

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Львівський національний університет імені Івана Франка
Факультет прикладної математики та інформатики
Кафедра обчислювальної математики

Оновлено та затверджено
на засіданні
кафедри обчислювальної математики
факультету прикладної математики та
інформатики
Львівського національного університету
імені Івана Франка
(протокол № 6 від 21 грудня 2023 р.)



Завідувач кафедри

Роман ХАПКО

Силабус з навчальної дисципліни
«Машинне навчання на графах»,
що викладається в межах ОНП Прикладна математика
другого (магістерського) рівня вищої освіти для здобувачів з
спеціальності 113 – прикладна математика

Львів 2023 р.

Назва дисципліни	Машинне навчання на графах
Адреса викладання дисципліни	Головний корпус ЛНУ ім. І. Франка м. Львів, вул. Університетська 1
Факультет та кафедра, за якою закріплена дисципліна	Факультет прикладної математики та інформатики Кафедра обчислювальної математики
Галузь знань, шифр та назва спеціальності	11 – математика та статистика 113 – прикладна математика
Викладачі дисципліни	Музичук Юрій Анатолійович, кандидат фізико-математичних наук, доцент кафедри обчислювальної математики, Гарасим Ярослав Степанович, старший викладач кафедри обчислювальної математики (лабораторні заняття)
Контактна інформація викладачів	Yuriy.Muzychuk@lnu.edu.ua ; https://ami.lnu.edu.ua/employee/muzychuk-yuriy Iaroslav.Harasym@lnu.edu.ua ; https://ami.lnu.edu.ua/employee/harasym Головний корпус ЛНУ ім. І. Франка, каб. 262. м. Львів, вул. Університетська, 1
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекцій/лабораторних занять (за попередньою домовленістю).
Сторінка курсу	https://ami.lnu.edu.ua/course/mashynne-navchannia-na-hrafakh-prykladna-matematyka-1-9
Інформація про дисципліну	Дисципліна «Машинне навчання на графах» є дисципліною на вибір зі спеціальності 113 – прикладна математика для освітньо-наукової програми «Прикладна математика», яка викладається в 2-му семестрі в обсязі 4,5 кредити (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Курс «Машинне навчання на графах» поєднує теоретичні та практичні аспекти застосування методів машинного навчання до даних, структурованих у вигляді графів. Курс починається з фундаментальних понять та алгоритмів теорії графів, моделей на основі графів, та веде до більш складних тем, як-от вивчення представлення графів і нейронні мережі на графах. Курс далі вивчає класифікацію вузлів і ребер та генерацію графів. Також розглядаються питання графів знань та побудови рекомендаційних систем. Значна увага приділяється отриманню практичного досвіду під час лабораторних занять, де студенти реалізують вивчені концепції за допомогою мови Python і бібліотек NetworkX, PyTorch Geometric та Neo4J.
Мета та цілі дисципліни	Метою курсу є ґрунтовне ознайомлення студентів із сучасними задачами машинного навчання на графах та алгоритмами і технологіями для їхнього розв’язування. Завданням вивчення навчальної дисципліни є формування у студентів теоретичних знань в області машинного навчання на графах; здобуття студентами практичних навичок для роботи з графами знань; розв’язування прикладних задач за допомогою моделей нейронних мереж на графах.

