЛР, 4,8 час СР)
Трансформеры	<p>Изучение архитектуры Transformer и механизма внимания</p> <p>Реализация scaled dot-product attention</p> <p>Построение multi-head attention механизма</p> <p>Создание полного трансформер-блока</p> <p>Визуализация матриц внимания и анализ результатов</p>	14,5 часов (8 часов ЛР, 6,5 часов СР)
Модификация архитектуры, PPO и основы RLHF	<p>Анализ ограничений стандартных трансформеров</p> <p>Реализация sparse attention с локальными окнами</p> <p>Построение linear attention через kernel approximation</p>	13,2 часа (5 часов ЛР, 8,2 часа СР)

Наименование Лабораторной работы	Содержание Лабораторной работы	Распределение часов
2	3	
	<p>Сравнительный анализ эффективности модификаций</p> <p>Оптимизация производительности для длинных последовательностей</p> <p>Основы обучения с подкреплением (Reinforcement Learning)</p> <p>Реализация алгоритма PPO (Proximal Policy Optimization)</p> <p>Сбор и обработка человеческих предпочтений</p> <p>Построение reward model для RLHF</p> <p>Fine-tuning политики с учетом человеческого feedback</p>	
Диффузионные модели, Контроль генерации	<p>Теория диффузионных процессов в генеративных моделях</p> <p>Реализация forward diffusion process</p> <p>Построение U-Net архитектуры с временными embedding</p> <p>Обучение модели с noise prediction loss</p> <p>Генерация samples с использованием DDPM/DDIM</p> <p>Методы управления выходом генеративных моделей</p> <p>Реализация classifier-free guidance</p> <p>Настройка контроля стиля и содержания</p> <p>Количественная оценка управляемости генерации</p> <p>Анализ trade-off между качеством и разнообразием</p>	16,5 часа (7 часов ЛР, 9,5 часов СР)
Нейро-символьная интеграция	<p>Принципы нейро-символьного искусственного интеллекта</p> <p>Извлечение сущностей и отношений из текста</p> <p>Построение графа знаний с использованием NetworkX</p> <p>Реализация логического вывода на графах</p> <p>Визуализация цепочек рассуждений</p>	12 часов (4 часа ЛР, 8 часов СР)
Оптимизация моделей	<p>Методы сжатия нейронных сетей</p> <p>Применение квантования (quantization)</p> <p>Техники прунинга (pruning) весов</p> <p>Дистилляция знаний (knowledge distillation)</p> <p>Сравнительный анализ эффективности методов</p>	9,2 часа (4 часа ЛР, 5,2 часа СР)
Мультимодальное обучение	<p>Основы мультимодального представления данных</p> <p>Реализация dual-encoder архитектуры</p> <p>Обучение с contrastive loss (InfoNCE)</p> <p>Кросс-модальный поиск и оценка качества</p> <p>Анализ совместного embedding-пространства</p>	9,2 часа (2 часа ЛР, 7,2 часа СР)
Оформление научной статьи	<p>Структура научной статьи в области ИИ</p> <p>Подготовка результатов экспериментов</p> <p>Визуализация данных и метрик</p>	8,4 часа (2 часа ЛР, 6,4 час СР)

Наименование Лабораторной работы	Содержание Лабораторной работы	Распределение часов
2	3	
	Формулировка выводов и ограничений Оформление в соответствии с требованиями конференций	

2) Разработка заданий:

- Пошаговые инструкции.
- Примеры кода.

Контрольные вопросы.

Разработка заданий для лабораторной работы «Нейро-символьная интеграция»

Цель:

Освоить принципы нейро-символьного искусственного интеллекта и их применение для построения интеллектуальных систем, сочетающих машинное обучение и логический вывод.

Пошаговые инструкции

Шаг 1: Установка библиотек

```
python
!pip install torch transformers spacy networkx matplotlib
!python -m spacy download en_core_web_sm
```

Шаг 2: Загрузка и подготовка данных

```
python
import pandas as pd
import spacy
```

Загрузка датасета медицинских текстов

```
data = {
    'text': [
        'Patient has fever and headache. Diagnosed with influenza.',
        'Patient presents with cough and chest pain. Suspected pneumonia.',
        'Patient reports fatigue and nausea. Possible food poisoning.'
    ]
}
df = pd.DataFrame(data)
```

Шаг 3: Реализация системы извлечения сущностей

```
python
import torch
import transformers
from transformers import AutoTokenizer, AutoModelForTokenClassification
import networkx as nx
import matplotlib.pyplot as plt
```

Загрузка модели для NER

```
nlp = spacy.load("en_core_web_sm")
```

```
def extract_medical_entities(text):
```

```
    doc = nlp(text)
    entities = []
    for ent in doc.ents:
        if ent.label_ in ['DISEASE', 'SYMPTOM']:
            entities.append((ent.text, ent.label_))
    return entities
```

```

# Тестирование извлечения сущностей
sample_text = "Patient has fever and headache. Diagnosed with influenza."
entities = extract_medical_entities(sample_text)
print("Извлеченные сущности:", entities)
Шаг 4: Построение графа знаний
python
class MedicalKnowledgeGraph:
    def __init__(self):
        self.graph = nx.DiGraph()

    def add_entities(self, text):
        entities = extract_medical_entities(text)
        for i, (entity1, label1) in enumerate(entities):
            self.graph.add_node(entity1, label=label1)
        for j, (entity2, label2) in enumerate(entities):
            if i != j:
                self.graph.add_edge(entity1, entity2, relationship='associated_with')

    def visualize(self):
        plt.figure(figsize=(10, 8))
        pos = nx.spring_layout(self.graph)
        nx.draw(self.graph, pos, with_labels=True, node_color='lightblue',
                font_size=10, node_size=2000)
        plt.title("Медицинский граф знаний")
        plt.show()

```

```

# Создание и визуализация графа

```

```

kg = MedicalKnowledgeGraph()
for text in df['text']:
    kg.add_entities(text)
kg.visualize()

