Міністерство освіти і науки України
Донбаська державна машинобудівна академія (ДДМА)

ОСНОВИ СУЧАСНИХ ТЕОРІЙ МОДЕЛЮВАННЯ ПРОЦЕСІВ

Методичні вказівки до лабораторно – практичних занять

для студентів спеціальності 131 «Прикладна механіка» всіх форм навчання

Краматорськ
ДДМА
2018

Міністерство освіти і науки України
Донбаська державна машинобудівна академія (ДДМА)

ОСНОВИ СУЧАСНИХ ТЕОРІЙ МОДЕЛЮВАННЯ ПРОЦЕСІВ

Методичні вказівки до лабораторно – практичних занять

для студентів спеціальності 131 «Прикладна механіка» всіх форм навчання

Затверджено
на засіданні
методичної ради
Протокол №4 від 2.10.2018

Краматорськ
ДДМА
2018

УДК

Методичні вказівки до лабораторно - практичних робіт з дисципліни «Основи сучасних теорій моделювання процесів» / Укладачі: С. В. Ковалевський, В. В. Ємець. – Краматорськ: ДДМА, 2018. – 75с.

Складається з відомостей щодо підготовки до занять, форми ведення занять, загальних теоретичних положень та варіантів завдань з основних тем дисципліни «Основи сучасних теорій моделювання».

Укладачі: С.В. Ковалевський, проф. д.т.н.
В.В. Ємець, аспірант (PhD)

Відп. за випуск С.В. Ковалевський, проф.

ЗМІСТ

Лабораторно-практична робота №1
Регресійний аналіз. Метод найменших квадратів

Мета роботи: оволодіння навичками моделювання з використанням регресійного аналізу за допомогою методу найменших квадратів

Теоретичні відомості

Рівняння лінійної парної регресії виглядає наступним чином:$Y=a_0+a_1X$.

За допомогою цього рівняння змінна Y виражається через константу a_0 і кут нахилу прямої (або кутовий коефіцієнт) a_1, помножене на значення змінної X. Константу a0 також називають вільним членом, а кутовий коефіцієнт - коефіцієнт регресії. Параметри рівняння можуть бути визначені за допомогою методу найменших квадратів (МНК).

Метод найменших квадратів (в довідкових системах англомовних програм - Least Squares Method, LS) є одним з основних методів визначення параметрів регресійних рівнянь, що дає найкращі лінійні незміщені оцінки. Саме він використовується в MS Excel. Лінійні – відноситься до характеру взаємозв'язку змінних. Незміщені означає, що очікувані значення коефіцієнтів регресії повинні бути дійсними коефіцієнтами. Тобто точки, побудовані за вихідними даними (x_i, y_i), повинні лежати якомога ближче до точок лінії регресії. Сутність даного методу полягає в знаходженні параметрів моделі, при яких сума квадратів відхилень емпіричних (фактичних) значень результуючого ознаки від теоретичних, отриманих за вибраним рівнянням регресії, тобто:

$$S=\sum_{i=1}^{n}\left(y_i^{\eth}-y_i\right)^2=\sum_{i}^{n}(y_i^p-a_0-a_1x)^2\rightarrow\min$$

де y_i^p – значення, обчислене за рівнянням регресії; $\left(y_i^p-y_i'\right)$ – **відхилення** (помилка, залишок) (рис. 1); n – кількість пар вихідних даних.

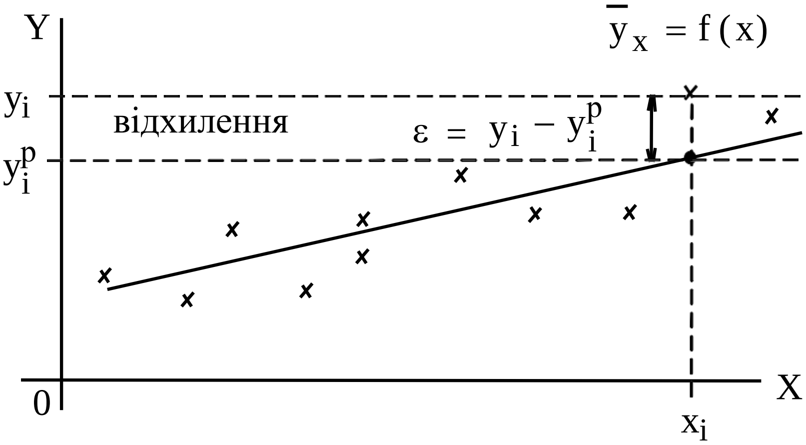


Рисунок. 1.1 – Поняття відхилення для випадку лінійної регресії

У регресійному аналізі передбачається, що математичне сподівання випадкової величини ε дорівнює нулю і її дисперсія однакова для всіх спостережуваних значень Y. Звідси випливає, що розсіювання даних біля лінії регресії має бути однаково при всіх значеннях параметра X. У випадку, показаному на рис. 1.2 дані розподіляються уздовж лінії регресії нерівномірно, тому метод найменших квадратів в цьому випадку непридатний.

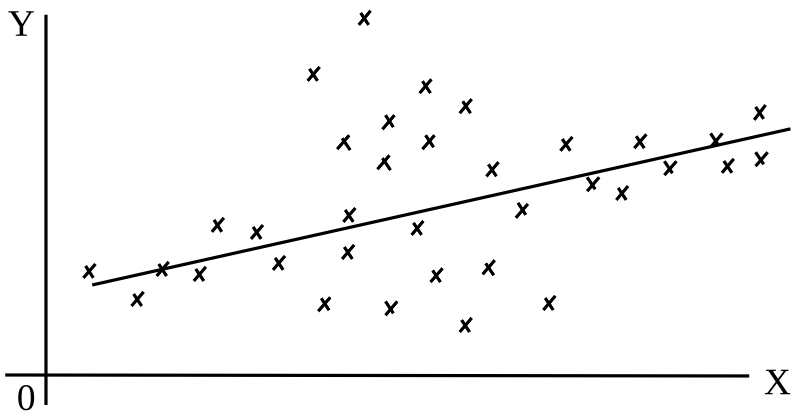


Рисунок 1.2 – Нерівномірний розподіл вихідних точок уздовж лінії регресії

Провівши необхідні перетворення, отримаємо систему двох рівнянь з двома невідомими a_0 і a_1, які знайдемо вирішивши систему.

$$a_1 = \frac{n\left(\sum y_i x_i\right) - \sum y_i \sum x_i}{n\left(\sum x_i^2\right) - \left(\sum x_i\right)^2};$$

$$a_0 = \frac{1}{n}\left(\sum y_i - a_1 \sum x_i\right)$$

Напрямок зв'язку між змінними визначається на підставі знаків (негативний або позитивний) коефіцієнта регресії (коефіцієнта a_1).

Якщо знак при коефіцієнті регресії-позитивний, зв'язок залежної змінної з незалежною буде позитивною. У нашому випадку знак коефіцієнта регресії позитивний, отже, зв'язок також є позитивною.

