Обчислювальні методи дають змогу покращити валідацію та тестування моделей аналізу даних. Широкий спектр обчислювальних методів допомагає проводити різні експерименти з даними, перевіряти різні гіпотези та визначати найбільш ефективний підхід до аналізу набору даних. Це підвищує достовірність результатів аналізу та робить процес прийняття рішень більш обґрунтованим [3].

Можна зазначити, що роль методів обчислень у вирішенні задач аналізу даних є критичною та незамінною. Сучасні підходи до обробки даних ґрунтуються на їх використанні, і вони відіграють ключову роль у всіх аспектах аналізу даних. Методи обчислень допомагають виявляти закономірності, які важливі для прийняття обґрунтованих рішень у різних сферах життя, від медицини до економіки. До того ж вони стимулюють розвиток нових напрямів досліджень у сфері аналізу даних, що відкриває нові можливості для вирішення складних проблем та досягнення нових знань. Тому важливо продовжувати розвивати та вдосконалювати методи обчислень, щоб забезпечити подальший прогрес в аналізі даних та в їх практичному застосуванні.

Список використаних джерел

1. Опрацювання даних як інформаційний процес. *Основні поняття з інформатики. UA5.org.* URL: https://ua5.org/osnovi/2167-opraczyuvannya-danyh-yak-informaczijnyj-proczes.html

2. Сем Фокс. OpenAI – нові можливості для бізнесу – розумовий аналіз даних та виявлення патернів. *Mediacom: світ новин.* 30.03.2024. URL: https://mediacom.com.ua/openai-rozumovij-analiz-danix-ta-viyavlennya-paterniv-novi-mozhlivosti-dlya-biznesu/

3. Верес О. М. Класифікація методів аналізу великих даних. *Вісник Національного університету «Львівська політехніка». Серія: Інформаційні системи та мережі.* Львів: Видавництво Львівської політехніки, 2017. № 872. С. 84–92.

УДК 519.2

Сапожнікова В. Є., *здобувачка 2 курсу спеціальності 122 Комп'ютерні науки,*
Сеник І. О., *асистент кафедри інформаційних технологій*

МЕТОД НАЙМЕНШИХ КВАДРАТІВ

Донецький національний університет імені Василя Стуса, м. Вінниця

У галузі статистики часто виникає потреба знайти зв'язок між змінними, і метод найменших квадратів є широко використовуваною технікою для підгонки математичних моделей до даних. Завдяки своїй універсальності та широкому спектру застосувань він відіграє важливу роль у розв'язанні різноманітних задач.

Цей метод використовується для знаходження найкращої прямої (або деколи кривої), що представляє взаємозв'язок між цими змінними. Англійською мовою лінія має назву *least squares line,* що означає лінія найменших квадратів, або *regression line,* тобто лінія регресії, чи також *best fit line*, що перекладається як найбільш підходяща лінія; до прикладу, це може бути $y = mx + b$ (або $y = b_0 + b_1 x$).

Головна мета цього методу – зменшити суму квадратів помилок настільки, наскільки це можливо, що і є причиною назви [1].

Треба розглянути наочний приклад задля кращого розуміння версатильності методу. Уявімо, що нам потрібно знайти параметри лінійного рівняння, що найкраще задовольняють такі дані:

Таблиця 1 – Дані

x	1	2	3	4	5	6	7
y	1,5	3,8	6,7	9,0	11,2	13,6	16

Перевіривши дані, можна побачити, що точки не лежать на одній лінії, тобто необхідно знайти таку пряму, щоб кожне задане умовою значення знаходилось якнайближче до неї.

Перший пункт виконання завдання – визначення шуканих параметрів. Для знаходження задовільного рівняння прямої $y = mx + b$ потрібно скористатися такими формулами:

$$m = \frac{n\sum xy - \sum x \sum y}{[n\sum x^2 - (\sum x)^2]}; \qquad b = \frac{\sum y - m\sum x}{n},$$

де n – кількість заданих точок.