```

Шаг 5: Реализация логического вывода

```

python
def diagnose_symptoms(symptoms, knowledge_graph):
    possible_diseases = []
    for symptom in symptoms:
        for node in knowledge_graph.graph.nodes:
            if knowledge_graph.graph.nodes[node]['label'] == 'DISEASE':
                if nx.has_path(knowledge_graph.graph, symptom, node):
                    possible_diseases.append(node)
    return list(set(possible_diseases))

```

```

# Пример диагностики

```

```

symptoms = ['fever', 'headache']
diagnoses = diagnose_symptoms(symptoms, kg)
print("Возможные диагнозы:", diagnoses)

```

Шаг 6: Интеграция с трансформерами

```

python
from transformers import pipeline

```

```

# Использование модели для классификации медицинских утверждений

```

```
classifier = pipeline("text-classification",
                      model="bvanaken/clinical-assertion-negation-bert")
```

```
def analyze_medical_statement(text):
    result = classifier(text)
    return result[0]
```

Тестирование

```
statement = "Patient has no fever."
analysis = analyze_medical_statement(statement)
print(f"Анализ утверждения: {analysis}")
```

Полный скрипт

```
python
```

1. Установка библиотек

```
!pip install torch transformers spacy networkx matplotlib
!python -m spacy download en_core_web_sm
```

2. Импорт библиотек

```
import pandas as pd
import spacy
import networkx as nx
import matplotlib.pyplot as plt
from transformers import pipeline
```

3. Загрузка данных

```
data = {
    'text': [
        'Patient has fever and headache. Diagnosed with influenza.',
        'Patient presents with cough and chest pain. Suspected pneumonia.',
        'Patient reports fatigue and nausea. Possible food poisoning.',
        'Patient with high temperature and sore throat. Diagnosed with strep throat.',
        'Patient complains of abdominal pain and vomiting. Suspected gastroenteritis.'
    ]
}
df = pd.DataFrame(data)
```

4. Инициализация моделей

```
nlp = spacy.load("en_core_web_sm")
classifier = pipeline("text-classification",
                      model="bvanaken/clinical-assertion-negation-bert")
```

5. Определение классов

```
class MedicalKnowledgeGraph:
    def __init__(self):
        self.graph = nx.DiGraph()

    def extract_entities(self, text):
        doc = nlp(text)
        entities = []
        for ent in doc.ents:
            if ent.label_ in ['DISEASE', 'SYMPTOM']:
```

```

        entities.append((ent.text, ent.label_))
    return entities

def add_entities(self, text):
    entities = self.extract_entities(text)
    for i, (entity1, label1) in enumerate(entities):
        self.graph.add_node(entity1, label=label1)
    for j, (entity2, label2) in enumerate(entities):
        if i != j:
            self.graph.add_edge(entity1, entity2, relationship='associated_with')

def diagnose(self, symptoms):
    possible_diseases = []
    for symptom in symptoms:
        for node in self.graph.nodes:
            if self.graph.nodes[node]['label'] == 'DISEASE':
                if nx.has_path(self.graph, symptom, node):
                    possible_diseases.append(node)
    return list(set(possible_diseases))

def visualize(self):
    plt.figure(figsize=(12, 10))
    pos = nx.spring_layout(self.graph)
    nx.draw(self.graph, pos, with_labels=True, node_color='lightgreen',
            font_size=10, node_size=2500, edge_color='gray')
    plt.title("Медицинский граф знаний", size=16)
    plt.show()

# 6. Основной цикл работы
kg = MedicalKnowledgeGraph()

# Построение графа знаний
for text in df['text']:
    kg.add_entities(text)

# Визуализация графа
kg.visualize()

# Диагностика по симптомам
symptoms = ['fever', 'headache']
diagnoses = kg.diagnose(symptoms)
print(f"При симптомах {symptoms} возможные диагнозы: {diagnoses}")

# Анализ медицинских утверждений
test_statements = [
    "Patient has fever",
    "Patient has no fever",
    "Patient denies headache"
]

for statement in test_statements:
    result = classifier(statement)

```

```
print(f"Утверждение: '{statement}' -> {result[0]}")
```

Контрольные вопросы

1. Теоретические:

- В чем отличие нейро-символьного ИИ от чистого глубокого обучения?
- Какие преимущества дает комбинация нейросетей и символьных методов?
- Что такое граф знаний и как он используется в интеллектуальных системах?

2. Практические:

- Напишите код для извлечения новых типов сущностей из медицинских текстов
- Реализуйте функцию для поиска всех путей между симптомами и заболеваниями в графе

- Модифицируйте систему для работы с другими предметными областями

3. Аналитические:

- Какие ограничения имеет текущая система извлечения сущностей?
- Как можно улучшить точность диагностики используя дополнительные данные?

Критерии оценки

Отлично:

- Полная реализация всех компонентов системы
- Корректное извлечение сущностей и построение графа
- Глубокий анализ результатов и предложения по улучшению

Хорошо:

- Рабочая система с небольшими недочетами
- Корректная визуализация графа знаний
- Базовый анализ результатов

Удовлетворительно:

- Частичная реализация основных функций
- Ошибки в логическом выводе или визуализации
- Поверхностный анализ

Неудовлетворительно:

- Невыполнение ключевых этапов работы
- Критические ошибки в реализации
- Отсутствие анализа результатов

Подготовка датасетов:

- Использование открытых медицинских датасетов (MIMIC-III, i2b2)
- Генерация синтетических медицинских случаев при необходимости
- Адаптация датасетов под конкретную предметную область

Порядок проверки:

1. Выполнение заданий:

- Код запускается без ошибок
- Граф знаний строится корректно
- Логический вывод работает адекватно

2. Качество кода:

- Соблюдение PEP-8
- Наличие комментариев и документации
- Модульная структура

3. Отчет:

- Описание архитектуры системы
- Анализ извлеченных сущностей и отношений
- Визуализация результатов

4. Своевременность:

- Сдача работы в установленный срок
- Соответствие требованиям оформления

4.5. Методические указания по организации проектной деятельности студентов

Цели, задачи и ожидаемые результаты

Цели организации проектной деятельности:

