
УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«25» мая 2026 г.
Протокол № 2

**Рабочая программа дисциплины (модуля)
«Графовые нейросети»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Методы исследований в
машинном обучении

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2026

**Москва
2026**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение	9
6. Материально-техническое обеспечение	9
7. Методические и оценочные материалы	11

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Графовые нейросети» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по направлению 02.04.01 Математика и компьютерные науки, профиль Методы исследований в машинном обучении, утвержденным приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Графовые нейросети» позволяет сформировать целостное понимание принципов машинного обучения на графовых структурах, включая формализацию прикладных задач, выбор и реализацию современных архитектур графовых нейронных сетей, а также развитие навыков проектирования, обучения и анализа моделей, способных эффективно извлекать знания из сложных связанных данных.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Методы исследований в машинном обучении и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений, как дисциплина по выбору.

Дисциплина (модуль) доступна к изучению на 2 курсе в 4 семестре. Доступна для прохождения при условии успешного завершения дисциплины (модуля) «Глубокое обучение (Deep Learning)».

Цель изучения дисциплины (модуля): подготовить обучающихся к решению задач анализа и прогнозирования на графовых данных путём освоения классических методов теории графов и современных подходов глубокого обучения, включая графовые нейронные сети, с акцентом на практическую реализацию, корректную постановку экспериментов и интерпретацию результатов в реальных прикладных контекстах.

Задачи изучения дисциплины (модуля):

- освоить базовые задачи графового машинного обучения: классификация вершин, прогнозирование рёбер, классификация графов, выявление сообществ и генерация структур;
- изучить классические методы анализа графов: меры центральности, PageRank, спектральные разложения, случайные блуждания и модели случайных графов;
- приобрести навыки реализации и настройки современных архитектур GNN: GCN, GraphSAGE, GAT, GIN, включая механизмы передачи сообщений, агрегации и пулинга;
- научиться формализовать прикладную задачу в терминах графовых данных и выбирать адекватную модель в зависимости от структуры графа и типа задачи;
- развить компетенции в проведении воспроизводимых экспериментов: корректное разбиение данных, негативное сэмплирование, оценка качества, абляционный анализ;
- освоить инструменты и практики масштабируемого обучения на больших графах с использованием эффективных даталоадеров и техник выборки соседей.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- базовые модели и задачи графового машинного обучения: классификация вершин, прогнозирование связей, классификация графов, выявление сообществ, генерация графов;
- классические методы анализа графов: меры центральности, PageRank, распространение меток, случайные блуждания, модели случайных графов и спектральные методы;
- современные архитектуры графовых нейронных сетей: передача сообщений, GCN, GraphSAGE, GAT, GIN, считывание/агрегация графового представления, пулинг, модели для гетерогенных и многореляционных графов.

уметь:

- формализовать прикладную задачу как задачу обучения на графе и выбрать корректный тип графа, признаки вершин/рёбер и целевую переменную;
- обучать и сравнивать модели для классификации вершин, прогнозирования связей и классификации графов, включая классические базовые методы сравнения и GNN;
- строить эксперимент с графовыми данными: разбиение рёбер/вершин, негативное сэмплирование, метрики качества, абляционный анализ, интерпретация ошибок.

владеть:

- инструментами Python для работы с графами и GNN: NetworkX, PyTorch, PyTorch Geometric/DGL по выбору преподавателя;
- практиками масштабируемого обучения на больших графах: мини-батчи, выборка соседей, графовые даталоадеры, контроль утечек данных.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-6.	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.1.	Знает основные методы самооценки и анализа своей деятельности, а также принципы управления временем и целеполагания
		УК-6.2.	Умеет ставить реалистичные и достижимые цели, определять приоритеты в своей деятельности, а также разрабатывать и внедрять планы по совершенствованию своих навыков и компетенций на основе полученной самооценки
		УК-6.3.	Имеет практический опыт применения методов самооценки в своей профессиональной деятельности, включая участие в тренингах, семинарах и проектах, направленных на развитие личной эффективности и профессионального роста
ОПК-1.	Способен находить, формулировать и решать актуальные и значимые проблемы прикладной и компьютерной математики	ОПК-1.1.	Знает современные методы прикладной математики и машинного обучения, а также программные средства анализа данных, применяемые в электронной коммерции для решения задач персонализации предложений, прогнозирования спроса и оптимизации пользовательского взаимодействия
		ОПК-1.2.	Умеет выявлять и формализовать задачи области профессиональной деятельности в виде математических постановок, выбирать и применять соответствующие методы и алгоритмы, интерпретировать полученные результаты и представлять их в форме, понятной специалистам различного профиля

