

Чернівецький національний університет імені Юрія Федьковича
(повне найменування вищого навчального закладу)

Факультет математики та інформатики

(назва інституту / факультету)

Кафедра математичного моделювання

(назва кафедри)

СИЛАБУС

навчальної дисципліни

Аналіз даних

(назва навчальної дисципліни)

Обов'язкова

(вказати: обов'язкова / вибіркова)

Освітньо-професійна програма Інформаційні технології та управління проектами

(назва програми)

Спеціальність 122 – Комп'ютерні науки

(вказати: код, назва)

Галузь знань 12 – Інформаційні технології

(вказати: шифр, назва)

Рівень вищої освіти перший (бакалаврський)

(вказати: перший бакалаврський/другий магістерський)

Факультет математики та інформатики

(назва факультету/інституту, на якому здійснюється підготовка фахівців за вказаною освітньо-професійною програмою)

Мова навчання українська

(вказати: на якій мові читається дисципліна)

Розробники: Черевко Ігор Михайлович, професор каф. мат. мод., доктор. фіз.-мат. наук,
професор

(П.І.Б. авторів, посада, науковий ступінь, вчене звання)

Профайл викладача: <http://matmod.fmi.org.ua/pro-kafedru/spivrobityky/cherevko-igor-myhailovych/>

Контактний тел.: (0372) 58-48-25

E-mail: i.cherevko@chnu.edu.ua

Сторінка курсу в Moodle: <https://moodle.chnu.edu.ua/course/view.php?id=4599>

Консультації: Зазначте формат і розклад проведення консультацій
Онлайн-консультації: среда з 14-20-15-20.
Очні консультації: за попередньою домовленістю.

1. Анотація дисципліни (призначення навчальної дисципліни).

Це є базовий курс для майбутніх аналітиків. Завдяки отриманим знанням вони зможуть проводити первинний статистичний аналіз даних, висувати гіпотези та оцінювати їх на основі зібраних даних. Прикладний аналіз даних служить базою прийняття управлінських рішень. На сучасному етапі його значимість значно зросла у зв'язку із величезним обсягом інформації, який необхідно проаналізувати. Аналіз та застосування згаданих підходів і складає основний зміст даного курсу.

2. Мета навчальної дисципліни. Задачі аналізу емпіричних даних займають центральне місце при проведенні експериментальних досліджень в будь-якій області знань. Аналіз даних складає алгоритмічну основу автоматизованих систем (САПР, АСУП, АСУТП та ін.). Метою та завданнями дисципліни «Аналіз даних» є систематичне вивчення задач, що пов'язані із етапами обробки даних, побудовою та оцінкою математичних моделей експериментальних даних, застосування сучасних ППП обробки статистичних даних на комп'ютері.

Дисципліна формує такі компетентності за ОП

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК10. Здатність бути критичним і самокритичним.

ЗК11. Здатність приймати обґрунтовані рішення.

ФК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.

ФК6. Здатність до системного мислення, застосування методології системного аналізу для дослідження складних проблем різної природи, методів формалізації та розв'язування системних задач, що мають суперечливі цілі, невизначеності та ризики.

ФК7. Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів.

ФК11. 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.

3. Пререквізити. Обчислювальні методи, Програмування, Системи та методи прийняття рішень, Теорія ймовірностей та математична статистика, Диференціальні рівняння, Аналіз даних та статистичне виведення на мові R (онлайн-курс платформи **Prometheus**).

4. Результати навчання

Завдання дисципліни. Знання і досвід, набуті в цьому курсі, будуть корисними в майбутній практичній діяльності студентів при моделюванні на ЕОМ різноманітних прикладних процесів та аналізі експериментальних даних.

У результаті вивчення навчальної дисципліни студент повинен

знати:

- основні теоретичні і практичні принципи, форми, методи, прийоми статистичної обробки чисельних даних, зокрема результатів експерименту із застосуванням ЕОМ;

вміти:

- застосовувати статистичні методи для розв'язання задач практичного змісту при моделюванні на ЕОМ різноманітних прикладних процесів та аналізі експериментальних даних.