Якщо знак при коефіцієнті регресії-негативний, зв'язок залежної змінної з незалежною є негативною (зворотною).

Для аналізу загальної якості рівняння рівняння регресії використовують зазвичай множинний коефіцієнт детермінації R^2, званий також квадратом коефіцієнта множинної кореляції R. R^2 (міра визначеності) завжди знаходиться в межах інтервалу [0; 1].

Якщо значення R^2 близько до одиниці, це означає, що побудована модель пояснює майже всю мінливість відповідних змінних. І навпаки, значення R-квадрата, близьке до нуля, означає погану якість побудованої моделі.

Коефіцієнт детермінації R^2 показує, на скільки відсотків ($R^2 \cdot 100\%$) знайдена функція регресії описує зв'язок між вихідними значеннями факторів X і Y.

$$R^2 = \frac{\sum_{i=1}^{n}\left(y_i^{\delta} - \bar{y}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}$$

де $\left(y_i^p - \bar{y}\right)^2$ – пояснена варіація; $\left(y_i - \bar{y}\right)^2$ – загальна варіація (рис.1.3).

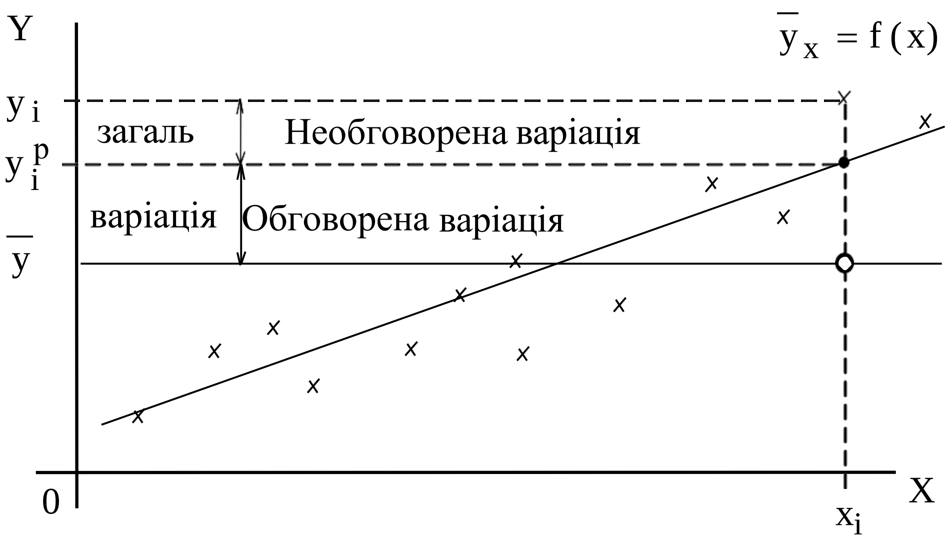


Рисунок 1.3 – Графічна інтерпретація коефіцієнта детермінації для випадку лінійної регресії

Відповідно, величина $(1-R^2)\cdot 100\%$ показує, скільки відсотків варіації параметра Y обумовлені факторами, не включеними в регресійну модель. При високому ($R^2 \geq 75\%$)значення коефіцієнта детермінації можна робити прогноз $y^* = f(x^*)$ для конкретного значення x^*.

Нелінійна регресія. Розглянемо найбільш прості випадки нелінійної регресії: гіперболу, експоненту і параболу. При знаходженні коефіцієнтів гіперболи і експоненти використовують прийом приведення нелінійної регресійної залежності до лінійного виду. Це дозволяє використовувати для обчислення коефіцієнтів функції регресії вищенаведені формули.

Гіпербола. Для приведення рівняння виду $y = a_0 + \frac{a_1}{x}$ до лінійного виду вводять нову змінну $z = \frac{1}{x}$, тоді рівняння гіперболи приймає лінійний вигляд $y = a_0 + a_1 z$. Після цього використовують формули для знаходження лінійної функції, але замість значень x_i використовуються значення

$$z_i = \frac{1}{x_i}.$$

$$a_1 = \frac{n\left(\sum y_i z_i\right) - \sum y_i \sum z_i}{n\left(\sum z_i^2\right) - \left(\sum z_i\right)^2}; \qquad a_0 = \frac{1}{n}\left(\sum y_i - a_1 \sum z_i\right).$$

Експонента. Для приведення до лінійного виду рівняння **експоненти** $y = a_0 e^{a_1 x}$ проведемо логарифмування:

$$\ln y = \ln\left(a_0 e^{a_1 x}\right);$$

$$\ln y = \ln a_0 + \ln\left(e^{a_1 x}\right);$$

$$\ln y = \ln a_0 + a_1 x.$$

Введемо змінні $b_0 = \ln a_0$ и $b_1 = a_1$, тоді $\ln y = b_0 + b_1 x$, звідки випливає, що можна застосовувати формули наведені вище, в яких замість значень y_i треба використовувати $\ln y_i$:

$$b_1 = \frac{n\left(\sum [\ln y_i]\, x_i\right) - \sum \ln y_i \sum x_i}{n\left(\sum x_i^2\right) - \left(\sum x_i\right)^2}; \quad b_0 = \frac{1}{n}\left(\sum \ln y_i - b_1 \sum x_i\right)$$

При цьому ми отримаємо чисельні значення коефіцієнтів b_0 і b_1, від яких треба перейти до a_0 і a_1, використовуваних в моделі експоненти. Виходячи з введених позначень і визначення логарифма, отримуємо:

$$a_0 = e^{b_0}, \ a_1 = b_1.$$

Парабола. Для знаходження коефіцієнтів рівняння параболи $y = a_0 + a_1 x + a_2 x^2$ необхідно вирішити лінійну систему з трьох рівнянь:

$$\begin{cases} n \cdot a_0 + \left(\sum x_i\right) a_1 + \left(\sum x_i^2\right) a_2 = \sum y_i, \\ \left(\sum x_i\right) a_0 + \left(\sum x_i^2\right) a_1 + \left(\sum x_i^3\right) a_2 = \sum\left(y_i x_i\right), \\ \left(\sum x_i^2\right) a_0 + \left(\sum x_i^3\right) a_1 + \left(\sum x_i^4\right) a_2 = \sum\left(y_i x_i^2\right). \end{cases}$$

Сила регресійної зв'язку для гіперболи і параболи визначається безпосередньо за тією ж формулою, що і для лінійної моделі. При обчисленні коефіцієнта детермінації для експоненти всі значення параметра Y (вихідні, регресійні, середня) необхідно замінити на їх логарифми, наприклад, y_i^p – на $\ln\left(y_i^p\right)$ и т.д.

Якщо функція регресії визначена, інтерпретована і обґрунтована, і оцінка точності регресійного аналізу відповідає вимогам, можна вважати, що побудована модель і прогнозні значення мають достатню надійність.

Прогнозні значення, отримані таким способом, є середніми значеннями, які можна очікувати.