		ОПК-1.3.	Имеет практический опыт решения прикладных задач в сфере профессиональной деятельности, включая работу с реальными наборами данных, построение и оценку качества моделей машинного обучения, а также участие в проектной и исследовательской деятельности, направленной на повышение эффективности цифровых платформ
ПК-3.	Способен решать задачи профессиональной деятельности, формулировать результат, увидеть следствия полученного результата	ПК-3.1.	Знает основные принципы и методы решения задач профессиональной деятельности, а также способы формулирования и представления результатов, включая анализ последствий и их значимость в контексте проекта
		ПК-3.2.	Умеет применять математические и компьютерные методы для решения конкретных задач, формулировать четкие и обоснованные результаты, а также анализировать их последствия для дальнейших действий и решений
		ПК-3.3.	Имеет практический опыт в решении профессиональных задач, включая участие в проектах, где были получены результаты и проанализированы их следствия, что способствовало принятию обоснованных решений

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостояте льная работа	
		Лекции	Семинары (практичес кие занятия)			
1	Основы графового машинного обучения	6	6		25	Домашнее задание
2	Графовые нейронные сети и прогнозирование связей	6	6		25	Домашнее задание
3	Классификация графов, графы знаний и масштабирование	6	6		25	Домашнее задание
4	Графовые алгоритмы, анализ сетей и генеративные модели	6	6		25	Домашнее задание
5	Исследовательский проект	6	6		26	Проект
	Экзамен			4		
Итого:		30	30	4	126	
Объем дисциплины (модуля) (в ак. ч.)		190				
Объем дисциплины (модуля) (в зач. ед.)		5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Основы графового машинного обучения	Графы как объект машинного обучения: типы графов, признаки вершин и ребер, прикладные задачи, алгоритм ранжирования страниц Классификация вершин в графах без атрибутов: распространение меток, структурные признаки, оценка качества Эмбединги вершин без учителя: случайные блуждания, DeepWalk, node2vec, негативное сэмплирование
2	Графовые нейронные сети и прогнозирование связей	Графовые нейронные сети: схема передачи сообщений, агрегаторы, инвариантность к перестановкам, выразительная способность Базовые архитектуры графовых нейронных сетей: графовая сверточная сеть, GraphSAGE, графовая сеть внимания, графовая изоморфная сеть; пересглаживание и сжатие дальних зависимостей Прогнозирование связей: декодеры, негативное сэмплирование, разбиение ребер и утечки данных
3	Классификация графов, графы знаний и масштабирование	Классификация графов: считывание графового представления, графовый пулинг, графовые ядра и нейросетевые модели для графов

		Гетерогенные и многореляционные графы: графы знаний, реляционная графовая сверточная сеть, метапути, рекомендации на графах Масштабируемое обучение на больших графах: мини-батчи, выборка соседей, методы выборки подграфов и графовые загрузчики данных
4	Графовые алгоритмы, анализ сетей и генеративные модели	Комбинаторная оптимизация на графах: обучение эвристик, нейросетевые методы комбинаторной оптимизации, ограничения и качество решений Поиск ближайших соседей в пространствах высокой размерности: графовые индексы, иерархические графы навигации в малом мире, приближенный поиск ближайших соседей Анализ сетей и модели случайных графов: модель Эрдеша—Реньи, модель Барабаши—Альберт, стохастическая блочная модель Выявление сообществ: модулярность, спектральная кластеризация, алгоритмы выявления сообществ и оценка результатов Генеративные модели графов: модели молекул и сетей, графовые автоэнкодеры, современные графовые генеративные модели
5	Исследовательский проект	Генеративные модели графов: модели молекул и сетей, графовые автоэнкодеры, современные графовые генеративные модели Итоговый проект: воспроизведение статьи или прикладной кейс по графовому машинному обучению, защита результатов

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Иванов, Б. Н. Дискретная математика и теория графов : учебник для вузов / Б. Н. Иванов. — Москва : Издательство Юрайт, 2025. — 177 с. — (Высшее образование). — ISBN 978-5-534-14470-3. — Текст : электронный // Образовательная платформа Юрайт. — URL: <https://urait.ru/bcode/567929>.

2. Гисин, В. Б. Дискретная математика : учебник и практикум для вузов / В. Б. Гисин. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 428 с. — (Высшее образование). — ISBN 978-5-534-16763-4. — Текст : электронный // Образовательная платформа Юрайт. — URL: <https://urait.ru/bcode/577329>.