Наведені результати навчання за відповідною дисципліною співвідносяться із такими **програмами результатами навчання:**

ПРН1. Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.

ПРН3. Використовувати знання закономірностей випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей.

ПРН8. Використовувати методологію системного аналізу об'єктів, процесів і систем для задач аналізу, прогнозування, управління та проектування динамічних процесів в макроекономічних, технічних, технологічних і фінансових об'єктах.

ПРН12. Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.

5. Опис навчальної дисципліни

5.1. Загальна інформація

Форма навчання	Рік підготовки	Семестр	Кількість		Кількість годин						Вид підсумкового контролю
			кредитів	годин	лекції	практичні	семінарські	лабораторні	самостійна робота	індивідуальні завдання	
Денна	4	8	3	90	22	–	–	22	46	–	екзамен

5.2. Дидактична карта навчальної дисципліни

Назви змістових модулів і тем	Кількість годин												
	денна форма						Заочна форма						
	усього	у тому числі					усього	у тому числі					
		Л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.	
1	2	3	4	5	6	7	8	9	10	11	12	13	
Змістовий модуль 1. Попередня обробка даних													
Тема 1. Задачі попередньої обробки даних	6	2				4							
Тема 2. Моделювання псевдовипадкових чисел.	10	2		2		6							
Тема 3. Емпіричний розподіл даних.	8	2		2		4							
Тема 4. Дослідження законів розподілу експериментальних даних.	14	4		4		6							
Разом за ЗМ 1	38	10		8		20							

Змістовий модуль 2. Аналіз експериментальних даних та статистичні моделі											
Тема 1. Лінійні регресійні моделі.	10	2		4		4					
Тема 2. Нелінійні багатофакторні регресійні моделі.	10	2		2		6					
Тема 3. Планування експерименту в задачах ідентифікації об'єктів.	6	2				4					
Тема 4. Аналіз даних в системі "Statistica". Опис, структура та головні функціональні модулі системи Statistica.	14	4		4		6					
Тема 5. Статистичний аналіз часових рядів в системі Statistica	12	2		4		6					
Разом за ЗМ 2	52	12		14		26					
Усього годин	90	22		22		46					

5.3. Самостійна робота

Самостійна робота студентів складає 46 годин. Розподіл самостійної роботи за видами навчальних робіт:

- 1) підготовка до лекційних занять – 8 годин;
- 2) підготовка до лабораторних занять та їх виконання – 20 годин;
- 3) самостійне опрацювання додаткового матеріалу – 8 годин;
- 4) підготовка до модульних контрольних робіт – 4 години;
- 5) підготовка до іспиту – 6 годин.

Зміст завдань для самостійної роботи

№	Назва теми
1	Особливості побудови програмного забезпечення для статистичної обробки даних. Типи статистичних пакетів і загальні вимоги до них.
2	Методика обчислення числових характеристик експериментальних даних. Додаткові відомості із теорії ймовірностей та математичної статистики. Оцінка числових характеристик та параметрів випадкових величин та процесів.
3	Регресійний аналіз простіших поліноміальних моделей. Особливості регресійного аналізу при порушенні базових припущень. Визначення показників зв'язку при криволінійній залежності.

4	Методи перевірки ортогональності планів. Побудова неповних факторних планів, генератор планів. Схеми проведення імітаційного експерименту
---	---

6. Система контролю та оцінювання

Види та форми контролю

1. Поточний (усне опитування)
2. Модульний (контрольні роботи, лабораторні роботи)
3. Підсумковий (екзамен)

Засоби оцінювання:

- контрольні роботи;
- сертифікати курсу «Аналіз даних та статистичне виведення на мові R» на платформі PROMETHEUS;
- аналітичні звіти про виконання індивідуальних завдань та самостійної роботи.