- Сформировать у студентов комплексное представление о применении современных методов ИИ для решения реальных исследовательских и прикладных задач
- Развить навыки самостоятельного проведения научных исследований и разработки инновационных решений в области искусственного интеллекта
- Подготовить студентов к работе над сложными междисциплинарными проектами на стыке науки и индустрии

Задачи преподавателя:

- Организация доступа к современным вычислительным ресурсам и исследовательским инструментам
- Курирование процесса формулировки и проверки научных гипотез
- Обеспечение интеграции теоретических знаний и практических навыков
- Формирование системы оценки проектов с учетом научной новизны и практической значимости

Ожидаемые результаты студентов:

- Умение формулировать и проверять исследовательские гипотезы в области фронтальных исследований ИИ
- Навыки работы с современными архитектурами нейронных сетей и методами оптимизации
- Опыт оформления результатов исследований в виде научных публикаций и технических отчетов
- Способность критически оценивать и сравнивать различные подходы к решению сложных задач

Порядок реализации

Задача №1: Формирование исследовательских тем на основе актуальных направлений ИИ

1. Разработка эффективных архитектур для обработки длинных последовательностей

Описание: Исследование методов оптимизации трансформеров для работы с длинными текстовыми последовательностями и временными рядами. Разработка модифицированных архитектур с sparse attention и линейной сложностью.

Цель: Создание эффективных моделей для обработки документов большого объема, геномных данных и финансовых временных рядов.

Ожидаемые результаты:

- Реализация модифицированных архитектур трансформеров
- Сравнительный анализ эффективности на различных типах данных
- Документация с рекомендациями по применению

2. Генеративные модели для научных исследований

Описание: Разработка систем контролируемой генерации для создания синтетических данных в научных исследованиях. Применение диффузионных моделей и методов контроля генерации.

Цель: Автоматизация процесса создания тренировочных данных и исследование свойств генеративных моделей.

Ожидаемые результаты:

- Реализация системы контролируемой генерации
- Метрики оценки качества и разнообразия генерируемых данных
- Анализ применимости в различных научных domain

Задача №2: Кейсы от научных лабораторий и исследовательских центров

1. Нейро-символьные системы для анализа научных публикаций

Постановка задачи: Научная лаборатория нуждается в системе автоматического анализа и структурирования исследовательских статей. Необходимо построить граф знаний из научных публикаций.

Требования:

- Реализовать извлечение ключевых концепций и отношений из текстов статей
- Построить семантическую сеть связанных исследований
- Разработать механизм логического вывода для поиска research gaps

2. Мультимодальные системы для анализа экспериментальных данных

Постановка задачи: Исследовательский центр собирает разнородные данные экспериментов (изображения, текстовые описания, числовые измерения). Необходимо создать систему кросс-модального анализа.

Требования:

- Реализовать совместное embedding-пространство для разных модальностей
- Обеспечить возможность кросс-модального поиска и сравнения
- Разработать методы визуализации мультимодальных данных

Задача №3: Формирование ТЗ на зачетный проект

1. Исследовательский проект: Сравнительный анализ методов оптимизации больших моделей

Используя методы квантования, прунинга и дистилляции, проведите сравнительный анализ эффективности различных подходов к оптимизации больших языковых моделей.

Требования к проекту:

- Проект выполняется в командах от 2 до 4 человек
- Длительность проекта: 4-6 недель
- Объем итогового отчета: 8-12 страниц

Этапы выполнения:

1. Подготовительный этап:

- Анализ существующих методов оптимизации
- Формулировка исследовательских гипотез
- Планирование эксперимента

2. Экспериментальный этап:

- Реализация методов оптимизации
- Проведение сравнительных экспериментов
- Сбор и обработка результатов

3. Аналитический этап:

- Статистический анализ полученных данных
- Формулировка выводов и рекомендаций
- Подготовка финального отчета

Критерии оценки:

Отлично:

- Полное выполнение всех этапов проекта
- Глубокий анализ результатов и статистическая значимость
- Научная новизна и практическая значимость работы
- Качественное оформление отчета и кода

Хорошо:

- Выполнение основных этапов проекта
- Корректный анализ результатов с незначительными недочетами
- Соответствие работы заявленным требованиям

Удовлетворительно:

- Частичное выполнение проекта
- Поверхностный анализ результатов
- Наличие существенных недочетов в реализации

Неудовлетворительно:

- Невыполнение ключевых этапов проекта
- Критические ошибки в методологии или реализации
- Несоответствие работы базовым требованиям

Методические рекомендации по организации работы

1. Формирование проектных групп:

- Распределение ролей в команде (research lead, implementation specialist, data analyst)
- Регулярные встречи с научным руководителем (не реже 1 раза в неделю)
- Ведение проектной документации и протоколов встреч

2. Доступ к ресурсам:

- Организация доступа к GPU-кластерам и облачным платформам
- Предоставление лицензий на специализированное ПО
- Обеспечение доступа к научным базам данных и публикациям

3. Система контроля качества:

- Code review и проверка воспроизводимости результатов
- Экспертная оценка научной составляющей проекта
- Презентация результатов на внутреннем семинаре

4. Шкала оценки:

- Научная ценность и новизна (40%)
- Техническая реализация и качество кода (30%)
- Оформление результатов и презентация (20%)
- Работа в команде и соблюдение сроков (10%)

Требования для повышения оценки:

- Публикация результатов в виде препринта или статьи
- Участие в научных конкурсах или хакатонах
- Разработка дополнительного функционала сверх требований
- Глубокий сравнительный анализ с state-of-the-art методами

Итоговая оценка формируется преподавателем на основе:

- Качества выполнения всех этапов проекта
- Глубины проработки теоретической основы
- Практической значимости полученных результатов
- Умения защищать и презентовать свою работу
- **Порядок проверки корректности**
- *Чек-лист для проверки лабораторных работ:*
- Набор кейсов индустриальных партнеров – 20 шт;
- Набор кейсов преподавателей практиков и лабораторий ВУЗа – 10 шт;

5. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

5.1 Учебная литература

Основная литература:

1. Бессмертный, И. А. Системы искусственного интеллекта : учебник для вузов / И. А. Бессмертный. — 3-е изд., испр. и доп. — Москва : Издательство Юрайт, 2025. — 164 с. — (Высшее образование). — ISBN 978-5-534-18416-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/561602> (дата обращения: 11.08.2025).

2. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для вузов / Л. А. Станкевич. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 478 с. — (Высшее образование). — ISBN 978-5-534-20363-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/560754> (дата обращения: 11.08.2025).

3. Бессмертный, И. А. Системы искусственного интеллекта : учебник для среднего профессионального образования / И. А. Бессмертный. — 3-е изд., испр. и доп. — Москва : Издательство Юрайт, 2025. — 163 с. — (Профессиональное образование). — ISBN 978-5-534-18417-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/565036> (дата обращения: 11.08.2025).

4. Рабчевский, А. Н. Синтетические данные и развитие нейросетевых технологий : учебник для вузов / А. Н. Рабчевский. — Москва : Издательство Юрайт, 2025. — 187 с. — (Высшее образование). — ISBN 978-5-534-17716-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/568661> (дата обращения: 11.08.2025).

5.2 Дополнительная литература:

1. Sun, X., Li, J., Kovalenko, A.V., Feng, W., Ou, Y. Integrating Reinforcement Learning and Learning From Demonstrations to Learn Nonprehensile Manipulation //IEEE Transactions on Automation Science and Engineering, 2023, 20(3), 1735–1744, DOI: 10.1109/TASE.2022.3185071, Q1

2. Petukhova, A.V.; Kovalenko, A.V.; Ovsyannikova, A.V. Algorithm for Optimization of Inverse Problem Modeling in Fuzzy Cognitive Maps. Mathematics 2022, 10, 3452. DOI: 10.3390/math10193452, Q1

3. Kirillova, E.; Kovalenko, A.; Urtenov, M. Study of the Current–Voltage Characteristics of Membrane Systems Using Neural Networks. AppliedMath 2025, 5, 10. <https://doi.org/10.3390/appliedmath5010010>.

4. Kadurin, Artur, et al. "The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology." Oncotarget 8.7 (2016): 10883.

5. Kadurin, Artur, et al. "druGAN: an advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico." Molecular pharmaceutics 14.9 (2017): 3098-3104.

6. Polykovskiy, Daniil, et al. "Molecular sets (MOSES): a benchmarking platform for molecular generation models." Frontiers in pharmacology 11 (2020): 565644.

7. Khrabrov, Kuzma, et al. " ∇^2 DFT: A Universal Quantum Chemistry Dataset of Drug-Like Molecules and a Benchmark for Neural Network Potentials." Advances in Neural Information Processing Systems 37 (2024): 36869-36889.

8. Polykovskiy, Daniil, et al. "Entangled conditional adversarial autoencoder for de novo drug discovery." Molecular pharmaceutics 15.10 (2018): 4398-4405.

9. Николенко, Сергей, Кадури, Артур и Архангельская Екатерина. Глубокое обучение. "Издательский дом"" Питер""", 2017.

10. Рабчевский, А. Н. Синтетические данные и развитие нейросетевых технологий : учебное пособие для вузов / А. Н. Рабчевский. — Москва : Издательство Юрайт, 2024. — 187 с. — (Высшее образование). — ISBN 978-5-534-17716-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/545036> (дата обращения: 19.07.2025).

11. Елисеев А. И., Минин Ю. В. Разработка программных интерфейсов веб-приложений с использованием фреймворка FastAPI : учебное пособие. Тамбов: ТГТУ, 2024. 81 с. <https://e.lanbook.com/book/472310> (дата обращения: 19.07.2025).

12. Чернышев, С. А. Основы программирования на Python : учебное пособие для вузов / С. А. Чернышев. — Москва : Издательство Юрайт, 2022. — 286 с. — (Высшее образование). — ISBN 978-5-534-14350-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/496893>. (дата обращения: 19.07.2025).

13. Златопольский Д. М. Основы программирования на языке Python. 2-е изд. Москва: ДМК Пресс, 2018.

5.3. Периодические издания и конференции:

1. Базы данных компании «Ист Вью» <http://dlib.eastview.com>
2. Электронная библиотека GREBENNIKON.RU <https://grebennikon.ru/>

Конференции А*:

3. <https://openreview.net/forum?id=FMMF1a9ifL>
4. <https://openreview.net/forum?id=EIUrNM9U8c#discussion>
5. <https://openreview.net/forum?id=JoO6mtCLHD>
6. <https://aclanthology.org/2024.findings-emnlp.760/>
7. <https://aclanthology.org/2020.coling-main.588/>
8. https://link.springer.com/chapter/10.1007/978-3-030-72113-8_30
9. https://link.springer.com/chapter/10.1007/978-3-031-42448-9_10
10. <https://aclanthology.org/2024.findings-naacl.288/>

5.4. Интернет-ресурсы, в том числе современные профессиональные базы данных и информационные справочные системы

Электронно-библиотечные системы (ЭБС):

1. ЭБС «ЮРАЙТ» <https://urait.ru/>
2. ЭБС «УНИВЕРСИТЕТСКАЯ БИБЛИОТЕКА ОНЛАЙН» <http://www.biblioclub.ru/>
3. ЭБС «BOOK.ru» <https://www.book.ru>
4. ЭБС «ZNANIUM.COM» www.znanium.com
5. ЭБС «ЛАНЬ» <https://e.lanbook.com>

