

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
Тернопільський національний економічний університет



ЗАТВЕРДЖУЮ  
Проректор з наукової роботи

*Задорожний З.-М. В.* Задорожний З.-М. В.

“ 24 ” 09 2019 р.

**РОБОЧА ПРОГРАМА**  
з дисципліни

«МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ»

рівень вищої освіти – третій (освітньо-науковий)  
галузь знань – 12 Інформаційні технології  
спеціальність – 121 Інженерія програмного забезпечення  
освітньо-наукова програма – «Інженерія програмного забезпечення»

Тернопіль – ТНЕУ  
2019

## СТРУКТУРА РОБОЧОЇ ПРОГРАМИ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

### «Математичне моделювання та обчислювальні методи»

Дисципліна «Математичне моделювання та обчислювальні методи»	Галузь знань, спеціальність, СВО	Характеристика навчальної дисципліни
Кількість кредитів – 6	галузь знань 12 Інформаційні технології	<b>Статус дисципліни</b> нормативна <b>Мова навчання</b> українська
Кількість залікових модулів – 1	спеціальність 121 Інженерія програм- ного забезпечення	Рік підготовки: <i>Денна – 1</i> <i>Заочна – 1</i> Семестр: <i>Денна – 1</i> <i>Заочна – 1</i>
Кількість змістових модулів – 2	рівень вищої освіти – третій (освітньо- науковий)	Аудиторні години - <i>Денна форма:</i> <i>Лекції – 45 год</i> <i>Практичні – 45 год</i> <i>Заочна форма:</i> <i>Лекції – 24 год</i> <i>Практичні – 10 год</i>
Загальна кількість годин – 180		Самостійна робота: <i>Денна – 90 год</i> <i>Заочна – 146 год</i>
Тижневих годин – 12, з них аудиторних – 6		Вид підсумкового контролю – залік

## **2. Мета і завдання вивчення дисципліни «Математичне моделювання та обчислювальні методи»**

### **2.1. Мета вивчення дисципліни**

Основною метою викладання дисципліни «Математичне моделювання та обчислювальні методи» є формування у майбутніх фахівців теоретичних основ побудови методів попередньої обробки інформації, яка фіксується у тематичних базах даних, а також набуття навичок щодо використання згаданих методів для видобування корисних знань.

### **2.2. Завдання вивчення дисципліни**

В результаті вивчення курсу «Математичне моделювання та обчислювальні методи» студенти повинні:

- знати основні поняття попереднього аналізу вибірок, метод ієрархічного кластерного аналізу, метод к-середніх, методи побудови дерев рішень, лінійних класифікаторів, випадкових лісів, відбору ознак, основи прогнозування часових рядів;

- вміти здійснювати попередній аналіз вибірок, реалізовувати кластеризацію вибірок, відбір ознак, побудову дерев класифікації, лінійних класифікаторів, випадкових лісів, будувати регресійні залежності прогнозувати часові ряди.

### **2.3. Найменування та опис компетентностей, формування котрих забезпечує вивчення дисципліни:**

Здатність застосовувати сучасні інформаційні технології, бази даних та інші електронні ресурси, спеціалізоване програмне забезпечення у науковій та навчальній діяльності

### **2.4. Передумови для вивчення дисципліни**

Математичний аналіз, теорія імовірностей та математична статистика.

### **2.5. Результати навчання**

Розробляти та досліджувати концептуальні, математичні і комп'ютерні моделі процесів і систем, ефективно використовувати їх для отримання нових знань та/або створення інноваційних продуктів у інженерії програмного забезпечення та дотичних міждисциплінарних напрямках.

### **3. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ»**

#### **Змістовий модуль 1. Аналіз часових рядів та інформаційне моделювання**

##### **Тема 1. Автопроекційні методи прогнозування часових рядів.**

Закони функціонування і властивості систем. Мета системи. Основи формалізації системних процесів. Збалансована система показників. Часові ряди. Параметризовані моделі. Ковзаюче середнє. Експоненціально зважене середнє. Метод Холта. Багатокроковий автопроекційний прогноз.