Методичні рекомендації

Для проведення регресійного аналізу та прогнозування необхідно:

1) **побудувати графік** вихідних даних і спробувати візуально, наближено визначити характер залежності;

2) **вибрати вид функції регресії**, яка може описувати зв'язок вихідних даних;

3) **визначити чисельні коефіцієнти** функції регресії методом найменших квадратів;

4) **оцінити силу** знайденої регресійної залежності на основі коефіцієнта детермінації R^2;

5) **зробити прогноз** (при $R^2 \geq 75\%$) або зробити висновок про неможливість прогнозування за допомогою знайденої регресійної залежності. При цьому не рекомендується використовувати модель регресії для тих значень незалежного параметра X, які не належать інтервалу, визначеному у вихідних даних.

Довідкова інформація по технології роботи з режимом "регресія" надбудови Пакет аналізу MS Excel

Режим роботи "регресія" служить для розрахунку параметрів рівняння лінійної регресії і перевірки його адекватності досліджуваному процесу.

Для вирішення завдання регресійного аналізу в MS Excel вибираємо в меню **Сервіс** команду **аналіз даних** та інструмент аналізу **"регресія".**

У діалоговому вікні задаємо наступні параметри:

1. Вхідний інтервал Y - це діапазон даних по результативному ознакою. Він повинен складатися з одного стовпця.

2. Вхідний інтервал X-це діапазон комірок, що містять значення факторів (незалежних змінних). Число вхідних діапазонів (стовпців) повинно бути не більше 16.

3. Прапорець мітки, встановлюється втому випадку, якщо в першому рядку діапазону стоїть заголовок.

4. Прапорець Рівень надійності активізується, якщо в поле, що знаходиться поруч з ним необхідно ввести рівень надійності, відмінний від встановленого за замовчуванням. Використовується для перевірки значущості коефіцієнта детермінації R^2 і коефіцієнтів регресії.

5. Константа нуль. Даний прапорець необхідно встановити, якщо лінія регресії має пройти через початок координат (a_0=0).

6. Вихідний інтервал / новий робочий лист / нова робоча книга-вказати адресу верхньої лівої комірки вихідного діапазону.

7. Прапорці в групі залишки встановлюються, якщо необхідно включити у вихідний діапазон відповідні стовпці або графіки.

8. Прапорець Графік нормальної ймовірності необхідно зробити активним, якщо потрібно вивести на лист точковий графік залежності спостережуваних значень Y від автоматичних інтервалів персентилей.

Після натискання кнопки ОК у вихідному діапазоні отримуємо звіт.

Приклад виконання лабораторної роботи

Задача: Деяка фірма займається поставками різних вантажів на короткі відстані всередині міста. Оцінити вартість таких послуг, що залежить від витрачається на поставку часу. В якості найбільш важливого фактора, що впливає на час поставки, вибрано пройдену відстань. Були зібрані вихідні дані про десять постачання (таблиця 1).

Таблиця 1.1 – Вихідні дані

Відстань, миль	3,5	2,4	4,9	4,2	3,0	1,3	1,0	3,0	1,5	4,1
Час, хв	16	13	19	18	12	11	8	14	9	16

Визначте характер залежності між відстанню і витраченим часом, використовуючи майстер діаграм MS Excel, проаналізуйте застосовність методу найменших квадратів, побудуйте рівняння регресії, використовуючи МНК, проаналізуйте силу регресійної зв'язку. Проведіть регресійний аналіз, використовуючи режим роботи "Регресія" в MS Excel і порівняйте з результатами, отриманими раніше. Зробіть прогноз часу поїздки на 2 милі. Порахувати і побудувати графічно міру помилки регресійної моделі використовуючи табличний процесор Excel.

Рішення

На графіку будуємо вихідні дані по десяти поїздках.

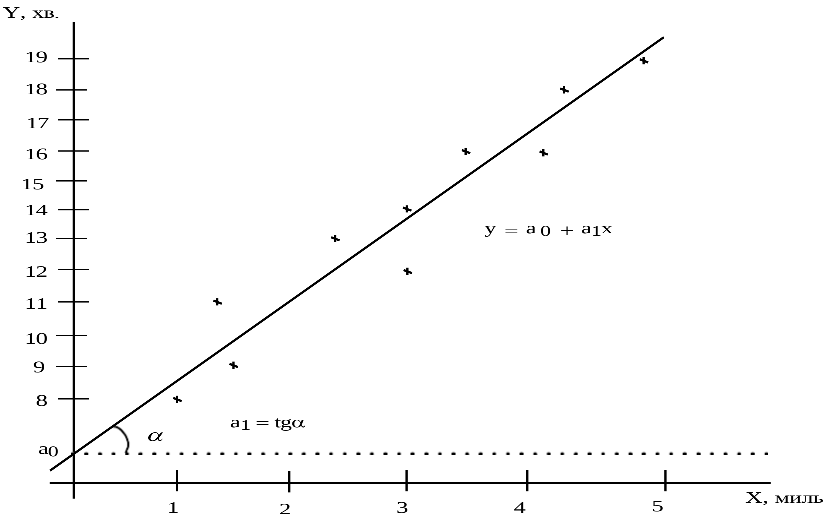


Рисунок 1.4 – Графік вихідних даних і передбачувана лінія регресії

Крім відстані на час поставки впливають пробки на дорогах, час доби, дорожні роботи, погода, кваліфікація водія, вид транспорту. Побудовані точки не знаходяться точно на лінії, що обумовлено описаними вище факторами. Але ці точки зібрані навколо прямої лінії, тому можна припустити лінійний зв'язок між параметрами. Всі вихідні точки рівномірно розподілені уздовж передбачуваної прямої лінії, що дозволяє застосувати метод найменших квадратів.

Обчислимо суми, необхідні для розрахунку коефіцієнтів рівняння лінійної регресії та коефіцієнта детермінації R_2 за допомогою допоміжної таблиці (таблиця 1.2).

Таблиця 1.2 –

x_i	y_i	x_i^2	$x_i y_i$	y_i^p	$(y_i^p - \bar{y})^2$	$(y_i - \bar{y})^2$
3,5	16	12,25	56,00	15,223	2,634129	5,76
2,4	13	5,76	31,2	12,297	1,697809	0,36
4,9	19	24,01	93,1	18,947	28,59041	29,16
4,2	18	17,64	75,60	17,085	12,14523	19,36
3,0	12	9,00	36,00	13,893	0,085849	2,56
1,3	11	1,69	14,30	9,371	17,88444	6,76
1,0	8	1,00	8,00	8,573	25,27073	31,36
3,0	14	9,00	42,00	13,893	0,085849	0,16
1,5	9	2,25	13,50	9,903	13,66781	21,16
4,1	16	16,81	65,60	16,819	10,36196	5,76
$\sum =$ 28,9	$\sum =$ 136	$\sum =$ 99,41	$\sum =$ 435,30	–	112,4242	122,4

$$\bar{y} = \frac{\sum y_i}{n} = \frac{16+13+19+18+12+11+8+14+9+16}{10} = 13,6.$$

Обчислимо коефіцієнти лінійної регресії за формулами:

$$a_1 = \frac{10 \cdot 435,30 - 136 \cdot 28,9}{10 \cdot 99,41 - 835,21} = 2,660;$$

$$a_0 = 0,1 \cdot (136 - 2,660 \cdot 28,9) = 5,913.$$

Таким чином, шукана регресійна залежність має вигляд:

$$y^p = 5,913 + 2,660\,x.$$

Нахил лінії регресії $a_1 =$ 2,66 хвилин на милю – це кількість хвилин, що припадає на одну милю відстані. Координата точки перетину прямої з віссю Y $a_0 =$ 5,913 хвилин – це час, який не залежить від пройденої відстані, а обумовлюється усіма іншими можливими факторами, явно не врахованими при аналізі.

