

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Чернівецький національний університет імені Юрія Федьковича
Факультет математики та інформатики
Кафедра прикладної математики та інформатики

ЗАТВЕРДЖЕНО

На засіданні

кафедри прикладної математики та ІТ

Чернівецького національного університету

імені Юрія Федьковича

(протокол № 1 від 30 серпня 2021 р.)

Завідувач кафедри _____ Я.Й.Бігун

Силабус з навчальної дисципліни
«Методи Data Science»,
що викладається в межах ОПШ
«Прикладна математика»
другого (магістерського) рівня вищої освіти для здобувачів з
спеціальності 113 «Прикладна математика»

Чернівці 2021 р.

	<p align="center">Силабус навчальної дисципліни « Методи Data Science» Галузь знань: 11 «Математика» Спеціальність: 113 «Прикладна математика»</p>
Адреса викладання дисципліни	м. Чернівці, вул. Університетська, 28
Факультет та кафедра, за якою закріплена дисципліна	Факультет математики та інформатики Кафедра прикладної математики та ІТ
Галузь знань, шифр та назва спеціальності	11 «Математика» 113 «Прикладна математика»
Викладачі дисципліни	Мельник Галина Василівна, к. е. н., доцент кафедри прикладної математики та ІТ https://amit.chnu.edu.ua/pro-kafedru/personalii/melnyk-halyna-vasylivna/
Контактна інформація викладачів	Електронні скриньки: g.melnik@chnu.edu.ua Місце знаходження: м. Чернівці, вул. Університетська, 28, ауд. 26 (кафедра прикладної математики та ІТ)
Консультації з питань навчання по дисципліні відбуваються	Згідно розкладу (вул. Університетська, 28, ауд. 26) Консультації в день проведення лекцій/лабораторних занять (за попередньою домовленістю). Можливі он-лайн консультації через Zoom, Google Meet. Для погодження часу он-лайн консультацій слід писати на електронну пошту викладача.
Сторінка курсу	Платформа Google Classroom: https://classroom.google.com/c/MTYwODk0ODIwNjQw?cjc=6nuuqmd
Інформація про дисципліну	Дисципліна «Методи Data Science» є логічним продовженням вивчення спеціальних дисциплін, здобуті знання знадобляться в подальшій практичній діяльності майбутнім фахівцям галузі ІТ. У курсі розглянуто основні математичні методи машинного навчання, аналізу та прогнозування інформаційних процесів та явищ, типи та етапи побудови та розв'язування математичних моделей, розглядаються практичні приклади економіко-математичних задач та методи їх реалізації.
Коротка анотація дисципліни	Дисципліна «Методи Data Science» є нормативною дисципліною зі спеціальності 113 «Прикладна математика» для освітньої програми «Прикладна математика», яка викладається в I семестрі в обсязі 5 кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Мета та цілі дисципліни	Мета вивчення дисципліни «Методи Data Science»: формування системи знань з методології та інструментарію побудови і використання різних типів економіко-математичних моделей. Основними завданнями дисципліни «Методи Data Science» є вивчення основних принципів та інструментарію постановки задач, побудови економіко-математичних моделей, методів їх розв'язування та аналізу з метою використання в подальшій діяльності.

<p>Література для вивчення дисципліни</p>	<p>Основна література:</p> <ol style="list-style-type: none"> 1. Agrawal R., Imielinski T., Swami A. (1993). Mining Associations between Sets of Items in Massive Databases. In Proc. of the 1993 ACM-SIGMOD Int’l Conf. on Management of Data, 207–216. 2. Agrawal R., Srikant R. (1994). "Fast Discovery of Association Rules", In Proc. of the 20th International Conference on VLDB, Santiago, Chile, September 1994. 3. Bartholomew, D. J. (1987). Latent variable models and factor analysis. Oxford University Press. 7. Basilevsky, A. (1994). Statistical Factor Analysis and Related Methods: Theory and Applications. Wiley. Bengio, Y. and LeCun, Y. (2007). Scaling learning algorithms towards AI. In Large Scale Kernel Machines. 8. Bertsekas, D. P. and Tsitsiklis, J. (1996). Neuro-Dynamic Programming. Athena Scientific. 9. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer. 10. Bottou, L. (1998). Online algorithms and stochastic approximations. In D. Saad, editor, Online Learning in Neural Networks. Cambridge University Press, Cambridge, UK. 11. Boyd, S. and Vandenberghe, L. (2004). Convex Optimization. Cambridge University Press, New York, NY, USA. <p>Інтернет-джерела:</p> <ol style="list-style-type: none"> 1. Офіційний сайт Державної служби статистики [Електронний ресурс]. – Режим доступу : www.ukrstat.gov.ua. 2. Офіційний сайт проекту «Doing Business» [Електронний ресурс]. – Режим доступу : http://www.doingbusiness.org. 3. Офіційний сайт Служби статистики Європейського союзу [Електронний ресурс]. – Режим доступу : http://ec.europa.eu/eurostat.
<p>Тривалість курсу</p>	<p>150 год.</p>
<p>Обсяг курсу</p>	<p>60 годин аудиторних занять. З них 30 годин лекцій, 30 годин лабораторних робіт занять та 90 годин самостійної роботи</p>