З формул робимо висновок, що спочатку треба знайти показник m, який у статистиці буде також мати позначення b_1. Для цього заповнюємо таблицю:

Таблиця 2 – Дані, необхідні для підрахунку m

x	y	xy	x^2
1	1.5	1.5	1
2	3.8	7.6	4
3	6.7	20.1	9
4	9	36	16
5	11.2	56	25
6	13.6	81.6	36
7	16	112	49

Також необхідно знайти суму кожного стовпчика таблиці для подальшої калькуляції тому, порахувавши, ми визначили, що $\sum x = 28, \sum y = 61.8, \sum xy = 61.8, \sum x^2 = 140$. Також занотовуємо, що $n = 7$.

Тепер підставляємо значення у попередньо вказані формули.

$$m = \frac{n\sum xy - \sum x \sum y}{[n\sum x^2 - (\sum x)^2]} = \frac{7(314.8) - (28)(61.8)}{7(140) - (28)^2} = 2.4142857;$$

$$b = \frac{\sum y - m\sum x}{n} = \frac{(61.8) - 2.4142857(28)}{7} = -0.828571.$$

Округлюючи значення і підставляючи їх у рівняння лінійної функції, отримуємо $y = 2.41x - 0.83$, що і є шуканою відповіддю.

Ми можемо переглянути точність виконання, підставивши початкові дані в отриману пряму. Візьмемо декілька випадкових аргументів для перевірки:

$$y = 2.41(2) - 0{,}83 = 3.99;$$

$$y = 2.41(5) - 0{,}83 = 11.22;$$

$$y = 2.41(2) - 0,83 = 16.04.$$

Отримані числа є дуже хорошими апроксимаціями, адже вони неймовірно близькі до вхідних.

Для вирішального переконання варто проілюструвати усі дані.

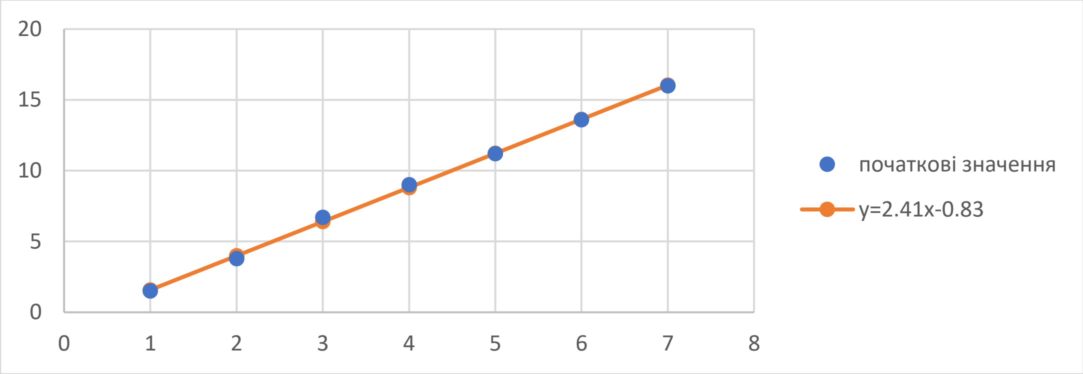

Рис. 1. Зображення початкових і отриманих даних

Тепер остаточного можна впевнитись, що виконані обчислення зроблені правильно.

Одне з головних використань методу найменших квадратів – регресійний аналіз. Він передбачає вивчення взаємозв'язку між залежною та однією або кількома незалежними змінними. Хоча значення методу найменших квадратів може виходити далі його застосування в регресійному аналізі.

Концепція мінімізації суми квадратів різниць між спостережуваними та прогнозованими значеннями виходить далеко за межі статистики. Вона знаходить застосування у різних галузях, наприклад, обробка сигналів, машинне навчання та навіть квантова механіка [2].