Обчислимо коефіцієнт детермінації:

$$R^2 = \frac{112,424}{122,400} = 0,918 \qquad 91,8\%.$$

Проведемо регресійний аналіз з використанням режиму Регресія MS Excel. Значення параметрів, встановлених в однойменному діалоговому вікні, представлені на рис.6.

Рисунок 1.6 – Згенеруються результати з регресійної статистики, представлені в таблиці 3.

Таблиця 1.3 –

ВИСНОВОК ПІДСУМКІВ	
Регресійна статистика	
Множинний R	0,958275757
R-квадрат	0,918292427
Нормований R-квадрат	0,90807898
Стандартна помилка	1,11809028
Спостереження	10

Розглянемо представлену в таблиці 3 регресійну статистику.

Величина R-квадрат, звана також мірою визначеності, характеризує якість отриманої регресійної прямої. Ця якість виражається ступенем відповідності між вихідними даними і регресійною моделлю (розрахунковими даними). Міра визначеності завжди знаходиться в межах інтервалу [0;1]. У нашому прикладі міра визначеності дорівнює 0,91829, що говорить про дуже гарну підгонці регресійної прямої до вихідним даним і збігається з коефіцієнтом детермінації R^2, обчисленим за формулою.

Таким чином, лінійна модель пояснює 91,8% варіації часу доставки, що означає правильність вибору фактора (відстані). Не пояснюється $100\% - 91{,}8\% = 8{,}2\%$ варіації часу поїздки, які обумовлені іншими факторами, що впливають на час поставки, але не включеними в лінійну модель регресії.

Розрахований рівень значущості α_p=1,26E-05<0,05(показник значимість F в таблиці Дисперсійний аналіз) підтверджує значимість R^2.

Множинний R - коефіцієнт множинної кореляції R - виражає ступінь залежності незалежних змінних (X) і залежною змінною (Y) і дорівнює квадратному кореню з коефіцієнта детермінації, ця величина приймає значення в інтервалі від нуля до одиниці. У простому лінійному регресійному аналізі множинний R дорівнює коефіцієнту кореляції Пірсона. Дійсно, множинний R в нашому випадку дорівнює коефіцієнту кореляції Пірсона (0,95827), який обчислюється за формулою:

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2] \cdot [n\sum y^2 - (\sum y)^2]}}$$

Тепер розглянемо середню частину розрахунків, представлену в таблиці 1.4 (наведена в скороченому варіанті). Тут дано коефіцієнт регресії a_1 (2,65970168) і зсув по осі ординат, тобто константа a_0 (5,913462144).

Таблиця 1.4 –

	Коефіцієнт	*Стандартна хибка*	*по-t-статистика*	*P- Значення*
Y-перетин	5,913462144	0,884389599	6,686489927	0,00015485
Змінна X 1	2,65970168	0,280497238	9,482095791	1,26072E-05

Виходячи з розрахунків, можемо записати рівняння регресії таким чином:

$$y^{\partial} = 5{,}91346 + 2{,}65970\,x. \quad (*)$$

Бачимо, що це рівняння, збігається з рівнянням, отриманим нами при розрахунку по МНК вручну з точністю до помилки округлення.

Напрямок зв'язку між змінними визначається на підставі знаків (негативний або позитивний) коефіцієнта регресії (коефіцієнта a1). У нашому випадку знак коефіцієнта регресії позитивний, отже, зв'язок також є позитивною.

Далі перевіримо значимість коефіцієнтів регресії: a_0 і a_1.Порівнюючи попарно значення стовпців Коефіцієнти і Стандартна помилка в таблиці 1.4, бачимо, що абсолютні значення коефіцієнтів більше, ніж їх стандартні помилки. До того ж ці коефіцієнти є значущими, про що можна судити за значеннями показника Р-значення в таблиці 4, які менше заданого рівня значущості α=0,05.

Таблица 1.5 –

ВИВОД ЗАЛИШКУ			
Спостереження	*Передбачене Y*	*залишки*	*Стандартні залишки*
1	15,22241803	0,777581975	0,737641894
2	12,29674618	0,703253823	0,667131568
3	18,94600038	0,053999622	0,051225961

4	17,0842092	0,915790799	0,868751695
5	13,89256718	-1,892567185	-1,795356486
6	9,371074328	1,628925672	1,545256778
7	8,573163824	-0,573163824	-0,543723571
8	13,89256718	0,107432815	0,101914586
9	9,903014664	-0,903014664	-0,8566318
10	16,81823903	-0,818239033	-0,776210624

В таблиці 1.5. представлені результати виведення залишків. За допомогою цієї частини звіту ми можемо бачити відхилення кожної точки від побудованої лінії регресії. Найбільше абсолютне значення залишку в нашому випадку - 1,89256, найменше - 0,05399. Для кращої інтерпретації цих даних скористаємося графіком вихідних даних і побудованої лінією регресії, представленими на рис. 7. Як бачимо, лінія регресії добре "підігнана" під значення вихідних даних.

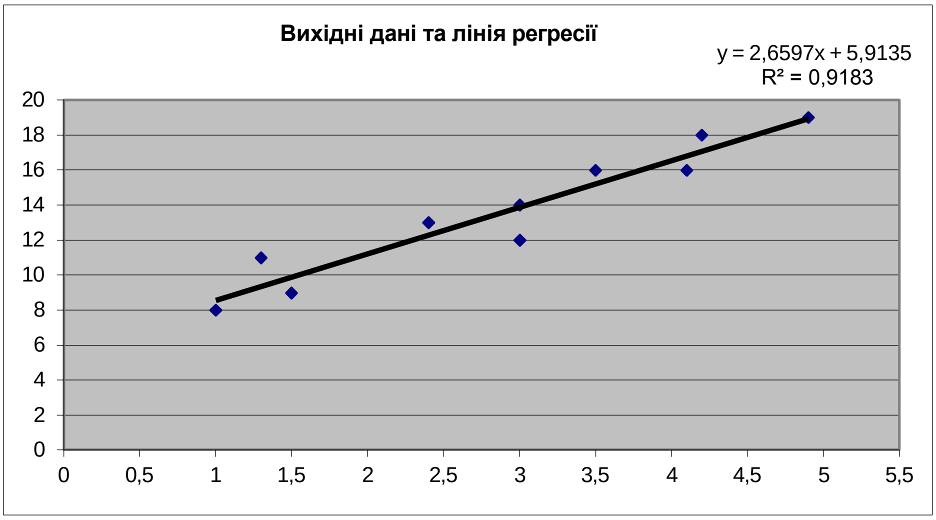


Рисунок 1.7 – Вихідні дані та лінія регресії

Приблизними, але самим простим і наочним способом перевірки задовільності регресійної моделі є графічне представлення відхилень:

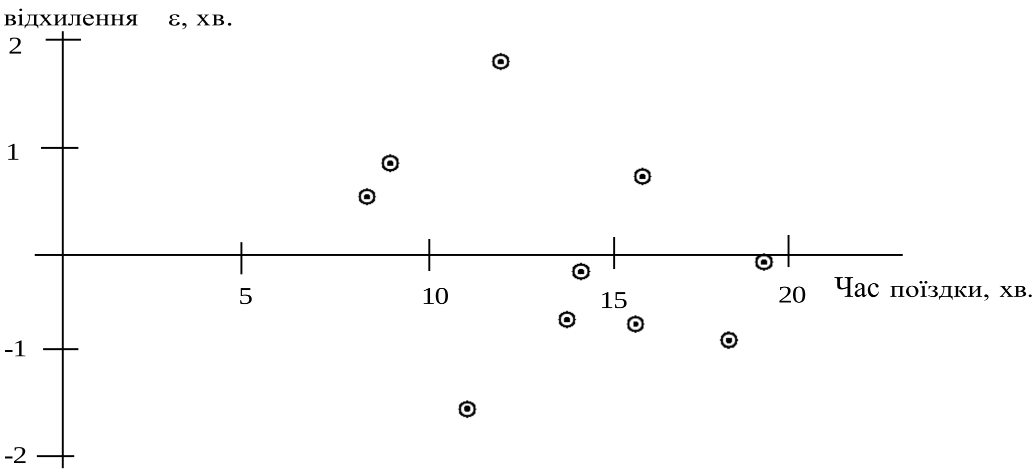


Рисунок 1.8 – Графік відхилень

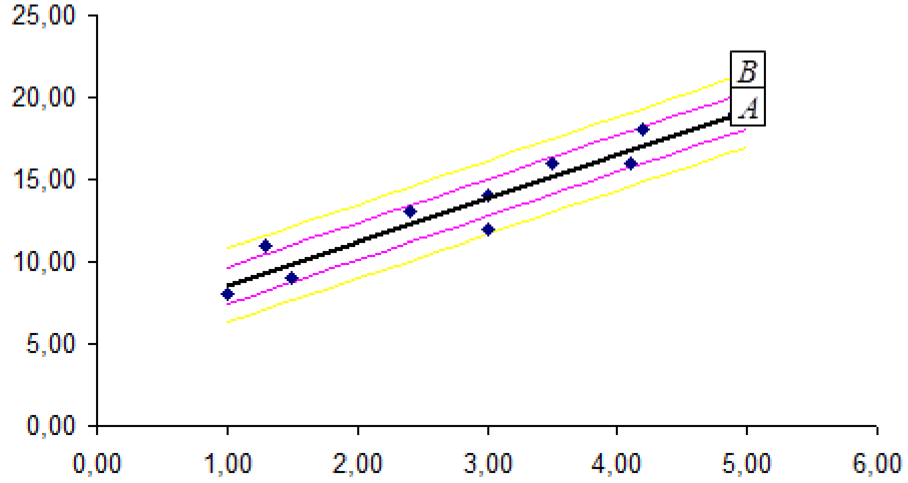


Рисунок 1.9 Графік відхилень A B

Відкладемо відхилення $\left(y_i^p - y_i\right)$ по осі Y, для кожного значення y_i. Якщо регресійна модель близька до реальної залежності, то відхилення будуть носити випадковий характер і їх сума буде близька до нуля. У розглянутому прикладі

$$\sum_{i=1}^{n}\left(y_i^p - y_i\right) = 0{,}004$$

Зазвичай мірою помилки регресійної моделі служить середнє квадратичне відхилення:

$$\sigma_\varepsilon = \left[\left(\sum_{i=1}^{n}\varepsilon_i^2\right)\Big/(n-2)\right]^{1/2} = \left\{\left[\sum_{i=1}^{n}(y_i^p - y_i)^2\right]\Big/(n-2)\right\}^{1/2}$$

Для нормально розподілених процесів приблизно 67% точок знаходиться в межах одного відхилення σ_ε від лінії регресії і 95% - в межах $2\sigma_\varepsilon$ (на рисунку 8 трубки A і B відповідно).

Вирішимо задачу прогнозування. Оскільки коефіцієнт детермінації R^2 має досить високе значення і відстань 2 милі, для якого треба зробити прогноз, перебуває в межах діапазону вихідних даних (таблиця 1), то ми можемо використовувати отримане рівняння лінійної регресії для прогнозування

$$y^*(2\text{мили}) = 5{,}913 + 2{,}660 \cdot 2 = 11{,}2 \text{ хв.}$$

При прогнозах на відстані, що не входять в діапазон вихідних даних, не можна гарантувати справедливість отриманої моделі. Це пояснюється тим, що зв'язок між часом і відстанню може змінюватися по мірі збільшення відстані. На час далеких перевезень можуть впливати нові фактори такі, як використання швидкісних шосе, зупинки на відпочинок, обід і т. п.

Таким чином, в результаті використання регресійного аналізу в пакеті Microsoft Excel ми:

- побудували рівняння регресії;
- встановили форму залежності та напрямок зв'язку між змінними - позитивна лінійна регресія, яка виражається в рівномірному зростанні функції;
- встановили напрямок зв'язку між змінними;
- оцінили якість отриманої регресійної прямої;
- змогли побачити відхилення розрахункових даних від даних вихідного набору;
- передбачили майбутнє значення залежної змінної.

Варіанти завдань для самостійного вирішення

Задача №1

Побудуйте регресійну модель (лінійну) для вихідних даних, наведених у таблиці. Для полегшення розрахунків вихідні дані містять тільки чотири пари значень (x_i, y_i).