##### **Тема 2. Адаптивні короткотермінові прогнози**

Сезонність часових рядів. Сезонні автопроекційні методи. Мультиплікативна сезонність. Метод Холта-Вінтера. Адаптивне експоненціальне згладжування. Метод Тріга-Ліча. Виявлення тенденцій в часових рядах. Регресійні моделі. Модель білого шуму. Нормально розподілені випадкові величини. Означення кореляції, автокореляції та їх властивості. Автокореляційні характеристики часових рядів. Аналіз залишків. Авторегресійні моделі.

##### **Тема 3. Засоби організації обчислень модельного аналізу**

Пакет прикладних програм R. Основні компоненти пакету. Мова матричних обчислень. Команди роботи з матрицями та векторами. Основні варіанти організації розгалужень та циклів. Засоби побудови графічних ілюстрацій. Управління форматами виводу.

##### **Тема 4. Прогнозування часових рядів із використанням ARIMA-моделей**

Основні поняття про параметричні моделі часових рядів властивості загальної моделі. Процеси ковзної середньої (MA(q)-процеси). Авторегресійні процеси (AR(p)-процеси). ARMA- та ARIMA-процеси. Узагальнена авторегресійна модель ARIMAX. Сезонна модель Бокса-Дженкінса. Ідентифікація моделей Бокса-Дженкінса за допомогою пакету R.

##### **Тема 5. Дерева рішень**

Узагальнення експертної інформації. Метод побудови ієрархії перемикачів. Ентропія. Жадібний приріст інформації. Невизначеність Джіні. Побудова границі розмежування. Пороги розмежування. Критерій класифікації для кількісних показників. Класи дерев рішень в пакетах Python.

## **Тема 6. Лінійні класифікатори**

Метод найменших квадратів. Гомоскедастичність. Оцінка за мінімальною дисперсією. Метод максимальної правдоподібності. Квадрат зміщення. Дисперсія. Неусувна похибка. Мультиколінеарність. Некоректні задачі. Регуляризація Тихонова. Гребенева регресія. Логістична регресія. Варіанти використання логістичної регресії. Проблема виключаючого «або».

## **Змістовий модуль 2. Методи аналізу даних**

### **Тема 7. Композиції. Беггін, випадковий ліс**

Генерація підвибірок. Бутстреп. Усереднення випадкових класифікаторів. Беггін. Помилки лінійного класифікатора та випадкового лісу. Алгоритм побудови випадкового лісу. Випадкові дерева. Плюси та мінуси випадкового лісу.

### **Тема 8. Відбір ознак**

Тексти. Мішок слів. Векторизовані слова. Аналіз зображень. Геодані. Дата і час. Браузер користувача. Нормалізація та трансформація розподілу. Відбір ознак. Статистичний критерій.

### **Тема 9. Ієрархічний кластерний аналіз**

Віддаль між кластерами. Центроїдний метод та його недоліки. Методи далекого та найближчого сусіда. Віддаль Соренсона. Рекомендовані методи попереднього кластерного аналізу. Метод покрокового агрегування кластерів. Дендрограми та алгоритми їх побудови. Критерій зупинки. Насипи (лікті) для масштабних кластеризацій.

### **Тема 10. Кластеризація методом $k$ -середніх.**

Етапи побудови моделі. Експертне задання центрів кластерів. Ініціалізація датчиків випадкових чисел. Обмеження кількості змін центрів кластерів. Кількість проведення процедур кластеризації. Вибір найкращої кластеризації за критерієм якості. Допустимість мінімальної зміни центрів кластерів. Метод голосування кількості кластерів. Середнє значення показників по кластеру. Інтерпретація типових представників кластеру.