Профессиональные базы данных

1. Scopus <http://www.scopus.com/>
2. ScienceDirect <https://www.sciencedirect.com/>
3. Журналы издательства Wiley <https://onlinelibrary.wiley.com/>
4. Научная электронная библиотека (НЭБ) <http://www.elibrary.ru/>
5. Полнотекстовые архивы ведущих западных научных журналов на Российской платформе научных журналов НЭИКОН <http://archive.neicon.ru>
6. Национальная электронная библиотека (доступ к Электронной библиотеке диссертаций Российской государственной библиотеки (РГБ) <https://rusneb.ru/>
7. Президентская библиотека им. Б.Н. Ельцина <https://www.prlib.ru/>
8. База данных CSD Кембриджского центра кристаллографических данных (CCDC) <https://www.ccdc.cam.ac.uk/structures/>
9. Springer Journals: <https://link.springer.com/>
10. Springer Journals Archive: <https://link.springer.com/>
11. Nature Journals: <https://www.nature.com/>
12. Springer Nature **Protocols and Methods**: <https://experiments.springernature.com/sources/springer-protocols>
13. Springer Materials: <http://materials.springer.com/>
14. Nano Database: <https://nano.nature.com/>
15. Springer eBooks (i.e. 2020 eBook collections): <https://link.springer.com/>
16. "Лекториум ТВ" <http://www.lektorium.tv/>
17. Университетская информационная система РОССИЯ <http://uisrussia.msu.ru>

Бесплатные образовательные ресурсы

1. Jupyter Notebook – интерактивные вычисления
2. Visual Studio Code – редактор кода с поддержкой Python
3. Google Scholar/arXiv – доступ к научным публикациям

Ресурсы свободного доступа

1. Официальная документация по TensorFlow <https://www.tensorflow.org/?hl=ru>
2. Официальная документация по Keras <https://www.tensorflow.org/guide/keras?hl=ru>
3. Официальная документация по pyTorch <https://docs.pytorch.org/docs/stable/index.html>
4. КиберЛенинка <http://cyberleninka.ru/>;
5. Американская патентная база данных <http://www.uspto.gov/patft/>
6. Министерство науки и высшего образования Российской Федерации <https://www.minobrnauki.gov.ru/>;
7. Федеральный портал "Российское образование" <http://www.edu.ru/>;
8. Информационная система "Единое окно доступа к образовательным ресурсам" <http://window.edu.ru/>;
9. Единая коллекция цифровых образовательных ресурсов <http://school-collection.edu.ru/> .
10. Проект Государственного института русского языка имени А.С. Пушкина "Образование на русском" <https://pushkininstitute.ru/>;
11. Справочно-информационный портал "Русский язык" <http://gramota.ru/>;
12. Служба тематических толковых словарей <http://www.glossary.ru/>;
13. Словари и энциклопедии <http://dic.academic.ru/>;
14. Образовательный портал "Учеба" <http://www.ucheba.com/>;
15. Законопроект "Об образовании в Российской Федерации". Вопросы и ответы http://xn--273--84d1f.xn--plai/voprosy_i_otvety

Собственные электронные образовательные и информационные ресурсы КубГУ

1. Электронный каталог Научной библиотеки КубГУ <http://megapro.kubsu.ru/MegaPro/Web>
2. Электронная библиотека трудов ученых КубГУ <http://megapro.kubsu.ru/MegaPro/UserEntry?Action=ToDb&idb=6>
3. Среда модульного динамического обучения <http://moodle.kubsu.ru>
4. База учебных планов, учебно-методических комплексов, публикаций и конференций <http://infoneeds.kubsu.ru/>
5. Библиотека информационных ресурсов кафедры информационных образовательных технологий [http://mschool.kubsu.ru/](http://mschool.kubsu.ru;)
6. Электронный архив документов КубГУ <http://docspace.kubsu.ru/>
7. Электронные образовательные ресурсы кафедры информационных систем и технологий в образовании КубГУ и научно-методического журнала "ШКОЛЬНЫЕ ГОДЫ" <http://icdau.kubsu.ru/>

6. Методические указания для обучающихся по освоению дисциплины (модуля)

Учебная деятельность проходит в соответствии с графиком учебного процесса. Процесс самостоятельной работы контролируется во время аудиторных занятий и индивидуальных консультаций. Самостоятельная работа студентов проводится в форме изучения отдельных теоретических вопросов по предлагаемой литературе и выполнении практических заданий по разобранным во время аудиторных занятий примерам.

Фонд оценочных средств дисциплины состоит из средств текущего контроля (см. список лабораторных работ, задач и вопросов) и итоговой аттестации (зачета, экзамена).

В качестве оценочных средств, используемых для текущего контроля успеваемости, предлагается перечень вопросов, которые прорабатываются в процессе освоения кур-

са. Данный перечень охватывает все основные разделы курса, включая знания, получаемые во время самостоятельной работы. Кроме того, важным элементом технологии является самостоятельное решение студентами и сдача заданий. Это полностью индивидуальная форма обучения. Студент рассказывает свое решение преподавателю, отвечает на дополнительные вопросы.

При самостоятельной работе студентам необходимо изучать рекомендованную литературу (учебники, научные статьи, документацию библиотек) для полного понимания теоретических основ и современных подходов в области глубокого обучения нейросетей. Выполняя проектные задания, студент должен уметь: определять с помощью какой архитектуры можно решить поставленную задачу, как дообучать и настраивать глубокую архитектуру под свою задачу, как оптимизировать процесс обучения, а также как оценивать качество работы модели с помощью соответствующих метрик.

Важнейшим компонентом курса является самостоятельная проектная работа, в ходе которой студент разрабатывает законченное CV-приложение. Такой проект позволяет закрепить навыки проектирования и реализации комплексных решений в области компьютерного зрения.

Для студентов с ограниченными возможностями здоровья предусмотрены индивидуальные консультации и адаптированные материалы. Преподаватель помогает осваивать интерфейсы взаимодействия с ИИ, объясняет ключевые понятия в доступной форме, предоставляет инструкции с альтернативным форматированием. При необходимости используются голосовые интерфейсы, увеличенный масштаб экрана, сопровождение при выполнении заданий. Индивидуальный подход обеспечивает равные условия участия в образовательном процессе и достижения запланированных результатов обучения.

Подход, определяющий установление соответствия кейсов ИП и УГТ (5-7), позволяет четко соотносить этапы развития технологии с вовлеченностью партнера и снижать риски при переходе от лабораторных испытаний к промышленному внедрению.

Кейс 1. Мультимодальный анализ клиентского опыта

Описание: Используя данные из различных источников (текстовые обращения в кол-центр, аудиозаписи разговоров, метрики использования мобильного приложения), необходимо разработать систему комплексной оценки клиентского опыта на основе современных мультимодальных архитектур.