Вихідні дані задачі №1

№ варіанта	Координати	Точки				x*
1	X	1	2	3	4	1.6
	Y	30	7	8	1	?
2	X	1	2	3	4	2.3
	Y	25	7	7	2	
3	X	9	5	2	3	2.9
	Y	25	7	7	2	?
4	X	1	2	3	4	2.6
	Y	15	10	7	0.5	?
5	X	10	3	6	4	8
	Y	25	7	7	2	?
6	X	9	5	2	3	2.5
	Y	15	8.5	7.5	5	?
7	X	2	3	7	8	7.5
	Y	11	8.5	6.5	5	?
8	X	10	3	6	4	9
	Y	15	7	8	6	?
9	X	2	3	4	5	4.5
	Y	13	9	8	7	?
10	X	1	2	3	4	1.5
	Y	7.5	7	5	3.5	?
11	X	1	2	3	4	3.6
	Y	13	9	8	7	?
12	X	3	4	6	10	8

	Y	7.5	7	6.5	3.5	?
13	X	3	4	5	6	7.8
	Y	9	7	5	3	?
14	X	7	5.6	13	14.7	15
	Y	7.5	7	5	3.5	?
15	X	9	5	2	3	5.7
	Y	13	9	8	7	?
16	X	3	4	6	8	5
	Y	7.5	7	6.5	5	?
17	X	2	3	7	8	7.5
	Y	9	9	8	7	?
18	X	9	10	11	12	10.5
	Y	13	9	8	7	?
19	X	1	2	3	4	3.5
	Y	5	4.5	3	3	?
20	X	11	12	13	16	13.6
	Y	7.6	8	6.5	4.2	?
21	X	5	6	7	8	6.5
	Y	5	4.5	3	3	?
22	X	9	10	12	14	12.5
	Y	8	7	6.5	4.2	?
23	X	7	8	9	10	9.6
	Y	8	7	6	4.2	?
24	X	1.5	2.5	3.5	4.5	3.9
	Y	5	4.5	3	3	?
25	X	1	2	5	6	3.9
	Y	5	4	3	3	?
26	X	1.5	2.4	3.8	6.9	4.1
	Y	5.5	5.5	4.8	1.1	?
27	X	1	2	3	4	3.6
	Y	12	3	9	5	?
28	X	1	2	3	7	2.8
	Y	5	5.5	4.8	1.1	?
29	X	11	12	13	16	14.1
	Y	0.25	0.19	5.2	8	?
30	X	1	2	3	4	3.4
	Y	13	4	10	6	?

Досліджуйте модель за допомогою режиму регресія в MS Excel і зробіть прогноз для x^*.

Задача № 2

Для вихідних даних, представлених в таблиці, були побудовані наступні регресійні моделі:

- $y = 6{,}067 - 0{,}085x$;
- $y = -2{,}017 + 3{,}957x - 0{,}367x^2$;
- $y = 5{,}918e^{-0{,}043x}$.

Вихідні дані завдання №2

X	3	8	5	10	7	6	4	9	1	2
Y	6	5	9	1	8	9	8	4	2	4

З допомогою графіка відхилень виберіть задовільну модель і перевірте свій вибір за допомогою відповідного розрахунку.

Задача №3

У таблиці представлені дані про ціни на комплектуючі для ПЕОМ. Комплектуючі виробляються різними компаніями-виробниками і розбиті на групи за своїми функціональними можливостями.

Вихідні дані завдання №3

Група	1	1	2	2	2	3	3	3	4	4
Ціна, $	50	60	70	80	95	100	115	120	105	120
Група	4	5	5	5	6	6	6	7	7	7
Ціна, $	130	110	150	190	120	130	220	145	265	270

Побудуйте графік вихідних даних і з його допомогою проаналізуйте застосовність методу найменших квадратів. Підтвердіть свої висновки за допомогою розрахунку (для лінійної моделі). Прокоментуйте економічні причини отриманого результату.

Лабораторно-практична робота № 2
Дослідження нейромережевих математичних моделей за допомогою програми Нейросимулятор

Мета роботи: набути навичок проектування нейронні мережі шаруватої структури, і дослідження отриманих нейромережевих математичних моделей

Теоретичні відомості

Програма нейросимулятор дозволяє проектувати, оптимізувати, навчати і тестувати нейронні мережі шаруватої структури, проводити дослідження отриманих нейромережевих математичних моделей, вирішувати за допомогою них завдання прогнозування, оптимізації, управління, розпізнавання образів. Рішення будь-якого завдання методом нейромережевого моделювання, як правило, включає виконання наступних етапів (рис.1).

Рисунок 2.1 – Алгоритм застосування методу нейромережевого математичного моделювання

Етап 1. Постановка завдання.

На цьому етапі визначаються цілі моделювання, встановлюються вхідні і вихідні параметри моделі, встановлюється структура (склад і довжина) вхідного вектора X, і вихідного вектора D. як компонент вхідного вектора X важливо вибрати значущі параметри, тобто ті, які мають істотний вплив на результат. Якщо є сумніви в значущості того чи іншого вхідного параметра, то його краще включити у вхідний вектор, розраховуючи, що в подальшому за допомогою створюваної нейромережі можна буде оцінити ступінь його впливу на результат і, якщо вона виявиться слабкою, то цей параметр виключити.

Вихідний вектор D формується таким, щоб його компоненти давали можливість отримати відповіді на всі поставлені питання.

Компоненти вхідного вектора X і вихідного вектора D являють собою числа. Значення може бути будь-яких величин, наприклад, температура тіла, артеріальний тиск, частота пульсу та ін. Це можуть бути також числа, що кодують наявність або відсутність яких-небудь ознак, наприклад, одиниця, якщо стать чоловіча і двійка, якщо стать жіноча. У деяких випадках, якщо дані нечіткі і є сумнів в їх правильності, корисно кодувати оцінку їх ймовірності.

Етап 2. Формування прикладів.

На цьому етапі формується вміст вхідних і вихідних векторів. В результаті створюється безліч пар Xq–Dq (q=1,...,Q). Кожна така пара становить приклад, що характеризує предметну область.

Значення компонент векторів Xq і Dq формують різними способами: шляхом проведення соціологічних досліджень, анкетування, спеціальних експериментів над предметною областю, беруть з мережі Інтернет, ЗМІ, з архівних матеріалів організацій та з інших джерел. Всі безліч прикладів розбивають на навчальне L і тестує T (рис. 2). Зазвичай обсяг тестуючого безлічі вибирають не менше 10% від навчальної. Який необхідний мінімальний обсяг навчального безлічі, залежить від завдання. Зазвичай рекомендується не менше п'ятдесяти прикладів. Проте в нашій практиці зустрічалися випадки, коли для вирішення завдання вистачало і десяти навчальних прикладів.

Рисунок 2.2 – Поділ прикладів предметної області на навчальне безліч L, тестуюча безліч T і підтверджує безліч P

В особливо відповідальних випадках рекомендується крім навчального безлічі L і тестуючого безлічі T, формувати ще й підтверджує безліч P з прикладів, що належать тієї ж самої предметної області, але не перетинається ні з безліччю L, ні з безліччю P (див. рис. 2.2).

Етап 3. Проектування мережі.

