
УТВЕРЖДЕНА

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

**Рабочая программа дисциплины (модуля)
«Графовые методы в машинном обучении»**

Направление подготовки: 38.03.05 Бизнес-информатика

Направленность (профиль) подготовки: Бизнес и аналитика

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Срок освоения программы: 4 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)	8
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 методы в машинном обучении» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по специальности 38.03.05 Бизнес-информатика, профиль Бизнес и аналитика, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 838 от 29.07.2020 года.

Дисциплина (модуль) «Графовые методы в машинном обучении» позволяет учитывать сложные взаимосвязи между объектами, улучшать качество предсказаний в рекомендационных системах, биоинформатике, сетевой безопасности и прочих областях, где данные естественно задаются в виде графов; они также способствуют развитию новых архитектур нейронных сетей (GNN), повышая эффективность и интерпретируемость моделей.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки бакалавриата по направлению 38.03.05 Бизнес-информатика, профиль Бизнес и аналитика и входит в часть, формируемую участниками образовательных отношений, Блока 1.

Дисциплина (модуль) является выборной и доступна для изучения на 3 или 4 курсе в 6, 7, 8 семестрах на выбор после успешного освоения дисциплин «Основы математического анализа и линейной алгебры» или "Математический анализ" и "Линейная алгебра и геометрия" или "Математический анализ. Продвинутый уровень" и "Линейная алгебра и геометрия. Продвинутый уровень".

Цель изучения дисциплины (модуля): освоить навыки применения графовых структур и алгоритмов к задачам машинного обучения, позволяющие эффективно работать с данными, обладающими отношениями и топологией.

Задачи изучения дисциплины (модуля):

- понять основные понятия графов (вершины, ребра, вес, направленность) и их математическое представление (матрицы смежности, лапласиан);
- изучить методы извлечения признаков из графов: спектральные эмбединги, random walk, node2vec, graph convolutional networks;
- освоить техники агрегации и трансформации графовых сигналов для классификации, регрессии, кластеризации и рекомендаций;
- реализовать и оценить графовые алгоритмы на реальных датасетах (социальные сети, биологические взаимодействия, Knowledge Graphs);
- сравнивать графовые подходы с традиционными методами и анализировать их преимущества и ограничения.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- понять основные понятия графов (вершины, ребра, вес, направленность) и их математическое представление (матрицы смежности, лапласиан);
- изучить методы извлечения признаков из графов: спектральные эмбединги, random walk, node2vec, graph convolutional networks;
- освоить техники агрегации и трансформации графовых сигналов для классификации, регрессии, кластеризации и рекомендаций;
- реализовать и оценить графовые алгоритмы на реальных датасетах (социальные сети, биологические взаимодействия, Knowledge Graphs);
- сравнивать графовые подходы с традиционными методами и анализировать их преимущества и ограничения;

уметь:

- понять основные понятия графов (вершины, ребра, вес, направленность) и их математическое представление (матрицы смежности, лапласиан);
- изучить методы извлечения признаков из графов: спектральные эмбединги, random walk, node2vec, graph convolutional networks;
- освоить техники агрегации и трансформации графовых сигналов для классификации, регрессии, кластеризации и рекомендаций;
- реализовать и оценить графовые алгоритмы на реальных датасетах (социальные сети, биологические взаимодействия, Knowledge Graphs);
- сравнивать графовые подходы с традиционными методами и анализировать их преимущества и ограничения;

владеть:

- понять основные понятия графов (вершины, ребра, вес, направленность) и их математическое представление (матрицы смежности, лапласиан);
- изучить методы извлечения признаков из графов: спектральные эмбединги, random walk, node2vec, graph convolutional networks;
- освоить техники агрегации и трансформации графовых сигналов для классификации, регрессии, кластеризации и рекомендаций;
- реализовать и оценить графовые алгоритмы на реальных датасетах (социальные сети, биологические взаимодействия, Knowledge Graphs);
- сравнивать графовые подходы с традиционными методами и анализировать их преимущества и ограничения.