Цель: Создание единой системы оценки удовлетворенности клиентов через интеграцию разрозненных данных и выявление скрытых паттернов недовольства.

Ожидаемые результаты:

- Мультимодальная модель анализа клиентского опыта с точностью >85%
- Система раннего предупреждения о проблемах в клиентском сервисе
- Рекомендации по улучшению качества обслуживания

Кейс 2. Генеративная система для персонального финансового планирования

Описание: Разработка диффузионной модели для генерации персонализированных финансовых планов на основе данных о доходах, расходах, финансовых целях и макроэкономических прогнозах.

Цель: Автоматизация создания индивидуальных финансовых стратегий с учетом уникальных особенностей каждого клиента.

Ожидаемые результаты:

- Генеративная модель персонального финансового планирования
- Система адаптации планов при изменении жизненных обстоятельств
- Увеличение вовлеченности в использование финансовых продуктов на 30%

Кейс 3. Нейро-символьная система для compliance-контроля

Описание: Создание гибридной системы для автоматического мониторинга соблюдения регуляторных требований, сочетающей глубокое обучение для анализа транзакций и символьные правила законодательства.

Цель: Снижение регуляторных рисков и автоматизация процессов compliance-контроля.

Ожидаемые результаты:

- Система автоматического выявления нарушений регуляторных требований
- Сокращение ложных срабатываний на 40%
- Уменьшение затрат на compliance на 25%

Кейс 4. Предсказание рыночных шоков с помощью временных рядов

Описание: Разработка модели предсказания экстремальных рыночных событий на основе анализа многомерных временных рядов с использованием модифицированных архитектур трансформеров.

Цель: Улучшение управления рыночными рисками и разработка превентивных стратегий хеджирования.

Ожидаемые результаты:

- Модель предсказания рыночных шоков с заблаговременностью 7-10 дней
- Система стресс-тестирования инвестиционного портфеля
- Снижение потерь от экстремальных событий на 20%

Кейс 5. RL-оптимизация кросс-канальных коммуникаций

Описание: Создание системы оптимизации кросс-канальных коммуникаций с клиентами с использованием обучения с подкреплением, учитывающей историю взаимодействий и индивидуальные предпочтения.

Цель: Максимизация эффективности маркетинговых коммуникаций и минимизация раздражения клиентов.

Ожидаемые результаты:

- Система персонализации каналов и времени коммуникаций
- Увеличение CTR на 35%
- Снижение отказов от рассылок на 50%

Кейс 6. Трансформеры для анализа корпоративных документов

Описание: Разработка системы автоматического анализа корпоративных отчетов и новостей с использованием больших языковых моделей для оценки кредитных рисков компаний.

Цель: Улучшение качества кредитного анализа корпоративных клиентов за счет обработки неструктурированных данных.

Ожидаемые результаты:

- Модель анализа корпоративных документов
- Система раннего предупреждения о проблемах заемщиков
- Повышение точности кредитного скоринга на 15%

Кейс 7. Диффузионные модели для синтеза финансовых данных

Описание: Создание системы генерации синтетических финансовых данных для тестирования моделей и алгоритмов без использования реальных персональных данных клиентов.

Цель: Разработка безопасного инструмента для исследований и тестирования с сохранением статистических свойств реальных данных.

Ожидаемые результаты:

- Генеративная модель синтетических финансовых данных
- Сохранение статистических свойств и корреляций

- Ускорение разработки новых продуктов на 40%

Кейс 8. Многоагентные системы для торговых операций

Описание: Разработка многоагентной системы для автоматизированной торговли на финансовых рынках, где агенты специализируются на разных стратегиях и активах.

Цель: Создание адаптивной торговой системы, способной эффективно работать в различных рыночных условиях.

Ожидаемые результаты:

- Многоагентная торговая система
- Адаптация к изменяющимся рыночным условиям
- Увеличение доходности на 12% при снижении риска

Кейс 9. Кросс-модальный поиск финансовых услуг

Описание: Создание системы семантического поиска финансовых продуктов по естественно-языковым запросам с использованием совместных эмбедингов текста и структурированных данных продуктов.

Цель: Упрощение процесса поиска и подбора финансовых продуктов для клиентов.

Ожидаемые результаты:

- Система естественно-языкового поиска продуктов
- Увеличение конверсии на 25%
- Снижение нагрузки на контакт-центр на 30%

Кейс 10. Графовые нейросети для выявления мошеннических сетей

Описание: Разработка системы обнаружения сложных мошеннических схем на основе анализа транзакционных графов с использованием графовых нейронных сетей.

Цель: Выявление скоординированных мошеннических действий, которые невозможно обнаружить традиционными методами.

Ожидаемые результаты:

- Модель анализа транзакционных графов
- Система обнаружения мошеннических сетей
- Снижение потерь от мошенничества на 35%

Критерии оценки для всех кейсов:

Научная новизна (30%):

- Применение современных архитектур и методов
- Инновационный подход к решению задачи
- Потенциал для публикации результатов

Техническая реализация (40%):

- Качество кода и документации
- Эффективность и масштабируемость решения
- Соответствие промышленным стандартам

Бизнес-impact (30%):

- Потенциальный экономический эффект
- Возможность интеграции в существующие процессы
- Соответствие стратегическим целям Сбербанка

Дополнительные требования:

- Обеспечение конфиденциальности данных
- Соответствие регуляторным требованиям
- Возможность A/B тестирования
- Документация по внедрению и эксплуатации

Каждый кейс предполагает не только разработку прототипа, но и:

- Анализ экономической эффективности
- План внедрения и масштабирования
- Оценку рисков и ограничений
- Рекомендации по дальнейшему развитию

Кейсы от «АВАЛАБ»

Кейс 1. Прогнозирование востребованности IT-специализаций

Описание: Компания «АВАЛАБ» хочет предсказывать востребованность различных IT-специализаций на рынке труда. Задача — разработать модель, которая будет анализировать данные о вакансиях, технологических трендах, запросах компаний-партнеров и глобальные экономические показатели.