<p>Література для вивчення дисципліни</p>	<p style="text-align: center;">Основна література</p> <ol style="list-style-type: none"> 1. Goodfellow I. Deep Learning / Ian Goodfellow, Yoshua Bengio, Aaron Courville. – MIT Press, 2016. – WWW: http://www.deeplearningbook.org 2. Hamilton W. Graph Representation Learning / William L. Hamilton. – McGill University, 2020. – WWW: https://www.cs.mcgill.ca/~wlh/grl_book/ 3. Labonne M. Hands-On Graph Neural Networks Using Python: Practical techniques and architectures for building powerful graph and deep learning apps with PyTorch / Maxime Labonne. – Packt Publishing, 2023. – 354p. 4. Negro A. Graph-Powered Machine Learning / Alessandro Negro. – Manning, 2021. – 496p. 5. Stamile C. Graph Machine Learning: Take graph data to the next level by applying machine learning techniques and algorithms / Claudio Stamile, Aldo Marzullo, Enrico Deusebio. - Packt Publishing, 2021. – 338p. 6. Van Bruggen R. Learning Neo4j / Rik Van Bruggen. – Packt Publishing, 2014. – 222p. 7. Fey M. Fast Graph Representation Learning with PyTorch Geometric / Matthias Fey, Jan Eric Lenssen, 2019. – WWW: https://arxiv.org/abs/1903.02428 <p style="text-align: center;">Додаткова література</p> <ol style="list-style-type: none"> 8. Easley D. Networks, Crowds, and Markets: Reasoning About a Highly Connected World / David Easley, Jon Kleinberg. – Cambridge University Press, 2010. – WWW: http://www.cs.cornell.edu/home/kleinber/networks-book/ 9. Barabasi A.-L. Network Science / Albert-Laszlo Barabasi. – WWW: http://networksciencebook.com/ 10. Robinson I. Graph Databases / Ian Robinson, Jim Webber, Emil Eifrem. – O’Reilly, 2015. – 220p. 11. Needham M. Graph Algorithms: Practical Examples in Apache Spark and Neo4j / Mark Needham, Amy E. Hodler. – O’Reilly, 2019. – 268p. 12. Gosnell D. The Practitioner's Guide to Graph Data / Denise Gosnell, Matthias Broecheler. – O’Reilly, 2020. – 250p. 13. Huyen C. Designing Machine Learning Systems: An Iterative Process for Production-Ready Applications / Chip Huyen. – O’Reilly, 2022. – 386p.
<p>Обсяг курсу</p>	<p>Загальний обсяг: 135 годин. Аудиторних занять: 48 год., з них 16 години лекцій та 32 години лабораторних робіт. Самостійної роботи: 87 год.</p>
<p>Очікувані результати навчання</p>	<p>Після завершення цього курсу студент буде:</p> <p>Знати:</p> <ul style="list-style-type: none"> - математичне формулювання поширених сучасних задач машинного навчання на графах, - найбільш продуктивні різновиди нейронних мереж на графах та способи їхнього застосування. <p>Вміти:</p> <ul style="list-style-type: none"> - на основі поставлених вимог спроектувати та реалізувати нейронну мережу на графах з використанням актуальної бібліотеки машинного навчання, - оцінити результати роботи побудованої мережі та запропонувати ідеї для покращення точності отриманих результатів. <p>Даний курс сприятиме формуванню та поглибленню таких соціальних, «м’яких» навичок (soft skills):</p> <ul style="list-style-type: none"> - <i>критичне мислення</i> – аналіз даних, розуміння складних алгоритмів та оцінка моделей машинного навчання стимулюють до аналітичного підходу до розв’язування поставлених задач.