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-1.	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1.	Знает методы поиска и анализа информации в области аналитики, основные принципы критической оценки источников информации и их релевантности.
		УК-1.2.	Умеет критически оценивать источники информации и синтезировать данные из различных источников для решения задач, применять системный подход к анализу и решению комплексных проблем
		УК-1.3.	Имеет практический опыт работы с современными инструментами и технологиями для обработки информации, формулировании и структурировании задач на основе полученной информации
ОПК-4.	Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений	ОПК-4.1.	Знает основные принципы работы информационных технологий и их влияние на бизнес-процессы
		ОПК-4.2.	Умеет использовать методы и программные средства для сбора, обработки и анализа информации, обеспечивая качественную информационно-аналитическую поддержку
		ОПК-4.3.	Имеет практический опыт в применении аналитических инструментов для поддержки принятия управленческих решений в организациях
ПК-2.	Способен использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации	ПК-2.1.	Знает основные математические методы и инструментальные средства, применяемые для обработки и анализа информации
		ПК-2.2.	Умеет эффективно использовать математический аппарат для систематизации данных и решения профессиональных задач

	информации по теме исследования для решения задач профессиональной деятельности	ПК-2.3.	Имеет практический опыт работы с инструментами анализа информации в рамках исследовательских проектов
--	---	---------	---

3. Тематический план

№п/ п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостоятельная работа	
		Лекции	Семинары			
1	Введение в графы и классические методы	4	4		20	Домашнее задание Аудиторная работа
2	Графовые эмбединги	4	4	2	20	Домашнее задание Аудиторная работа Контрольная работа
3	Введение в графовые нейронные сети (GNN)	4	4		20	Домашнее задание Аудиторная работа
4	Механизмы внимания и продвинутые GNN	6	6	2	20	Домашнее задание Аудиторная работа
5	Применения GNN	4	4		20	Домашнее задание Аудиторная работа
6	Генеративные модели и масштабирование	4	4		10	Домашнее задание Аудиторная работа
7	Графовые методы и нейронные сети на графах	4	4	2	10	Домашнее задание Аудиторная работа Контрольная работа
	Зачет с оценкой			4		Коллоквиум
Итого:		30	30	10	120	
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	Введение в графы и классические методы	Типы графов. Матрица смежности, списки. BFS, DFS Извлечение признаков: степени вершин, треугольники, кластеризация Алгоритмы распространения меток: PageRank, Label Propagation
2	Графовые эмбединги	Node2Vec, DeepWalk Эмбединги для графов знаний и целых графов. Контрольная работа 1
3	Введение в графовые нейронные сети (GNN)	Парадигма передачи сообщений Графовая сверточная сеть (GCN) GraphSAGE: семплирование и агрегация
4	Механизмы внимания и продвинутые GNN	Графовые сети внимания (GAT). Коллоквиум Регуляризация Классификация узлов, предсказание связей, классификация графов
5	Применения GNN	Графовые нейросети в компьютерном зрении (сцены, объекты) Графовые нейросети в NLP (синтаксические графы, извлечение отношений)
6	Генеративные модели и масштабирование	Генерация графов (GraphVAE, MolGAN)
7	Графовые методы и нейронные сети на графах	Выбор архитектуры GNN в зависимости от задачи. Современные направления и прикладные вызовы. Контрольная работа 2

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. Сухарев, А. Г. Графовые методы в машинном обучении оптимизации : учебник и практикум для вузов / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. — 3-е изд., испр. и доп. — Москва : Издательство Юрайт, 2026. — 367 с. — (Высшее образование). — ISBN 978-5-534-17381-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/584770>

2. Аверина, Т. А. Графовые методы в машинном обучении. Верификация алгоритмов решения систем со случайной структурой : учебник для вузов / Т. А. Аверина. — Москва : Издательство Юрайт, 2025. — 179 с. — (Высшее образование). — ISBN 978-5-534-07205-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/564782>

Дополнительная литература:

1. Локтионов, И. К. Графовые методы в машинном обучении : учебник / И. К. Локтионов, Л. П. Мироненко, В. В. Турупалов ; под общ. ред. канд. техн. наук, проф. В. В. Турупалова. - Москва ; Вологда : Инфра-Инженерия, 2022. - 380 с. - ISBN 978-5-9729-0786-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1902598>.

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 (модуле), 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 методы в машинном обучении» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекция, семинары, аудиторная работа, контрольные работы и домашние задания, коллоквиум, 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.

Аудиторная работа — активная работа студента на семинаре, его ответы на вопросы преподавателя и участие в дискуссии.