Ще одним важливим застосуванням методу найменших квадратів є апроксимація кривих, що дає змогу знаходити найкращу криву, яка проходить через набір точок. Цей підхід широко використовується, коли дані мають нелінійну природу і потребують більш складної моделі для точного відображення залежності між змінними. Наприклад, поліноміальна апроксимація кривих дає змогу створити математичну модель, яка мінімізує суму квадратів відхилень між фактичними та розрахованими значеннями [3]. Такий підхід є надзвичайно корисним у фізиці, біології, економіці та інших галузях, де складні залежності можна описати за допомогою поліномів, експонент або логарифмів.

Цей метод також був використаний у галузі геодезії для визначення найкращої сфери, що представляє форму Землі [4]. Це застосування передбачає складний математичний процес, який враховує безліч геодезичних вимірювань для створення моделі, що мінімізує суму квадратів різниць між спостережуваними та розрахованими значеннями. Також він використовувався для оцінки параметрів у цій же сфері, як-от гравітаційне поле та форма планети [4].

До того ж метод найменших квадратів знайшов своє місце у сфері обробки сигналів, де його також використовують для оцінки параметрів моделі сигналу, щоб найкраще відповідати спостережуваним даним. Це застосування є особливо важливим у таких галузях, як-от телекомунікації, радіолокація та сонар, де точна оцінка параметрів сигналу є необхідною для ефективного аналізу даних [4].

Ці різноманітні застосування та теоретичні основи підкреслюють метод найменших квадратів як потужний та універсальний інструмент, що проникає в різні дисципліни, стаючи незамінною технікою для аналізу даних та підгонки моделей.

Висновки. Отже, метод найменших квадратів пропонує надійний підхід до підгонки математичних моделей до даних. Зводячи до мінімуму суму квадратів різниць між спостережуваними та прогнозованими значеннями, він не лише забезпечує спосіб кількісної оцінки відповідності, але й дає змогу ідентифікувати закономірності та тенденції в даних. Загалом метод найменших квадратів є фундаментальною технікою в області аналізу та моделювання даних.

Список використаних джерел

1. Least Square Method: вебсайт. URL: https://www.cuemath.com/data/least-squares/
2. Chakraborty S., Morolia A., Peduri A. Quantum Regularized Least Squares. 2023. Vol. 7. 988 p. URL: https://quantum-journal.org/papers/q-2023-04-27-988/
3. Seber G. A. F., Wild C. J. Nonlinear Regression. John Wiley & Sons. 2003. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/0471725315
4. Eshagh M. The Earth's Gravity Field Role in Geodesy and Large-Scale Geophysics. 23 October 2020. URL: https://www.intechopen.com/chapters/76431

УДК: 004.58:004.8:007.5

Сивак Д. І., *здобувач 2 курсу*
спеціальності 122 Комп'ютерні науки,
Комаров В. Ф., *канд. техн. наук,*
старший викладач кафедри
інформаційних технологій

РОЗВИТОК ТЕХНОЛОГІЙ МАШИННОГО НАВЧАННЯ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ТА БЕЗПЕКИ АВТОНОМНИХ ТРАНСПОРТНИХ ЗАСОБІВ

Донецький національний університет імені Василя Стуса, м. Вінниця

У сучасному світі автомобільна промисловість швидко рухається в напрямі автономності, переходячи від традиційних автомобілів до безпілотних систем. Зростання інтересу до автономних автомобілів спонукає нас до розуміння та розробки комплексу технологій, спрямованих на підвищення ефективності та безпеки транспортних систем. Одним із ключових компонентів цього переходу є використання технік машинного навчання для розвитку інтелектуальних систем керування безпілотними автомобілями.

В сучасних безпілотних автомобілях широко використовуються різноманітні методи машинного навчання для забезпечення автономної роботи. Декілька основних методів включають:

1. Нейронні мережі: зокрема, згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN) широко використовуються з технологією глибокого навчання для розпізнавання об'єктів на дорозі, виявлення дорожніх знаків та інших дорожніх об'єктів, а також для прогнозування руху інших учасників дорожнього руху [1].