#### 4. Структура залікового кредиту дисципліни "МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ"

##### денна форма навчання

	Кількість годин	
	Аудиторні години	Самостійна робота
<b>Змістовий модуль 1. Аналіз часових рядів та інформаційне моделювання</b>		
Тема 1. Автопроекційні методи прогнозування часових рядів.	8	8
Тема 2. Адаптивні короткотермінові прогнози	8	8
Тема 3. Засоби організації обчислень модельного аналізу	8	8
Тема 4. Прогнозування часових рядів із використанням ARIMA-моделей	8	8
Тема 5. Дерева рішень	8	8
Тема 6. Лінійні класифікатори	10	10
<b>Змістовий модуль 2. Методи аналізу даних</b>		
Тема 7. Композиції. Беггін, випадковий ліс	10	10
Тема 8. Відбір ознак.	10	10
Тема 9. Ієрархічний кластерний аналіз	10	10
Тема 10. Кластеризація методом к-середніх.	10	10
Разом	90	90

##### заочна форма навчання

	Кількість годин	
	Аудиторні години	Самостійна робота
<b>Змістовий модуль 1. Аналіз часових рядів та інформаційне моделювання</b>		
Тема 1. Автопроекційні методи прогнозування часових рядів.	3	14
Тема 2. Адаптивні короткотермінові прогнози	3	14
Тема 3. Засоби організації обчислень модельного аналізу	3	14
Тема 4. Прогнозування часових рядів із використанням ARIMA-моделей	3	14
Тема 5. Дерева рішень	3	14
Тема 6. Лінійні класифікатори	3	14
<b>Змістовий модуль 2. Методи аналізу даних</b>		

Тема 7. Композиції. Беггін, випадковий ліс	4	14
Тема 8. Відбір ознак.	4	16
Тема 9. Ієрархічний кластерний аналіз	4	16
Тема 10. Кластеризація методом k-середніх.	4	16
Разом	34	146

### 5. Самостійна робота

№ п/п	Тематика	К-сть годин ДФН	К-сть годин ЗФН
1.	Закони функціонування і властивості систем. Мета системи. Основи формалізації системних процесів. Збалансована система показників.	3	7
2.	Часові ряди. Параметризовані моделі. Ковзаюче середнє. Експоненціально зважене середнє. Метод Холта. Багатокроковий автопроекційний прогноз.	3	7
3.	Сезонність часових рядів. Сезонні автопроекційні методи. Мультиплікативна сезонність. Метод Холта-Вінтера. Адаптивне експоненціальне згладжування. Метод Трига-Ліча.	3	7
4.	Виявлення тенденцій в часових рядах. Регресійні моделі. Модель білого шуму. Нормально розподілені випадкові величини.	3	7
5.	Означення кореляції, автокореляції та їх властивості. Автокореляційні характеристики часових рядів. Аналіз залишків. Авторегресійні моделі.	3	7
6.	Пакет прикладних програм R. Основні компоненти пакету. Мова матричних обчислень. Команди роботи з матрицями та векторами. Основні варіанти організації розгалужень та циклів. Засоби побудови графічних ілюстрацій. Управління форматами виводу.	3	7
7.	Основні поняття про параметричні моделі часових рядів властивості загальної моделі. Процеси ковзної середньої (MA(q)-процеси). Авторегресійні процеси (AR(p)-процеси).	3	7
8.	ARMA- та ARIMA-процеси. Узагальнена авторегресійна модель ARIMAX. Сезонна модель Бокса-Дженкінса. Ідентифікація моделей Бокса-Дженкінса за допомогою пакету R.	3	7
9.	Узагальнення експертної інформації. Метод побудови ієрархії перемикачів. Ентропія. Жадібний приріст інформації. Невизначеність Джіні.	3	7
10.	Побудова границі розмежування. Пороги розмежування. Критерій класифікації для кількісних показників. Класи дерев рішень в пакетах Python.7	4	7
11.	Метод найменших квадратів. Гомоскедастичність. Оцінка за	4	7