Структура персептрона вибирається з наступних міркувань. Число вхідних нейронів Nx має дорівнювати розмірності вхідного вектора X. Число вихідних нейронів Ny має дорівнювати розмірності вихідного вектора D. Число прихованих шарів, згідно теоремі Арнольда-Колмогорова-Хехт-Нільсена, має бути не менше одного. На наступних етапах число прихованих шарів може коригуватися, якщо це дозволить поліпшити якість роботи мережі.Число нейронів в прихованих шарах N розраховується за допомогою формул, що є наслідком з теореми Арнольда-Колмогорова-Хехт-Нільсена:

$$\frac{N_y Q}{1+\log_2(Q)} \le N_w \le N_y\left(\frac{Q}{N_x}+1\right)\left(N_x+N_y+1\right)+N_y,$$

$$N=\frac{N_w}{N_x+N_y},$$

де N_Y – розмірність вихідного сигналу; Q – число елементів безлічі навчальних прикладів; N_W – необхідне число синаптичних зв'язків;N_X – розмірність вхідного сигналу.

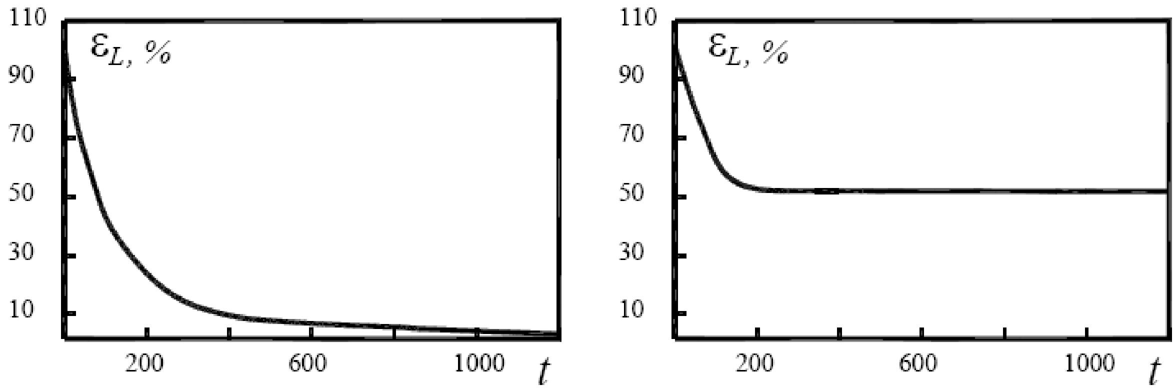
На наступних етапах число нейронів в прихованих шарах може ректуватися, якщо це дозволить поліпшити якість роботи мережі. Активаційні функції прихованих нейронів, згідно теоремі Арнольда-Колмогорова-Хехт-

Нільсена, рекомендується задати сигмоїдними, проте надалі, їх вид може бути змінений, якщо це дозволить поліпшити якість роботи мережі.

При коригуванні структури персептрона слід мати на увазі, що збільшення прихованих нейронів зазвичай дозволяє домогтися меншої помилки навчання, проте надмірне їх збільшення призводить до ефекту гіпервірності-втрати узагальнюючих властивостей мережі, що виражається в зростанні помилки Узагальнення.

Етап 4. Навчання мережі.

Навчання мережі – дуже важливий, але не остаточний етап створення нейромережевої інтелектуальної системи. Мета навчання – підібрати синап-тичні ваги Wij так, щоб на кожен вхідний вектор Xq безлічі навчальних прикладів мережа видавала вектор Yq, мінімально відрізняється від заданого вихідного вектора Dq. Ця мета досягається шляхом використання алгоритмів навчання нейронної мережі. Зазвичай пробується відразу кілька АЛ-горитмів навчання. Характерна крива навчання-залежність квадратичної і максимальної помилок навчання від числа епох навчання, наведена на рис. 3.



а *б*

Рисунок 2.3 – Характерні залежності помилок навчання від числа епох в разі, коли мережа навчається успішно (а), і коли процес навчання не дає бажаного результату (б)

Приклад здійснення навчання нейронної мережі в програмі Нейросимулятор наведено на рис 2.4.

Рисунок 2.4 – Приклад навчання нейронної мережі в програмі Нейросимулятор

Але може трапитися так, що мережа не захоче навчатися – помилка навчання зі збільшенням числа епох не буде прагнути до нуля (див. рис. 2.3, б). Розглянемо можливі причини цього небажаного явища.

1-я причина. Недостатня кількість прихованих шарів і прихованих нейронів.

Рекомендується збільшити число прихованих шарів і прихованих нейронів.

2-я причина. Попадання в локальний мінімум. Рекомендується» струснути " мережу, почати навчання з іншими початковими значеннями сил синаптичних зв'язків, або змінити метод навчання.

3-я причина. Попадання в яр. Рекомендується використовувати антиовражный метод навчання, наприклад, метод зворотного поширення з моментом.

4-я причина. Наявність в навчальному безлічі сторонніх викидів, що випадають із загальних закономірностей предметної області.Рекомендується виявити сторонні викиди, постаратися переконатися, що вони дійсно є не характерними для предметної області, є наслідком помилок, неточності вимірювань та ін., і видалити сторонні викиди з навчальної множини прикладів.

5-я причина. Наявність в навчальному безлічі суперечать один одному прикладів. Наприклад, одному і тому ж вектору симптомів відповідають різні діагнози захворювань.

Виявити такі приклади в навчальній множині можна шляхом його візуального аналізу, або шляхом застосування тих же методик, які використуються при виявленні сторонніх викидів. Потім слід розібратися в причинах виникнення суперечать прикладів. Деякі з прикладів можуть просто виявитися помилковими, і їх потрібно видаляти. Інша причина може бути пов'язана з тим, що в самій структурі вхідного вектора відсутні якісь параметри (наприклад, вік хворого, зріст, вага, колір його очей та ін), також надають вплив на результат. У цьому випадку рекомендується переглянути постановку задачі, збільшити розмірність вхідного вектора X, додавши додаткові параметри, які своїми значеннями забезпечать несуперечність прикладів навчальної вибірки.

6-я причина. Параліч мережі. Рекомендується змінити сигмоїдну активаційну функцію на ЛОДА-римическую, або на функцію Ле-Кана, або змінити спосіб передобробки вхідних параметрів.

7-я причина. Занадто велика швидкість навчання. Рекомендується зменшити швидкість навчання.

Етап 5. Тестування (перевірка) та оптимізація мережі.

Мета тестування – переконатися, що математична нейромережева модель адекватна модельованої предметної області і придатна для подальшого використання. Мета оптимізації – домогтися найкращих узагальнюючих властивостей мережі, тобто – мінімальної помилки узагальнення. Перевірка узагальнюючих властивостей проводиться на тестуючому безліч прикладів, тобто на тих прикладах, які не брали участь у навчанні мережі. Результати тестування корисно представити графічно у вигляді гістограми, на якій значення бажаних виходів персептрона (Dq) можна зіставити з прогнозними (Yq), тобто тими, які обчислив персептрон. Приклад такої гістограми наведено на рис. 4.

Рисунок 2.5 – Приклад гістограми, що показує співвідношення бажаних виходів мережі з прогнозними

Якщо різниця між компонентами бажаного вихідного вектора тестуючого безлічі прикладів Dq і прогнозними значеннями Yq виявиться незначною, то можна переходити до наступного етапу 6, не виконуючи оптимізацію мережі. Однак, щоб зайвий раз переконатися в адекватності разрабатываемой нейромережевої математичної моделі корисно повернутися на етап 2 і ті приклади, які були тестуючими, (або частина тестуючих прикладів) включити в навчальне множини, а частина прикладів, колишніх навчальними, зробити тестуючими. Після цього знову повторити етапи 3, 4, 5.