Для успешного участия в семинаре студентам рекомендуется заранее ознакомиться с темой обсуждения, прочитать необходимые материалы и подготовить вопросы. Важно активно слушать и вовлекаться в дискуссию, высказывая свои мнения и аргументируя их. При ответах на вопросы преподавателя стоит быть уверенным, четким и логичным,

опираясь на изученный материал. Также полезно поддерживать диалог с однокурсниками, чтобы обогатить обсуждение и расширить свои знания.

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.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы - получить специальные знания по одной или нескольким темам дисциплины (модуля) и продемонстрировать навыки их практического применения.

Коллоквиум – устные ответы на вопросы, список которых известен студенту заранее.

В процессе подготовки к коллоквиуму необходимо проанализировать учебные материалы, ознакомившись с лекциями, учебниками и дополнительными источниками, акцентируя внимание на ключевых темах. Рекомендуется создать структурированные конспекты, выделяя основные идеи, термины и формулы.

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 методы в машинном обучении»

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 **зачета с оценкой**, 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	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Зачтено	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину. Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать,
9	Отлично	Зачтено	
8	Отлично	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, 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	Зачтено	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	Зачтено	
5	Удовлетворительно	Зачтено	Студент обладает базовыми знаниями по дисциплине (модулю),
4	Удовлетворительно	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения 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.
3	Не сдан	Не зачтено	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «Графовые методы в машинном обучении» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	25%	10	Набор задач по темам недели
Контрольные работы	35%	2	Письменная работа с набором задач, которые нужно решить за ограниченное время
Аудиторная работа	15%	15	Ответы на вопросы, список которых известен студенту заранее
Зачет с оценкой	25%	1	Коллоквиум - устные ответы на вопросы, список которых известен студенту заранее.

Формула расчёта итоговой оценки по дисциплине (модулю) «Графовые методы в машинном обучении»: « $0,25 \times \text{среднее за домашние задания} + 0,35 \times \text{среднее за контрольные работы} + 0,15 \times \text{среднее за аудиторную работу} + 0,25 \times \text{зачет с оценкой}$ ».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные вопросы для семинаров

- какие основные представления графа используются в машинном обучении (матрица смежности, список смежности, атрибуты вершины и ребра);
- как формируется нормализованный лапласиан $L_{\text{sym}} = I - D^{-1/2} A D^{-1/2}$ и зачем он нужен в спектральных методах;
- в чём различие между задачами node-level, edge-level и graph-level;
- опишите основной принцип алгоритма PageRank (одна-строчная формула);
- как работает label-propagation и в каких задачах его применяют;
- что такое oversmoothing в GNN и какие симптомы наблюдаются;
- какие механизмы внимания применяются в GAT (self-attention, multi-head и т.п.);
- в чём заключается агрегация соседей в GraphSAGE (mean, max, LSTM);

9. назовите два применения графовых нейронных сетей в реальных областях (например, молекулярный дизайн, рекомендационные системы);
10. какие метрики обычно используют для оценки link-prediction (AUC-ROC, AP) и для graph-classification (accuracy, F1).

Примерные задания по контрольной работе

1. матричные представления

- a) задайте ориентированный граф из 5 вершин, выпишите его матрицу смежности A и степень-диагональ D ;
- b) вычислите несмонтированный лапласиан $L=D-A$ и симметричный нормализованный лапласиан L_{sym} .

2. спектральное кластерирование

используя матрицу L_{sym} из пункта 1, найдите два собственных вектора, соответствующие наименьшим ненулевым собственным значениям;

выполните k-means ($k=2$) по этим векторам и укажите полученные кластеры вершин.

3. графовая сверточная сеть (GCN)

дано $\tilde{A}=A+I$ и $\tilde{D}$ – степень-диагональ $\tilde{A}$. Запишите формулу обновления признаков для одного слоя GCN с функцией активации ReLU;

используя входные признаки $X=1010101100$ и весовую матрицу $W=[10-11]$, вычислите $H(1)$.

4. GAT-механизм

- a) запишите формулу коэффициента внимания α_{ij} для одного голоса (single-head) с использованием линейных проекций $aT[Wh_i||Wh_j]$ и softmax;
- b) для трёх соседних вершин с признаками $[1][0],[0][1],[1][1]$ и весом $W=I$ вычислите ненормализованные коэффициенты e_{ij} (положите $\text{LeakyReLU}(x)=\max(0.2x,x)$).

5. oversmoothing

объясните, почему при увеличении количества GCN-слоёв признаки всех вершин стремятся к одинаковому вектору; предложите два способа борьбы с этим (например, skip-connections, регуляризация градиента).