	мінімальною дисперсією. Метод максимальної правдоподібності. Квадрат зміщення. Дисперсія. Неусувна похибка. Мультиколінеарність.		
12.	Некоректні задачі. Регуляризація Тихонова. Гребенева регресія. Логістична регресія. Варіанти використання логістичної регресії. Проблема виключаючого «або».	4	7
13	Генерація підвбірок. Бутстреп. Усереднення випадкових класифікаторів. Беггін.	4	7
14	Помилки лінійного класифікатора та випадкового лісу. Алгоритм побудови випадкового лісу. Випадкові дерева. Плюси та мінуси випадкового лісу.	4	7
15	Тексти. Мішок слів. Векторизовані слова. Аналіз зображень. Геодані. Дата і час. Браузер користувача.	4	7
16	Нормалізація та трансформація розподілу. Відбір ознак. Статистичний критерій.	5	7
17.	Віддаль між кластерами. Центроїдний метод та його недоліки. Методи далекого та найближчого сусіда. Віддаль Соренсона. Рекомендовані методи попереднього кластерного аналізу.	5	7
18.	Метод покрокового агрегування кластерів. Дендрограми та алгоритми їх побудови. Критерій зупинки. Насипи (лікті) для масштабних кластеризацій.	5	7
19.	Відбір змінних кластеризації. Сурогатні змінні. Роль стандартизації змінних. Методи стандартизації. Стандартизація за розмахом та за допомогою z-міток. Інтерпретація результатів кластерного аналізу. Обмеженість на розмірності в методі ієрархічного кластерного аналізу. Приклади використання кластерного аналізу.	5	7
20.	Етапи побудови моделі. Експертне задання центрів кластерів. Ініціалізація датчиків випадкових чисел. Обмеження кількості змін центрів кластерів. Кількість проведення процедур кластеризації.	5	7
21.	Вибір найкращої кластеризації за критерієм якості. Допустимість мінімальної зміни центрів кластерів. Метод голосування кількості кластерів. Середнє значення показників по кластеру. Інтерпретація типових представників кластеру.	5	6
<b>Разом:</b>		90	146

## 6. Засоби оцінювання та методи демонстрування результатів навчання

У процесі вивчення дисципліни «Математичне моделювання та обчислювальні методи» використовуються такі засоби оцінювання та методи демонстрування результатів навчання:

- поточне опитування;
- підсумкове тестування по кожному змістовому модулю;



- розрахункові та розрахунково-графічні роботи;
- залік

### 7. Критерії, форми поточного та підсумкового контролю

Підсумковий бал (за 100-бальною шкалою) з дисципліни «Математичне моделювання та обчислювальні методи» визначається за шкалою оцінювання:

За шкалою ТНЕУ	За національною шкалою	За шкалою ECTS
90–100	відмінно	A (відмінно)
85–89	добре	B (дуже добре)
75-84		C (добре)
65-74	задовільно	D (задовільно)
60-64		E (достатньо)
35-59	незадовільно	FX (незадовільно з можливістю повторного складання)
1-34		F (незадовільно з обов'язковим повторним курсом)

### 8. Інструменти, обладнання та програмне забезпечення, використання яких передбачає навчальна дисципліна

№	Найменування	Номер теми
1.	Пакет R, програмна оболонка R Studio	1-10

### РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

1. Зарядов И. С. Введение в статистический пакет R: типы переменных, структуры данных, чтение и запись информации, графика. М, 2010.
2. Мاستицкий С.Э., Шитиков В.К. (2014) Статистический анализ и визуализация данных с помощью R. – Электронная книга, адрес доступа: <http://r-analytics.blogspot.com>
3. Тамбовцева А. Проверка гипотез в R. Режим доступа: <https://rpubs.com/AllaT/hypo-test-r>
4. Реусова А. Иерархическая кластеризация категориальных данных в R. <https://habr.com/ru/company/otus/blog/461741/>
5. Coghlan А. Использование R для анализа временных рядов/ [http://templet.ssau.ru/wiki/\\_export/pdf/translate/using\\_r\\_for\\_time\\_series\\_analysis?rev=0](http://templet.ssau.ru/wiki/_export/pdf/translate/using_r_for_time_series_analysis?rev=0).