	<ul style="list-style-type: none"> - <i>комунікація</i> – необхідність презентувати результати своїх індивідуальних проєктів та пояснювати тонкощі роботи алгоритмів вдосконалюють здатність чітко та ефективно доносити до аудиторії складну інформацію. - <i>прийняття рішень</i> – вибір необхідних алгоритмів, наборів даних та параметрів для ефективного розв’язування поставлених задач спонукатиме до прийняття проінформованих рішень, орієнтованих на факти. - <i>гнучкість</i> – машинне навчання, як сфера комп’ютерних наук, яка активно розвивається через швидку появу нових технологій та методик, вимагають гнучкості мислення та постійного навчання. 																																										
Ключові слова	Машинне навчання, глибокі нейронні мережі, нейронні мережі на графах, граfi знань, pytorch geometric, networkx, neo4j.																																										
Формат курсу	Очний. Проведення лекцій, лабораторних робіт і консультацій.																																										
Теми	Подано нижче у таблиці Схема курсу “Машинне навчання на графах”																																										
Підсумковий контроль, форма	Залік																																										
Пререквізити	Для вивчення курсу студенти потребують базових знань з <ul style="list-style-type: none"> - Основ програмування - Програмного забезпечення - Теорії імовірності та математичної статистики - Основ машинного навчання - Моделей глибинного машинного навчання 																																										
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Презентації, лекції (лекція-розповідь, лекція-бесіда), модульний контроль. Домашні та індивідуальні завдання																																										
Необхідне обладнання	Комп’ютер із програмним забезпеченням Visual Studio Code.																																										
Критерії оцінювання (окремо для кожного виду навчальної діяльності)	<p>Оцінювання проводиться за 100-бальною шкалою.</p> <table border="1"> <thead> <tr> <th colspan="2">Оцінка за шкалою ECTS</th> <th>Оцінка в балах</th> <th colspan="2">Оцінка за національною шкалою</th> <th rowspan="2">залік</th> </tr> <tr> <th colspan="2"></th> <th></th> <th>Екзамен, диференційований залік</th> <th></th> </tr> </thead> <tbody> <tr> <td>A</td> <td>Відмінно</td> <td>100 - 90</td> <td>Відмінно</td> <td>5</td> <td rowspan="4">зараховано</td> </tr> <tr> <td>B</td> <td>Дуже добре</td> <td>81- 89</td> <td rowspan="2">Добре</td> <td>4</td> </tr> <tr> <td>C</td> <td>Добре</td> <td>71 -80</td> <td>3</td> </tr> <tr> <td>D</td> <td>Задовільно</td> <td>61 - 70</td> <td></td> <td></td> </tr> <tr> <td>E</td> <td>Достатньо</td> <td>51- 60</td> <td></td> <td></td> <td rowspan="2">не зараховано</td> </tr> <tr> <td>FX (F)</td> <td>Незадовільно</td> <td>0 - 50</td> <td>Незадовільно</td> <td>2</td> </tr> </tbody> </table> <p>Впродовж семестру студент може отримати 100 балів. З них:</p> <ul style="list-style-type: none"> - домашні завдання: 30% семестрової оцінки; максимальна кількість балів 30 (3 завдання по 10 балів); - індивідуальне завдання: 40% семестрової оцінки; максимальна кількість балів 40; завдання оформлено у вигляді змагання на 	Оцінка за шкалою ECTS		Оцінка в балах	Оцінка за національною шкалою		залік				Екзамен, диференційований залік		A	Відмінно	100 - 90	Відмінно	5	зараховано	B	Дуже добре	81- 89	Добре	4	C	Добре	71 -80	3	D	Задовільно	61 - 70			E	Достатньо	51- 60			не зараховано	FX (F)	Незадовільно	0 - 50	Незадовільно	2
Оцінка за шкалою ECTS		Оцінка в балах	Оцінка за національною шкалою		залік																																						
			Екзамен, диференційований залік																																								
A	Відмінно	100 - 90	Відмінно	5	зараховано																																						
B	Дуже добре	81- 89	Добре	4																																							
C	Добре	71 -80		3																																							
D	Задовільно	61 - 70																																									
E	Достатньо	51- 60			не зараховано																																						
FX (F)	Незадовільно	0 - 50	Незадовільно	2																																							

платформі Kaggle; для нього встановлено терміни здачі. Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку (кожного лабораторного заняття після терміну здачі на 4 бали менше);

- контрольні заміри (модулі): 30% семестрової оцінки; максимальна кількість балів 30 (2 тестові модулі в середовищі MS Teams по 15 балів);

Підсумкова максимальна кількість балів 100.

Критерії оцінювання домашніх завдань:

10 балів – студент повністю виконав умови завдання; алгоритм реалізовано правильно; продемонстровано приклади використання; оформлено належним чином за допомогою Jupyter Notebook з коментарями та візуалізаціями;

5-9 балів – завдання реалізовано повністю, але з незначними помилками; оформлено не в повній відповідності до вимог: частково відсутні візуалізації або коментарі;

1-4 бали – завдання реалізовано частково або із значними помилками; відсутні приклади використання; оформлено неналежним чином;

0 балів - студент не виконав завдання.

Критерії оцінювання індивідуальних завдань:

40 балів – студент повністю виконав умови завдання; проведено попередній аналіз даних та побудовано візуалізації; здійснено вибір та порівняння архітектур для розв'язування завдання; підбрано оптимальні гіперпараметри для алгоритму; проведено аналіз та візуалізацію результатів; оформлено за допомогою Jupyter Notebook та збережено в системі контролю версій GitHub або на платформі Kaggle;

32-39 балів – студент повністю виконав умови завдання, але з незначними помилками або зауваженнями щодо оформлення роботи;

24-31 балів – студент виконав завдання частково, але при цьому зміг побудувати кінцеву модель, яка показує високий результат; відсутній один з обов'язкових пунктів роботи: дослідження та візуалізації даних, підбору гіперпараметрів алгоритму або аналізу результатів;

16-23 балів – студент виконав завдання частково; алгоритм реалізовано з помилками; кінцева модель характеризується низьким результатом точності; пропущені два або більше етапів роботи;

8-15 балів – студент виконав завдання частково; проведено лише завантаження даних та їх поверхневе дослідження; не побудована кінцева модель для задачі;

1-8 балів – студент виконав завдання частково з грубими помилками; присутні помилки виконання програми; відсутня кінцева модель для задачі;

0 балів – студент не виконав завдання.