Критерії оцінювання результатів навчання з навчальної дисципліни

Система оцінювання рівня навчальних досягнень ґрунтується на принципах ECTS та є накопичувальною. На протязі семестру студенти виконують дві модульні контрольні роботи та 4 лабораторних роботи. Кожна модульна та лабораторна робота оцінюється максимум 10 балами. Отримання сертифікату курсу «Аналіз даних та статистичне виведення на мові R» на платформі PROMETHEUS може замінити виконання лабораторних робіт і оцінюється в 40 балів. Активна та постійна робота на протязі семестру оцінюється до 10 балів. Підсумковим контролем з дисципліни є іспит у формі тесту, максимальна оцінка – 30 балів. Таким чином, максимальна оцінка за курс – 100 балів.

Оцінка виставляється за результатами суми балів набраних на змістовних модулях під час семестру та підсумковому модулі (екзамені) згідно таблиці

Шкала оцінювання: національна та ECTS

Оцінка за національною шкалою	Оцінка за шкалою ECTS	
	Оцінка (бали)	Пояснення за розширеною шкалою
Відмінно	A (90-100)	відмінно
Добре	B (80-89)	дуже добре
	C (70-79)	добре
Задовільно	D (60-69)	задовільно
	E (50-59)	достатньо
Незадовільно	FX (35-49)	(незадовільно) з можливістю повторного складання
	F (1-34)	(незадовільно) з обов'язковим повторним курсом

7. Політика освітнього процесу

Студенти зобов'язані своєчасно та якісно виконувати всі отримані завдання і акуратно їх оформлювати у вигляді звіту з наданням програмного коду у випадку чисельних розрахунків. За необхідністю з метою з'ясування всіх незрозумілих під час самостійної та індивідуальної роботи питань, відвідувати консультації викладача. Кожний студент зобов'язаний дотримуватися принципів академічної доброчесності. Виконаний студентом не свій варіант завдання не оцінюється. Складання (перескладання) екзамену проводиться за встановленим деканатом розкладом.

8. Рекомендована література

Основна

1. Бахрушин В.Є. Методи аналізу даних : навчальний посібник для студентів. – Запоріжжя : КПУ, 2011. – 268 с.
2. Грицюк П.М., Остапчук О.П. Аналіз даних: Навчальний посібник.– Рівне: НУВГП, 2008. – 218 с.
3. Kahaner D., Moler C., Nash S. Numerical Methods and Software. Prentice-Hall, 1989. 508 p.
4. Майборода Р.Є. "Комп'ютерна статистика". ВПЦ "Київський університет", 2019. 589 с.
<http://probability.univ.kiev.ua/userfiles/mre/cscolor.pdf>
5. Пашинський В.А. Статистичні методи в інженерних дослідженнях. Навчальний посібник для здобувачів вищої освіти з інженерних спеціальностей. – Кропивницький: ЦНТУ, 2020. – 106 с.
6. Слабоспицький А.С. Аналіз даних. Попередня обробка. Навчальний посібник. – К.: Київський національний університет імені Тараса Шевченка (КНУ), 2001. – 52 с.
7. Турчин В.М. Теорія ймовірностей і математична статистика. Основні поняття, приклади, задачі. Підручник.– Дніпропетровськ: ІМА-прес, 2014. –556 с.
8. Аналіз даних : Частина 1. Попередня обробка експериментальних даних. Методичні рекомендації та завдання для лабораторних робіт / Укл.: І.М. Черевко. – Чернівці: 2023. –27 с.

Додаткова

1. Benjamin S. Data Analytics: A Comprehensive Beginner's Guide To Learn About The Realms Of Data Analytics From A-Z Independently Published. – 2020. – 154 p.
2. Robinson Eward L. Data Analysis for Scientists and Engineers. Princeton University Press.– 2016.– 408 p.
3. D. E. Knuth. The Art of Computer Programming, v. 2, Seminumerical Algorithms. Stenford University Press.– 1997.– 774 p.
4. Forsythe, G. E. / Malcolm, M. A. / Moler, C. B., Computer Methods for Mathematical Computations. Englewood Cliffs, New Jersey. 1977. – 259 p.
5. Аналіз даних – Prometheus.– <https://prometheus.org.ua> > [dataanalysis](#)

9. Інформаційні ресурс

<https://moodle.chnu.edu.ua/course/view.php?id=4599>