6. А.В. Трегуб, И.В. Трегуб. Методика построения модели Arima для прогнозирования динамики временных рядов. Режим доступа: <https://cyberleninka.ru/article/n/metodika-postroeniya-modeli-arima-dlya-prognozirovaniya-dinamiki-vremennyh-ryadov>
7. Ю.А. Крюков, Д.В. Чернягин. ARIMA – модель прогнозирования значений трафика. Режим доступа: [http://www.isa.ru/jitcs/images/documents/2011-02/41\\_49.pdf](http://www.isa.ru/jitcs/images/documents/2011-02/41_49.pdf)
8. О. Дзедзелюк, Л. Костів, В. Рабик. Побудова Arima моделей часових рядів для прогнозування метеоданих на мові програмування R. Електроніка та інформаційні технології. 2013. Випуск 3. С. 211–219.
9. Грас Дж. Data Science. Наука о данных с нуля. БХВ-Петербург. 2016. 336с.
10. К. Дейт. Введение в системы баз данных М.; СПб.: Вильямс, 2005– 576с.
11. Д. Мейер. Теория реляционных баз данных. М., Мир, 1987. – 368с.
12. Фаулер М., Скотт К. UML в кратком изложении. Применение стандартного языка объектного моделирования. М., Мир, 2009. – 234с.
13. Кендэл М. Временные ряды. - М. Финансы и статистика, 1981, -199с.
14. Льюис К. Методы прогнозирования экономических показателей. - М. Финансы и статистика, 1986, -133с.
15. Айвазян С.А. Прикладная статистика. Анализ зависимостей. - М. Финансы и статистика, 1989, -606с.
16. Open Data Science. Первичный анализ данных с Pandas. Режим доступа: <https://habr.com/ru/company/ods/blog/322626/>
17. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. - М. Финансы и статистика. Книга\_1, 1986-366с. Книга 2, 1987- 351с.
18. Lampros М. Кластеризация с пакетом ClusterR. Режим доступа: <https://habr.com/ru/company/infopulse/blog/310288/>
19. Open Data Science. Классификация, деревья решений и метод ближайших соседей. Режим доступа: <https://habr.com/ru/company/ods/blog/322534/>
20. Open Data Science. Линейные модели классификации и регрессии Режим доступа: <https://https://habr.com/ru/company/ods/blog/323890/>
21. Open Data Science. Композиции: бэггинг, случайный лес. Режим доступа: <https://https://habr.com/ru/company/ods/blog/323890/>
22. John Verzani. Using R for Introductory Statistics. Access mode: <https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf>
23. Rita Giordano. Visualizing statistics in R. Access mode: <http://www.cryst.chem.uu.nl/lutz/ecacomsig/pdfs-warwick/RitaGiordano-VisualisingStatisticsInR.pdf>
24. Anne Segonds-Pichon. Introduction to Statistics with R. [https://www.bioinformatics.babraham.ac.uk/training/R\\_Statistics/Introduction%20to%20Statistics%20with%20R%20manual.pdf](https://www.bioinformatics.babraham.ac.uk/training/R_Statistics/Introduction%20to%20Statistics%20with%20R%20manual.pdf)
25. Joel Grus. Data Science from Scratch. Access mode:

[http://math.ecnu.edu.cn/~lfzhou/seminar/\[Joel\\_Grus\]\\_Data\\_Science\\_from\\_Scratch\\_First\\_Princ.pdf](http://math.ecnu.edu.cn/~lfzhou/seminar/[Joel_Grus]_Data_Science_from_Scratch_First_Princ.pdf)

26. C J Date. An Introduction to Database Systems. Access mode:<https://www.pdfdrive.com/an-introduction-to-database-systems-8e-by-c-j-datepdf-e31093920.html>
27. David Maier. The Theory of Relational Databases. Access mode: <http://web.cecs.pdx.edu/~maier/TheoryBook/TRD.html>
28. Sergey Subbotin. A Random Forest Model Building Using A priori Information for Diagnosis. Access mode:<http://ceur-ws.org/Vol-2353/paper76.pdf>
29. Cedric Beaulac Jeffrey S. Rosenthal. Predicting University Students' Academic Success and Major using Random Forests. . January 15, 2019. <https://arxiv.org/pdf/1802.03418.pdf>
30. Eric Ariel L. Salas ,Sakthi Kumaran Subburayalu. Modified shape index for object-based random forest image classification of agricultural systems using airborne hyperspectral datasets. Published: March 7, 2019. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0213356>