Критерії оцінювання контрольних замірів (модулів):

1 бал – відповідь на запитання правильна;

0 балів – відповідь на завдання неправильна.

Неформальна освіта: Додаткові бали за дисципліну можна отримати, пройшовши один або більше курсів за даною тематикою на популярних онлайн платформах. Запропоновані курси на вибір (інші курси можливі лише після підтвердження від лектора):

- <https://www.skillsoft.com/course/gnns-an-introduction-to-graph-neural->

	<p>networks-9541eea6-572c-438a-b64a-286ec80553d3</p> <ul style="list-style-type: none"> - https://www.udemy.com/course/graph-neural-network/ - https://www.udemy.com/course/the-ultimate-graph-neural-network-course/ - https://www.udemy.com/course/graphgeneration/ - https://www.edx.org/learn/python/imt-advanced-algorithmics-and-graph-theory-with-python - https://www.coursera.org/learn/graphs - https://www.coursera.org/learn/big-data-graph-analytics - https://www.coursera.org/learn/algorithms-graphs-data-structures <p>В такому випадку кожний курс оцінюватиметься по 30 балів. Необхідні вимоги:</p> <ol style="list-style-type: none"> 1. Курс завершено успішно (з результатом більше, ніж 80%); 2. Доступний іменний сертифікат про завершення; 3. Продемонстровано коротку презентацію-звіт про проходження курсу на лабораторному занятті. <p>Академічна доброчесність: Очікується, що роботи студентів будуть їх оригінальними дослідженнями чи міркуваннями. Відсутність посилань на використані джерела, фабрикування джерел, списування, втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в письмовій роботі студента є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману.</p> <p>Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання домашніх та індивідуальних завдань, передбачених курсом.</p> <p>Література. Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих.</p> <p>Політика виставлення балів. Враховуються бали, набрані на лабораторних заняттях (здача домашніх та індивідуальних завдань) та під час контрольних замірів. При цьому обов'язково враховуються присутність на заняттях та активність студента під час лабораторного заняття; недопустимість пропусків та запізень на заняття; користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін.</p> <p>Жодні форми порушення академічної доброчесності не толеруються.</p>
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

Схема курсу «Машинне навчання на графах»

Тиж.	Тема, план, короткі тези	Форма діяльності (заняття)	Література, Ресурси в інтернеті	Завдання, год.	Термін виконання
1	Тема 1. Вступ (<i>Означення машинного навчання. Типи машинного навчання. Машинне навчання на графах. Прикладні задачі</i>)	Лекція (2 год.)	[1-5], [8-9]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
	Тема 1. Робота з бібліотеками PyTorch, PyG, NetworkX (<i>Робота з графами. Багаторівневий перцептрон. Модельні приклади</i>)	Лабораторне (2 год.)	[7]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
2	Тема 2. Представлення даних за допомогою графів. (<i>Ознайомлення з бібліотекою Neo4j</i>) Домашнє завдання №1. Побудова та аналіз бази даних за допомогою Neo4j	Лабораторне (2 год.)	[6], [10]	Виконання завдання №1 (6 год)	3 тижні
3	Тема 3. Інженерія характеристик для машинного навчання на графах (<i>Характеристики вершин. Графлети. Передбачення ребер. Характеристики ребер</i>)	Лекція (2 год.)	[2], [8-9]	Опрацювання лекційного матеріалу (4 год.)	1 тиждень
	Тема 3. Представлення вершин (<i>Побудова вбудовування для вузлів та графів</i>)	Лабораторне (2 год.)	[2-3], [8-9]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
4	Тема 4. Обхід графа. Пошук найкоротшого шляху (<i>Алгоритми BFS, DFS, Dijkstra</i>)	Лабораторне (2 год.)	[3], [8]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
5	Тема 5. Нейронні мережі на графах (<i>Основи НМ. Ідея глибокого навчання на графах. Згорткові мережі на графах</i>)	Лекція (2 год.)	[2], [8-9], [11]	Опрацювання лекційного матеріалу (6 год.)	1 тиждень
	Тема 5. Реалізація нейронних мереж на графах засобами PyTorch Geometric (<i>Передбачення властивостей вершин</i>) Здача домашнього завдання №1.	Лабораторне (2 год.)	[2-5], [7]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень

6	<p>Тема 6. Реалізація нейронних мереж на графах засобами PyTorch Geometric (<i>Передбачення властивостей графів. Видача індивідуального завдання</i>)</p> <p>Індивідуальне завдання №1. Розв'язування задачі класифікації на графах. Доповідь про результати.</p>	Лабораторне (2 год.)	[2-5], [7]	Підготувати доповідь (індивідуальне завдання №1) (10 год)	6 тижнів
7	<p>Тема 7. Варіації нейронних мереж на графах та особливості тренування (GraphSAGE, Graph Attention Networks)</p>	Лекція (2 год.)	[2-5]	Опрацювання лекційного матеріалу (4 год.)	1 тиждень
	<p>Тема 7. Реалізація GraphSAGE за допомогою PyG (<i>Побудова архітектури. Тренувальний цикл. Оцінка результатів</i>)</p>	Лабораторне (2 год.)	[2-5], [7]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
8	<p>Тема 8. Реалізація GAN за допомогою PyG (<i>Побудова архітектури. Тренувальний цикл. Оцінка результатів</i>)</p>	Лабораторне (2 год.)	[7]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
9	<p>Тема 9. Поширення міток в графах (<i>Підхід «Виправи та Згладь»</i>)</p>	Лекція (2 год.)	[2-4]	Опрацювання лекційного матеріалу (3 год.)	1 тиждень
	<p>Тема 9. Гетерогенні графи (<i>Реляційні згорткові мережі на графах</i>)</p> <p>Контрольний замір №1.</p>	Лабораторне (2 год.)	[2-4]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
10	<p>Тема 10. Реалізація гетерогенних графів засобами PyG та DeepSNAP</p> <p>Домашнє завдання №2. Побудова гетерогенних графів.</p>	Лабораторне (2 год.)	[7]	Виконання завдання №2 (8 год)	3 тижні
11	<p>Тема 11. Графи знань (<i>Задача доповнення графів знань. TransE, TransR, DistMult, ComplEx методи</i>)</p>	Лекція (2 год.)	[3-6]	Опрацювання лекційного матеріалу (6 год.)	1 тиждень
	<p>Тема 11. Прийняття рішень на основі графів знань (<i>Види запитів. Вбудовування запиту</i>)</p>	Лабораторне (2 год.)	[3-6]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень

12	<p>Тема 12. Будівельні блоки графів (<i>Підграфи. Мотиви. Розпізнавання мотивів</i>)</p> <p>Здача індивідуального завдання №1.</p>	Лабораторне (2 год.)	[2-4]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
13	<p>Тема 13. Рекомендаційні системи (<i>Постановка задачі. Методи оцінки. Моделі на основі вбудувань</i>)</p>	Лекція (2 год.)	[2-6]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
	<p>Тема 13. Використання нейронних мереж на графах для розв'язування задачі рекомендації (<i>Колаборативне фільтрування. Моделі NGCF, LightGCN, PinSage</i>)</p> <p>Здача домашнього завдання №2.</p> <p>Домашнє завдання №3. Розв'язування задачі рекомендації.</p>	Лабораторне (2 год.)	[2-6]	Виконання завдання №3 (8 год)	3 тижні
14	<p>Тема 14. Побудова моделі LightGCN засобами PyG (<i>Робота з даними. Ініціалізація. Повідомлення та агрегація. Штрафна функція. Оцінка результатів</i>)</p>	Лабораторне (2 год.)	[7]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
15	<p>Тема 15. Глибокі генеративні моделі на графах (<i>Постановка задачі. Графи як послідовності. GraphRNN</i>)</p>	Лекція (2 год.)	[2-4]	Опрацювання лекційного матеріалу (6 год.)	1 тиждень
	<p>Тема 15. Атаки на нейронні мережі на графах (<i>Здача індивідуального завдання</i>)</p>	Лабораторне (2 год.)	[2]	Опрацювання лекційного матеріалу (2 год.)	1 тиждень
16	<p>Тема 16. Пояснюваність нейронних мереж на графах.</p> <p>Здача домашнього завдання №3.</p> <p>Контрольний замір №2.</p>	Лабораторне (2 год.)	[2]		