<p>Очікувані результати навчання</p>	<p>Після завершення цього курсу студент буде :</p> <p>а) знати:</p> <ul style="list-style-type: none"> - основні наукові підходи та сучасні концепції розвитку машинного навчання; - проблеми розвитку машинного навчання та шляхи їх оптимального розв'язання; - можливості використання машинного навчання для аналізу, та прогнозу формування ефективної діяльності економіки; - пріоритетні дослідження в машинному навчанні; <p>б) уміти</p> <ul style="list-style-type: none"> - аналізувати та розробляти алгоритми для розв'язання задач за допомогою методів машинного навчання; - розробляти програмні застосування з використанням методів машинного навчання засобами візуального проектування; - використовувати засоби математичних програм та пакетів; - оцінювати похибки та проаналізувати отриманий розв'язок; - будувати математичні моделі економічних задач; - оцінювати параметри економетричної моделі; - перевіряти адекватність економетричних моделей та їх параметрів; - виконувати точковий та інтервальний прогноз. <p>в) загальні компетентності:</p> <ul style="list-style-type: none"> - ЗК2. Здатність до абстрактного мислення, аналізу та синтезу. Вміння застосовувати знання у практичній діяльності. - ЗК4. Здатність вести професійну і науково-дослідну діяльність, зокрема у міжнародному середовищі. - ЗК5. Здатність спілкуватися державною мовою як усно, так і письмово. Здатність спілкуватися іноземною мовою; - ЗК7. Здатність бути критичним і самокритичним. Здатність генерувати нові ідеї і методи їх реалізації. <p>г) фахові компетентності:</p> <ul style="list-style-type: none"> - ФК10. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач. <p>д) та отримуються наступні програмні результати навчання:</p> <ul style="list-style-type: none"> - ПРН2. Здійснювати збір, систематизацію та аналіз науково-технічної інформації з питань професійної діяльності. Вміти використовувати поглиблені професійні знання та практичні навички для розв'язування дослідницьких задач; - ПРН9. Створювати та програмно реалізовувати алгоритми розв'язання задач, розробляти системне та прикладне забезпечення інформаційних систем і технологій з урахуванням вимог до його якості, надійності, виробничих характеристик; - ПРН11. Вміти проектувати архітектуру системи з великими обсягами даних.
<p>Ключові слова</p>	<p>Методи машинного навчання, система рівнянь, методи аналізу та прогнозування, робота з даними, робота з блоками даних, методи роботи з матрицями, оптимізація.</p>
<p>Формат курсу</p>	<p>Очний</p>

	Проведення лекцій, лабораторних робіт та консультації для кращого розуміння тем. Викладання навчальної дисципліни передбачає поєднання традиційних форм аудиторного навчання з елементами електронного навчання, в якому використовуються спеціальні інформаційні технології, такі як комп'ютерна графіка, аудіо та відео, інтерактивні елементи, онлайн консультування і т.п.
Теми	Подано у формі Схеми курсу
Підсумковий контроль, форма	Іспит в кінці семестру/ (письмовий тест). Оцінка складається із кількості балів нарахованих за: здачу лабораторних робіт, виконання самостійних робіт та індивідуального завдання, складання екзаменаційних тестів. Методи контролю: спостереження за навчальною діяльністю здобувачів вищої освіти, усне опитування, письмовий контроль, тестовий контроль, виконання навчальних та індивідуальних завдань.
Пререквізити	Для вивчення курсу студенти потребують базових знань з дисциплін «Теорія імовірності та математична статистика», «Теорія алгоритмів», «Методи оптимізації», «Програмування мовою Python» для: а) сприйняття категоріального апарату програмування математичних методів; б) оперування методами професійної дискусії для формування власної аргументованої позиції.
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Презентація, лекція-бесіда, лекція-візуалізація, колаборативне навчання (форми – групові проекти, спільні розробки і т. д.), проектно-орієнтоване навчання, навчальна дискусія, мозкова атака, кейс-метод, демонстрування, самостійна робота, лабораторні роботи, метод порівняння, метод узагальнення, метод конкретизації, метод виокремлення основного, обговорення, робота над помилками,
Необхідне обладнання	Вивчення курсу не потребує використання спеціального програмного забезпечення, крім загально вживаних програм і операційних систем. Для лектора - персональний комп'ютер, проектор.

**Критерії оцінювання
(окремо для кожного
виду навчальної
діяльності)**

Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням:

- лабораторні/самостійні тощо: 60% семестрової оцінки; максимальна кількість балів – 60;
- контрольні заміри (модулі): 10% семестрової оцінки; максимальна кількість балів – 10;
- іспит: 30% семестрової оцінки (максимальна кількість балів – 30).

Підсумкова максимальна кількість балів – 100.