Цель: Создать модель, которая сможет прогнозировать спрос на IT-специализации с горизонтом планирования 6-12 месяцев.

Ожидаемый результат: Глубокая нейронная сеть, способная анализировать многомерные временные ряды и выдавать точные прогнозы по востребованности IT-направлений.

Кейс 2. Персонализация треков обучения

Описание: Необходимо разработать систему адаптивного формирования индивидуальных образовательных траекторий для пользователей платформы. Модель должна анализировать текущий уровень знаний, скорость обучения, предпочтения и карьерные цели студентов.

Цель: Создать AI-систему, которая автоматически подбирает оптимальную последовательность курсов и практических заданий для каждого пользователя.

Ожидаемый результат: Рекуррентная нейросеть с механизмом внимания, генерирующая персонализированные учебные планы с учетом динамики прогресса студентов.

Кейс 3. Оптимизация процесса проверки заданий

Описание: Разработка системы автоматической проверки и оценки качества выполнения программистских заданий. Модель должна анализировать код, тестировать его функциональность и оценивать соответствие best practices.

Цель: Создать интеллектуальную систему проверки заданий, снижающую нагрузку на ревьюеров и обеспечивающую объективную оценку.

Ожидаемый результат: Мультимодальная нейросеть, сочетающая анализ синтаксиса, семантики кода и результатов автоматического тестирования.

Кейс 4. Предсказание успешности трудоустройства

Описание: Компания хочет прогнозировать вероятность успешного трудоустройства выпускников на основе их прогресса в обучении, активности на платформе и текущей ситуации на рынке труда.

Цель: Разработать модель, идентифицирующую факторы, наиболее влияющие на успешное трудоустройство, и прогнозирующую вероятность найма.

Ожидаемый результат: Ансамблевая модель машинного обучения, объединяющая градиентный бустинг и нейросети для точного предсказания карьерных outcomes.

Кейс 5. Динамическое ценообразование образовательных программ

Описание: Создание системы динамического ценообразования для образовательных продуктов «АВАЛАБ». Модель должна учитывать спрос, сезонность, конкурентную среду и платежеспособность целевой аудитории.

Цель: Разработать AI-систему, оптимизирующую ценовую политику для максимизации доступа к образованию и прибыльности.

Ожидаемый результат: Система на основе обучения с подкреплением, автоматически adjusting цены в реальном времени.

Кейс 6. Генерация персональных учебных материалов

Описание: Разработка генеративной модели для создания персонализированных учебных материалов, адаптированных под стиль обучения и уровень подготовки каждого студента.

Цель: Создать систему автоматической генерации учебного контента, повышающую эффективность усвоения материала.

Ожидаемый результат: Diffusion-модель, генерирующая текстовые и визуальные учебные материалы с контролем сложности и стиля.

Кейс 7. Оптимизация расписания менторских сессий

Описание: Создание системы оптимизации расписания менторских сессий и групповых занятий с учетом временных зон, загрузки менторов и предпочтений студентов.

Цель: Разработать AI-алгоритм, минимизирующий конфликты расписания и максимизирующий доступность менторинга.

Ожидаемый результат: Нейросетевая модель с ограничениями, решающая задачу оптимизации расписания в условиях множества переменных.

Кейс 8. Выявление аномалий в учебном процессе

Описание: Разработка системы детектирования аномального поведения студентов (пропуски занятий, списывание, резкие изменения успеваемости) для своевременного вмешательства.

Цель: Создать систему раннего предупреждения проблем в учебном процессе.

Ожидаемый результат: Автоэнкодер с механизмом внимания, обнаруживающий аномалии в многомерных временных рядах учебной активности.

Кейс 9. Кластеризация студентов по стилям обучения

Описание: Разработка системы автоматической кластеризации студентов по стилям обучения и когнитивным характеристикам для формирования эффективных учебных групп.

Цель: Создать модель сегментации студентов, позволяющую оптимизировать образовательный процесс.

Ожидаемый результат: Гибридная модель кластеризации, сочетающая методы глубокого обучения и анализ временных рядов.

Кейс 10. Прогнозирование оттока студентов

Описание: Создание системы предсказания оттока студентов из образовательных программ на основе их активности, прогресса и внешних факторов.

Цель: Разработать модель для своевременного выявления студентов с риском прекращения обучения.

Ожидаемый результат: Рекуррентная нейросеть с механизмом внимания, прогнозирующая вероятность оттока с высокой точностью.

Критерии оценки для всех кейсов:

Техническая сложность (40%):

- Использование современных архитектур нейросетей
- Качество feature engineering и предобработки данных
- Оптимизация гиперпараметров модели

Практическая применимость (30%):

- Интегрируемость с платформой «АВАЛАБ»
- Измеримый бизнес-эффект
- Масштабируемость решения

Иновационность (30%):

- Новизна подхода к решению задачи
- Потенциал для научной публикации
- Использование frontier методов ИИ

Каждый кейс предполагает создание рабочего прототипа, включая:

- Подготовку и анализ данных
- Разработку и обучение модели
- Валидацию результатов
- Документацию и рекомендации по внедрению

7. Перечень информационных технологий, используемых при осуществлении образовательного процесса

7.1 Перечень информационных технологий.