Дополнительная литература:

1. Галушкин, А. И. Нейронные сети: основы теории : монография / А. И. Галушкин. — Москва : Горячая линия-Телеком, 2024. — 496 с. — ISBN 978-5-9912-0082-0. — Текст : электронный // ЭБС Znanium. — URL: <https://znanium.ru/catalog/product/2186186>.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система «Единое окно доступа к образовательным ресурсам»	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		

Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Графовые нейросети» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, проект, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар — это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал, использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Проект – исследовательская работа по дисциплине (модулю) и презентация результатов.

Для успешной подготовки к проекту рекомендуется: четко определить цели и задачи проекта; составить план работы, разбив проект на этапы с указанием сроков выполнения каждого из них; использовать разнообразные источники информации и инструменты для исследования темы; регулярно проверять прогресс и вносить коррективы в план, если это необходимо.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Графовые нейросети»

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме **экзамена**, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в силлабусе дисциплины:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы,
9	Отлично	
8	Отлично	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	
5	Удовлетворительно	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Графовые нейросети» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	50%	5	Набор задач по темам недели
Проект	20%	1	Исследовательская работа по дисциплине (модулю) и презентация результатов
Экзамен	30%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Графовые нейросети»: $0,5 \times \text{среднее за домашние задания} + 0,2 \times \text{проект} + 0,3 \times \text{экзамен}$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1. Классификация вершин и эмбединги без учителя

Задание:

1. Используйте граф из набора данных Cora или PubMed (цитирование научных статей).
2. Реализуйте:
 - Распространение меток (Label Propagation) для классификации вершин
 - Генерацию эмбедингов с помощью node2vec (случайные блуждания + Word2Vec)
 - Обучите классификатор (например, логистическую регрессию) на эмбедингах
3. Сравните:
 - Качество классификации (accuracy, F1)
 - Зависимость от параметров (длина блужданий, p/q в node2vec)
4. Проанализируйте:
 - Как структура графа влияет на качество методов?
 - Какие вершины классифицируются хуже (низкая степень, граница сообществ)?

Требования:

- Код на Python (NetworkX, node2vec, scikit-learn)
- Визуализация графа (через t-SNE или Gephi)
- Отчёт с анализом и сравнением подходов

Домашнее задание 2. Прогнозирование связей с GNN**Задание:**

1. Выберите граф (например, Cora, PubMed или Facebook из PyG).
2. Реализуйте:
 - GCN или GraphSAGE для получения эмбедингов вершин
 - Пространственный декодер (например, dot product) для предсказания наличия рёбер
3. Подготовьте данные:
 - Разделите рёбра на обучающие/валидационные/тестовые
 - Проведите негативное сэмплирование
 - Убедитесь, что нет утечки данных (например, через пересечение подграфов)
4. Обучите модель:
 - Оцените AUC, AP на тесте
 - Проведите абляционный анализ (сравните с случайными эмбедингами, без GNN)
5. Обсудите:
 - Как GNN учитывает структуру?
 - Какие ошибки характерны (ложные срабатывания между сообществами)?

Требования:

- Реализация на PyTorch Geometric или DGL
- Чёткая стратегия разбиения
- Анализ утечек данных и их предотвращение

Домашнее задание 3. Классификация графов и пулинг**Задание:**

Электронный документ

1. Используйте набор данных для классификации графов (например, MUTAG, PROTEINS или MNIST Superpixels).
2. Реализуйте GNN-архитектуру с механизмом пулинга:
 - GraphSAGE + global mean pooling
 - GIN + attention pooling или hierarchical pooling (DiffPool)
3. Обучите модель на задаче классификации графов:
 - Сравните качество (accuracy, ROC-AUC)
 - Проанализируйте, как пулинг влияет на представление графа
4. Проведите визуализацию:
 - Активации узлов в разных слоях
 - Веса внимания при пулинге (если применимо)
5. Объясните:
 - Почему важно избегать пересглаживания (over-squashing)?
 - Как иерархический пулинг помогает в анализе крупных графов?

Требования:

- Код с использованием PyG/DGL
- Графики обучения и анализа
- Интерпретация результатов в контексте выразительности моделей

Примерное описание задания и критерии оценивания к проекту

Цель проекта: разработать и протестировать графовую модель для решения комплексной задачи — например, предсказание новых фактов в графе знаний, классификация сущностей или выявление аномалий — с использованием современных GNN и методов анализа графов.

