

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
імені В. Н. КАРАЗИНА

**К. Ю. Кононова**

# **МАШИННЕ НАВЧАННЯ**

## **МЕТОДИ ТА МОДЕЛІ**

Підручник

Харків – 2020

УДК 519.2(075.8)

К 64

**Рецензенти:**

**А. В. Матвійчук** – д. е. н., професор кафедри економіко-математичного моделювання Державного вищого навчального закладу «Київський національний економічний університет імені Вадима Гетьмана»;

**Л. С. Гур'янова** – д. е. н., професор, завідувач кафедри економічної кібернетики Харківського національного економічного університету імені Семена Кузнеця.

*Затверджено до друку рішенням Вченої ради  
Харківського національного університету імені В. Н. Каразіна  
(протокол № 13 від 23 грудня 2019 року)*

**Кононова К. Ю.**

К 64 Машинне навчання: методи та моделі : підручник для бакалаврів, магістрів та докторів філософії спеціальності 051 «Економіка» / К. Ю. Кононова. Харків: ХНУ імені В. Н. Каразіна, 2020. 280 с.

ISBN 978-966-285-654-5

У підручнику розглядаються основні принципи аналізу даних та машинного навчання, приділяється увага CRISP-DM методології та питанням підготовки даних. Розглянуто базові моделі лінійної та нелінійної регресії, класифікації (зокрема логістична регресія, методи опорних векторів та k-найближчих сусідів, Байєсова класифікація, дерева рішень і випадковий ліс), кластеризації (ієрархічна і k-середніх), а також методи побудови асоціативних правил. Серед більш складних методів машинного навчання розглянуто методи обробки природної мови (модель «Мішок слів», класифікація текстів та аналіз настроїв), використання штучних нейронних мереж в задачах прогнозування, класифікації та кластеризації, подано алгоритми глибокого навчання (згорткові та рекурентні мережі).

Кожен розділ підручника містить теоретичний матеріал, ілюстративні матеріали та лабораторні роботи, виконання яких дозволить студентам оволодіти методами та моделями машинного навчання, а також питання для самоперевірки та завдання для самостійного виконання.

Для бакалаврів, магістрів та докторів філософії, а також викладачів ВНЗ. Підручник буде корисним вченим і фахівцям, які використовують в роботі технології машинного навчання.

УДК 519.2(075.8)

ISBN 978-966-285-654-5

© Харківський національний університет імені В. Н. Каразіна, 2020

© Кононова К. Ю., 2020

© Чорна О. Д., макет обкладинки, 2020

# ЗМІСТ

<b>ВСТУП</b> .....	5
<b>РОЗДІЛ 1. Основні поняття інтелектуального аналізу даних</b> .....	8
Тема 1. Завдання аналізу даних .....	8
Лабораторна робота 1 .....	12
Тема 2. CRISP-DM методологія .....	13
Тема 3. Підготовка даних .....	19
Лабораторна робота 2 .....	23
Питання для самоперевірки .....	30
Самостійна робота 1 .....	30
<b>РОЗДІЛ 2. Регресія</b> .....	31
Тема 4. Лінійна регресія .....	31
Лабораторна робота 3 .....	32
Тема 5. Дерева рішень і випадковий ліс .....	42
Лабораторна робота 4 .....	44
Питання для самоперевірки .....	55
Самостійна робота 2 .....	55
<b>РОЗДІЛ 3. Класифікація</b> .....	56
Тема 6. Логістична регресія .....	56
Лабораторна робота 5 .....	60
Тема 7. Метод опорних векторів .....	71
Лабораторна робота 6 .....	75
Тема 8. Метод k-найближчих сусідів .....	83
Лабораторна робота 7 .....	84
Тема 9. Байєсова класифікація .....	88
Лабораторна робота 8 .....	91
Тема 10. Дерева рішень та ансамблеві методи .....	95
Лабораторна робота 9 .....	96
Питання для самоперевірки .....	108
Самостійна робота 3 .....	108
<b>РОЗДІЛ 4. Кластеризація</b> .....	109
Тема 11. Ієрархічна кластеризація .....	109
Лабораторна робота 10 .....	110
Тема 12. Кластеризація на основі k-means .....	116
Лабораторна робота 11 .....	119
Питання для самоперевірки .....	124
Самостійна робота 4 .....	124

<b>РОЗДІЛ 5. Асоціативні правила</b> .....	125
Тема 13. Побудова асоціативних правил .....	125
Тема 14. Алгоритми APRIORI та ECLAT .....	128
Лабораторна робота 12 .....	135
Питання для самоперевірки .....	142
Самостійна робота 5 .....	142
<b>РОЗДІЛ 6. Обробка природної мови</b> .....	143
Тема 15. Модель «мішок слів» .....	143
Лабораторна робота 13 .....	145
Тема 16. Моделі з урахуванням семантики .....	150
Питання для самоперевірки .....	151
Самостійна робота 6 .....	151
<b>РОЗДІЛ 7. Штучні нейронні мережі</b> .....	152
Тема 17. Теоретичні основи нейронних мереж .....	152
Тема 18. НМ в задачах апроксимації та прогнозування .....	158
А. НМ в задачах апроксимації .....	158
Лабораторна робота 14 .....	159
Б. НМ в задачах прогнозування .....	169
Лабораторна робота 15 .....	170
Тема 19. НМ в задачах класифікації .....	177
Лабораторна робота 16 .....	177
Тема 20. НМ в задачах кластеризації: карти Кохонена .....	184
Лабораторна робота 17 .....	185
Питання для самоперевірки .....	193
Самостійна робота 7 .....	194
<b>РОЗДІЛ 8. Глибоке навчання</b> .....	195
Тема 21. Рекурентні нейронні мережі .....	196
Лабораторна робота 18 .....	199
Тема 22. Згорткові нейронні мережі .....	205
Лабораторна робота 19 .....	208
Питання для самоперевірки .....	210
Самостійна робота 8 .....	211
<b>ДОДАТКИ</b> .....	212
<b>ЛІТЕРАТУРА</b> .....	275