1. Облачные платформы и сервисы для исследований ИИ

- **Google Colab** – облачная среда для выполнения кода на Python с бесплатным доступом к GPU/TPU ускорителям
- **Kaggle** – платформа для работы с датасетами и участия в соревнованиях по машинному обучению
- **Hugging Face Spaces** – развертывание демонстрационных версий моделей компьютерного зрения и NLP
- **Paperspace Gradient** – специализированная платформа для обучения нейросетей с мощными GPU
- **Yandex Cloud, SberCloud, AWS/GCP/Azure** – облачные вычисления для промышленного внедрения моделей

2. Системы управления версиями и коллаборации

- **Git/GitHub/GitLab** – распределенный контроль версий кода и совместная разработка
- **Weights & Biases** – трекинг экспериментов машинного обучения и визуализация метрик
- **MLflow** – платформа для управления жизненным циклом ML-моделей
- **DagsHub** – GitHub для проектов машинного обучения с интеграцией DVC и MLflow

3. Инструменты для работы с данными

- **Label Studio** – разметка мультимодальных датасетов (изображения, текст, аудио)
- **DVC (Data Version Control)** – управление версиями данных и экспериментов
- **Apache Spark** – распределенная обработка больших текстовых корпусов и графов
- **Snorkel** – создание размеченных датасетов с помощью программного кода
- **Airflow** – оркестрация ETL пайплайнов и ML workflows

4. Фреймворки глубокого обучения

- **PyTorch** – фреймворк для исследований с динамическими графами вычислений
- **TensorFlow** – промышленный фреймворк для развертывания моделей
- **JAX** – высокопроизводительные численные вычисления с автоматическим дифференцированием
- **Hugging Face Transformers** – библиотека предобученных моделей NLP
- **Diffusers** – библиотека для работы с диффузионными моделями

5. Специализированные библиотеки для фронтальных исследований

- **PyTorch Geometric** – графовые нейронные сети и анализ сетевых данных

- **Einops** – тензорные операции с удобным синтаксисом для сложных архитектур
- **Optuna** – байесовская оптимизация гиперпараметров
- **Stable-Baselines3** – реализация современных RL алгоритмов

6. Визуализация и анализ результатов

- **TensorBoard** – визуализация процесса обучения и метрик моделей
- **Plotly** – интерактивная визуализация данных и результатов экспериментов
- **Netron** – просмотр архитектур нейронных сетей
- **Matplotlib/Seaborn** – статическая визуализация данных

7. Системы управления обучением и коллаборацией

- **Moodle** – сдача работ, тестирование и управление учебными материалами
- **Notion** – ведение исследовательских дневников и документации проектов
- **Overleaf** – совместная работа над научными статьями в LaTeX
- **Slack/Discord** – коммуникация в исследовательских группах

8. Высокопроизводительные вычисления

- **Docker** – контейнеризация сред воспроизведения экспериментов
- **Kubernetes** – оркестрация распределенного обучения моделей
- **NVIDIA NGC** – оптимизированные контейнеры для глубокого обучения
- **Horovod** – распределенное обучение моделей на multiple GPU

9. Инструменты для развертывания моделей

- **FastAPI** – создание API для ML моделей
- **Streamlit** – быстрая разработка веб-интерфейсов для демонстрации моделей
- **Gradio** – создание UI для ML моделей за несколько строк кода
- **ONNX Runtime** – кроссплатформенное выполнение оптимизированных моделей

10. Базы данных для исследовательских проектов

- **Neo4j** – графовая база данных для нейро-символьных систем
- **Chroma** – векторная база данных для semantic search
- **Pinecone** – управление векторными эмбедами в продакшене
- **SQLite/PostgreSQL** – реляционные базы данных для метаданных экспериментов

Примечание: Все перечисленные технологии используются в рамках курса для обеспечения практической ориентированности и соответствия современным стандартам исследований в области искусственного интеллекта.

7.2 Перечень лицензионного и свободно распространяемого программного обеспечения

1. Лицензионное ПО

VSCoде – IDE для Python (свободнораспространяемое)

LibreOffice– оформление отчетов (свободнораспространяемое)

2. Свободное ПО (Open Source)

Hugging Face Transformers – предобученные модели (BERT, GPT, визуальные трансформеры)

Hugging Face Optimum – оптимизация и развёртывание моделей

Фреймворки и библиотеки для DL:

PyTorch/TensorFlow/Keras– разработка нейросетей

Инструменты для визуализации:

Streamlit/Gradio – создание веб-интерфейсов для моделей

Matplotlib/Seaborn – графики и анализ данных

СУБД:

SQLite/PostgreSQL – хранение структурированных данных
FAISS/Annoy – векторный поиск

8. Материально-техническое обеспечение по дисциплине (модулю)

Виртуальные машины и ресурсы GPU в облаке предоставляется индустриальным партнером ПАО «Сбербанк»:

№	Продукт	Параметры продукта	Кол-во	Кол-во конфигураций	Ед. изм.
1	Виртуальная машина	Виртуальная машина 10% vCPU 2 vCPU 4 RAM	1	60	Шт
		ОС Ubuntu 22.04	1		Шт
		Системный диск SSD	1		Шт
			10		Гб
		Аренда публичного IP	1		Шт
2	Виртуальная машина с GPU	Виртуальная машина с GPU NVIDIA® Tesla® V100 2 GPU 8 vCPU 128 ГБ RAM	1	1	Шт
		ОС Ubuntu_24.04	1		Шт
		Системный диск SSD	1		Шт
			2000		Гб
		Диск SSD	1		Шт
			4096		Гб
		Диск SSD	1		Шт
	4096	Гб			
3	K8S	Аренда публичного IP	1	1	Шт
		Master node 8 vCPU 16 RAM	1		Шт
		Worker node 10% доля 4 vCPU 32 RAM	5		Шт
		Worker node SSD-NVME	64		Гб
		Аренда публичного IP	1		Шт
4	ML Inference Instance Type GPU	Время работы в месяц	40	1	Ч
		Инстанс 8 x NVIDIA® H100 NVLink PCIe 160 vCPU 1520 GB RAM	1		Шт
		Количество запросов к ML-моделям	1		Млн. Шт
		Кэш ML-моделей	160		Гб
5	LLM	Токены GigaChat 2 Max	50		Млн. Шт
		Токены Embeddings	400		Млн.

Дополнительные облачные ресурсы предоставляются технологическим партнером Yandex Cloud.

№	Вид работ	Наименование учебной аудитории, ее оснащенность оборудованием и техническими средствами обучения
1.	Лекционные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
2.	Лабораторные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, проектором, программным обеспечением
3.	Практические занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
4.	Групповые (индивидуальные) консультации	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
5.	Текущий контроль, промежуточная аттестация	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
6.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.

Примечание: Конкретизация аудиторий и их оснащение определяется ОПОП.