6. оценка модели

дан набор предсказаний вероятностей существования ребра и истинные метки: $\{(0.9,1),(0.4,0),(0.6,1),(0.2,0),(0.7,1)\}$.

вычислите AUC-ROC (по формуле тройки) и precision@k при $k=3$.

7. программная реализация

опишите последовательность шагов (pipeline) для обучения GNN на датасете Cora в PyTorch-Geometric: загрузка данных, построение edge-index, нормализация, определение модели, обучение, оценка.

Примерные домашние задания

1. реализация базовых представлений

- напишите функцию `adjacency_matrix(edge_index, num_nodes)` и `degree_matrix(adj)`.
- проверьте корректность на случайном графе из 7 вершин (генерация через `torch_geometric.utils.erdos_renyi_graph`).

2. spectral embedding

- реализуйте алгоритм `spectral_clustering(A, k)` (поиск k -наименьших собственных векторов нормализованного лапласиана и k-means).
- визуализируйте полученные кластеры на графе с 2-мя признаками координат вершин.

3. однослойный GCN

- реализуйте класс **SimpleGCN(torch.nn.Module)** с единственным слоем согласно формуле из контрольной работы.
- обучите модель на задаче node-classification на датасете **Cora** (разделите 20 % обучающих, 10 % валидационных).
- выведите точность на тесте и сравните с базовым LogisticRegression (используя только исходные признаки).
- 4. **GAT-модель**
 - используя **torch_geometric.nn.GATConv**, постройте двухслойный GAT с 4-ю головами в первом слое и 1-й головой во втором.
 - проведите обучение на той же задаче node-classification и сравните метрику F1 с GCN-моделью.
- 5. **oversmoothing-эксперимент**
 - обучите GCN с разным числом слоёв от 1 до 8, фиксируя остальные гиперпараметры.
 - для каждого количества слоёв посчитайте среднее расстояние (cosine similarity) между признаками разных вершин; постройте график зависимости от числа слоёв.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1	укажите, как в матричной форме задаётся матрица смежности графа G .	$A_{ij}=1$, если $(i,j) \in E$, иначе 0;	УК 1
2	назовите матрицу степенной (degree) диаграммы D .	диагональная матрица, $D_{ii}=\sum_j A_{ij}$;	УК 1
3	запишите формулу ленточного (graph) лапласиана L .	$L=D-A$;	УК 1
4	как называется нормализованный лапласиан L_{sym} .	L_{sym} $=I-D^{-1/2}AD^{-1/2}$;	ПК 2
5	укажите основной принцип PageRank (одно слово).	случайный-перехват;	УК 1
6	назовите алгоритм распространения меток (одно слово).	label-propagation;	УК 1
7	запишите однослойную графовую сверточную сеть (gcn) : формула обновления признаков $H(l+1)$.	$H(l+1)=\sigma(D^{-1/2}A^{-1/2}H(l))$, где $A^{-1}=A+I$;	ПК 2
8	какое механизм внимания используется в gat (одно слово).	self-attention;	ПК 2
9	чем graphsage отличается от традиционного gcn (одно слово).	агрегация;	УК 1
10	назовите тип задачи node-level (одно слово).	классификация;	УК 1
11	назовите тип задачи edge-level (одно слово).	предсказание;	УК 1
12	назовите тип задачи graph-level (одно слово).	регрессия;	УК 1
13 **	как называется проблема, когда после нескольких слоёв gnn признаки вершин становятся одинаковыми (одно слово).	oversmoothing	УК 1
14	укажите один способ борьбы с oversmoothing.	добавление skip-connections;	ПК 2
15	назовите один эмбединг-метод для получения векторных представлений узлов без нейронных сетей.	node2vec;	ПК 2
16	как называется процесс агрегирования признаков соседей в graphsage (одно слово).	mean-pooling;	ОПК-4
17	укажите один критерий оценки качества модели link-prediction .	auc-roc;	ОПК-4
18	назовите метрику для задачи graph-classification .	accuracy;	ОПК-4

19	какой фреймворк в python позволяет быстро строить gnn-модели (одно слово).	pytorch-geometric;	ОПК-4
20	укажите один этап построения пайплайна для обучения gnn: подготовка признаков, нормализация, обучение, ... (одно слово).	нормализация;	ОПК-4