Этапы выполнения:

1. **Выбор задачи и данных**
 - Определить тип задачи: предсказание рёбер, классификация узлов, выявление сообществ.
 - Выбрать или построить граф: например, подмножество Wikidata, DBpedia, или синтетический граф.
2. **Предобработка и анализ графа**
 - Построить статистику: число узлов, рёбер, распределение степеней, плотность.
 - Визуализировать структуру (через NetworkX или Gephi).
 - Выделить сообщества (например, через Louvain) и оценить модулярность.
3. **Реализация и обучение моделей**
 - Реализовать как минимум два подхода:
 - Классический (например, node2vec + классификатор).
 - GNN (например, R-GCN для гетерогенного графа или GAT).
 - Обучить модели с контролем утечек данных
4. **Оценка и анализ**
 - Сравнить по метрикам (AUC, F1, accuracy).
 - Провести абляционный анализ (вклад структуры, признаков, архитектуры).
 - Проанализировать ошибки: где модель терпит неудачу и почему.
5. **Презентация результатов**

- Подготовить презентацию (5–7 слайдов).
- Защитить проект: постановка, методы, результаты, выводы, ограничения.

Критерии оценивания проекта:

- **Корректность постановки задачи.**
Задача чётко сформулирована, соответствует возможностям графовых методов, выбраны адекватные целевые переменные и тип графа.
- **Качество предобработки и анализа данных.**
Проведён структурный анализ графа (степени, компоненты связности, сообщества), визуализация, выявлены особенности структуры.
- **Адекватность выбранных методов.**
Использованы как классические, так и современные подходы (GNN), учтены особенности графа (гетерогенность, масштаб, тип связей).
- **Глубина анализа результатов.**
Проведено сравнение моделей, абляционный анализ, интерпретация ошибок, обсуждение ограничений.
- **Качество презентации и защита.**
Материал представлен структурированно, логично, с визуализацией. Ответы на вопросы демонстрируют понимание методов и данных.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Что такое граф в контексте машинного обучения?	Структура из узлов и рёбер с признаками, отражающая отношения между сущностями	ПК-3
2.	Как работает Label Propagation?	Алгоритм распространяет метки от размеченных узлов к соседям итеративно	ПК-3
3.	Что такое node2vec?	Метод генерации эмбедингов на основе параметризованных случайных блужданий	ПК-3
4.	Что такое GCN (Graph Convolutional Network)?	GNN, агрегирующая представления соседей с нормализацией по степеням узлов	ПК-3
5.	Почему важно избегать утечек данных в link prediction?	Чтобы не использовать информацию из теста при обучении, что ведёт к переоценке качества	ПК-3
6.	Что такое передача сообщений в GNN?	Обобщённая схема, включающая агрегацию, обновление и чтение сообщений от соседей	ОПК-1
7.	Почему GNN инвариантны к перестановке узлов?	Архитектура использует симметричные функции агрегации (например, сумма, среднее)	ОПК-1
8.	Что такое модулярность в выявлении сообществ?	Мера, оценивающая качество разбиения графа на плотные подграфы	ОПК-1

9.	Как работает алгоритм Louvain?	Жадная оптимизация модулярности через итеративное кластеризование	ОПК-1
10.	Что такое HNSW (Hierarchical Navigable Small World)?	Структура для эффективного приближённого поиска ближайших соседей в высоких размерностях	ОПК-1
11.	Что такое GIN (Graph Isomorphism Network)?	Максимально выразительная GNN, способная различать изоморфные графы	ПК-3
12.	Как работает GraphSAGE?	Индуктивный метод, использующий выборку соседей для генерации эмбедингов	ПК-3
13.	Что такое метапуть в гетерогенном графе?	Последовательность типов узлов и рёбер, описывающая семантическую связь (например, автор→статья→конференция)	ПК-3
14.	Что такое DiffPool?	Иерархический механизм пулинга, обучающий кластеризацию узлов на каждом слое	ПК-3
15.	Как R-GCN обрабатывает многореляционные графы?	Использует разные веса для разных типов рёбер при агрегации	ПК-3
16.	В чём заключается идея негативного сэмплирования?	Генерация "ложных" рёбер для обучения модели различать реальные и вымышленные связи	УК-6
17.	Почему важно фиксировать random seed в экспериментах?	Для обеспечения воспроизводимости результатов и научной достоверности	УК-6
18.	Какие метрики используются в link prediction?	AUC, Average Precision (AP), F1, Precision@K	УК-6
19.	Как организовать корректное разбиение графа на обучение/тест?	Разделять рёбра или узлы без пересечения подграфов, избегая утечек данных	УК-6
20.	Почему важно документировать эксперименты по GNN?	Чтобы обеспечить прозрачность, воспроизводимость и возможность критического анализа	УК-6