Академічна доброчесність: Очікується, що роботи студентів будуть їх оригінальними дослідженнями чи міркуваннями. Списування, втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в практичній (письмовій) роботі студента є підставою для її не зарахування викладачем, незалежно від масштабів плагіату чи обману.

Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції і лабораторні заняття курсу. Студенти мають інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися усіх строків визначених для виконання усіх видів робіт, передбачених курсом.

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.

Політика виставлення балів. Враховуються бали набрані на поточному тестуванні, самостійній роботі та бали підсумкового тестування. При цьому обов'язково враховуються присутність на заняттях та активність студента під час лабораторного заняття; недопустимість пропусків та запізнь 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.

Жодні форми порушення академічної доброчесності не толеруються.

<p>Питання до заліку чи екзамену.</p>	<ol style="list-style-type: none"> 1. Регресія. 2. Лінійна регресія. 3. Дерева рішень і випадковий ліс. 4. Класифікація. Логістична регресія. 5. Метод опорних векторів. 6. Метод k-найближчих сусідів. 7. Байєсова класифікація. 8. Дерева рішень та ансамблеві методи. 9. Кластеризація. Ієрархічна кластеризація. Кластеризація на основі k-means. 10. Асоціативні правила. 11. Побудова асоціативних правил. 12. Алгоритми APRIORI та ECLAT. 13. Обробка природної мови. Модель «мішок слів». 14. Моделі з урахуванням семантики 15. Штучні нейронні. 16. Теоретичні основи нейронних мереж. 17. Нейронні мережі в задачах апроксимації. 18. Нейронні мережі в задачах прогнозування. 19. Нейронні мережі в задачах класифікації. 20. Нейронні мережі в задачах кластеризації: карти Кохонена 21. Глибоке навчання. 22. Рекурентні нейронні мережі. 23. Згорткові нейронні мережі
<p>Опитування</p>	<p>Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.</p>

Схема курсу

Тиждень / дата / год.	Тема, план, короткі тези	Форма діяльності (заняття)	Література. Ресурси в інтернеті	Завдання, год.	Термін виконання
1	2	3	4	5	6
Тиждень 1 4 год.	Тема 1. Основні поняття інтелектуального аналізу даних Завдання аналізу даних.	Лекція Лабораторна робота	Осн. [1-4]. Інт. [2, 3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 2 4 год.	Тема 2. CRISP-DM методологія. Підготовка даних.	Лекція Лабораторна робота	Осн. [1-4]. Інт. [2, 3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 3 4 год.	Тема 3. Регресія. Лінійна регресія.	Лекція	Осн. [4-5]. Інт. [1,2].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 4 4 год.	Тема 4. Древа рішень і випадковий ліс.	Лабораторна робота	Осн. [4-5]. Інт. [1,2].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 5 4 год.	Тема 5. Класифікація. Логістична регресія. Метод опорних векторів.	Лекція	Осн. [5-8]. Дод. [5]. Інт. [1].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 6 4 год.	Тема 6. Метод k-найближчих сусідів. Байєсова класифікація. Древа рішень та ансамблеві методи.	Лабораторна робота	Осн. [5-8]. Дод. [5]. Інт. [1].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 7 4 год.	Тема 7.Кластеризація. Ієрархічна кластеризація.	Лекція	Осн. [5-8]. Дод. [5]. Інт. [1].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 8 4 год.	Тема 8.Кластеризація на основі k-means.	Лабораторна робота	Осн. [5-8]. Дод. [5]. Інт. [1].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом

Тиждень 9 4 год.	Тема 9. Асоціативні правила. Побудова асоціативних правил.	Лекція	Осн. [10,13]. Дод. [4,5]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 10 4 год.	Тема 10. Алгоритми APRIORI та ECLAT.	Лабораторна робота	Осн. [10,13]. Дод. [4,5]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 11 4 год.	Тема 11. Обробка природної мови. Модель «мішок слів».	Лекція	Осн. [10,13]. Дод. [4,5]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 12 4 год.	Тема 12. Обробка природної мови. Моделі з урахуванням семантики.	Лабораторна робота	Осн. [10,13]. Дод. [4,5]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 13 4 год.	Тема 13. Штучні нейронні. Теоретичні основи нейронних мереж. НМ в задачах апроксимації та прогнозування.	Лекція	Осн. [15,16]. Дод. [7,8]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 14 4 год.	Тема 14. НМ в задачах класифікації. НМ в задачах кластеризації: карти Кохонена	Лабораторна робота	Осн. [15,16]. Дод. [7,8]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом
Тиждень 15 4 год.	Тема 15. Глибоке навчання. Рекурентні нейронні мережі. Згорткові нейронні мережі	Лекція	Осн. [15,16]. Дод. [7,8]. Інт. [3].	Опрацювати лекційний матеріал, виконання завдань до лабораторного заняття 4 год.	До проведення наступного аудиторного заняття за розкладом

Викладач _____ Г.В.Мельник