Якщо похибка узагальнення мережі виявиться неприйнятно великий, можна спробувати оптимізувати мережу. Оптимізація мережі полягає в підборі найбільш підходящої для даного завдання структури мережі-кількості прихованих шарів, кількості прихованих нейронів, кількості синаптичних зв'язків, виду і параметрів активаційних функцій нейронів. У деяких нейропакетах передбачена автоматична оптимізація мережі. Але іноді буває корисно виконати оптимізацію вручну, побудувавши графік залежності похибки узагальнення від числа прихованих нейронів (див. рис. 6) та інших параметрів персептрона, і вибравши за допомогою цих графіків структуру мережі, що забезпечує мінімальну похибку узагальнення.

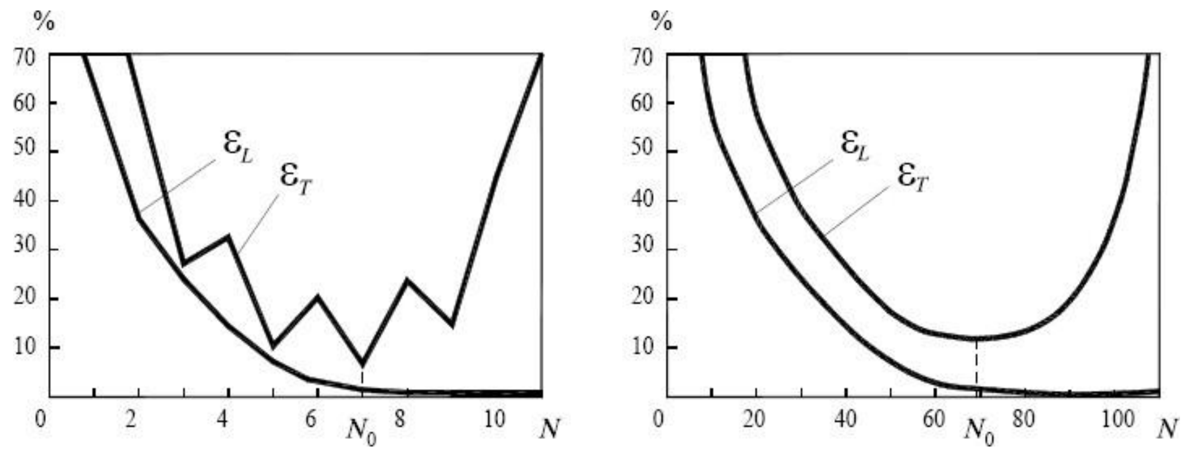


Рисунок 2.6 – Характерні залежності помилки навчання ε_L і помилки Узагальнення (тестування) ε_T від кількості нейронів прихованих шарів персептрона N

При виборі оптимальної структури мережі слід пам'ятати, що мета оптимізації мережі полягає в мінімізації похибки Узагальнення ε_T, але непогрішності навчання εL. Саме за величиною похибки узагальнення судять про якість мережі, про її узагальнюючих і, отже, прогностичних властивості. Похибка ж навчання-це всього лише проміжний результат. Бажано, щоб вона була невеликою, але домагатися її мінімального значення зовсім не обов'язково і навіть шкідливо, оскільки, як це видно з рис. 5, призводить до ефекту гіперрозмірності (перенавчання), тобто зростання похибки Узагальнення. Зрозуміло, що оптимізація нейронної мережі має на увазі многократ-ні повернення назад - на етапи 4, 3, 2, 1.

Після оптимізації мережі, її узагальнюючі властивості рекомендується про-вірити на прикладах підтверджує множини P. Справа в тому, що в процесі оптимізації мережа могла пристосуватися до прикладів тестуючого множини. А якщо ці приклади з будь-яких причин не характерні для всієї предметної області, то на інших прикладах, яких не було ні в тестуючому, ні в навчальному множинах, вона може дати несподівано велику помилку прогнозування. Для виключення такого явища, і щоб остаточно переконатися, що мережа має добрі узагальнюючі властивості не тільки на тестуючому безліч прикладів, обчислюють помилку прогнозу мережі εP на підтримуючому множині, тобто на тих прикладах, які ні в навчанні, ні в тестуванні не брали.

Результатом оптимізації та перевірки мережі є готова до використання нейромережева математична модель предметної області – інтелектуальна інформаційна система.

Етап 6. Дослідження моделі, прогнозування.

Шляхом проведення обчислювальних експериментів над математичною нейромережевою моделлю досягаються цілі моделювання, знаходяться відповіді на всі поставлені питання. Наприклад, можуть бути вирішені такі задачі, як оптимізація модельованого об'єкта, прогнозування його буду-них властивостей, виявлення закономірностей предметної області та ін Нейромережева математична модель, якщо вона правильно спроектована і навчена, ввібрала в себе закономірності модельованої предметної області. Вона реагує на зміну вхідних параметрів і поводиться так само, як вела б себе сама предметна область. І треба поставити над моделлю якомога більше експериментів. Треба постаратися витягти з цих віртуальних експериментів якомога більше корисної інформації.

Таким чином, в результаті роботи з програмою НЕЙРОСИМУЛЯТОР, в результаті виконання наведеного на рис. 1 алгоритму роботи з цією програмою виходять два види продуктів:

- після виконання етапу 5 створюється готова до використання інтелектуальна інформаційна система, яка є математичною моделлю предметної області,
- після виконання етапу 6 виходять результати дослідження математичної моделі, представлені у вигляді графіків, номограм, гістограм, корисних рекомендацій і висновків.

Порядок виконання роботи

1. Отримати завдання у викладача
2. Запустити програму Нейростимулятор
3. Спроектувати мережу
4. Навчити мережу
5. При виникненні помилки в мережі з'ясувати і усунути її
6. Протестувати спроектовану мережу
7. Результати дослідження математичної моделі представити у вигляді графіків, гістограм, корисних рекомендацій і висновків.

Лабораторно–практична робота №3
Побудова статичних і динамічних моделей

Мета роботи: оволодіти навичками побудови статичних і динамічних моделей

Теоретичні відомості

Дана практична робота виконується методом творчих завдань. На прикладі, що розбирається на практичному занятті, студенти вчаться аналізувати експериментальні дані, знаходити математичні залежності. В якості творчих завдань пропонуються таблиці експериментальних даних, отриманих студентами заздалегідь під час занять науковою роботою. При вирішенні творчих завдань, студенти можуть виконувати роль експертів, допомагаючи іншим студентам в групі знайти правильне рішення. Експертами вибираються студенти, які швидко побудували модель процесу.

Регресійний аналіз дозволяє оцінити ступінь зв'язку між змінними, пропонуючи механізм обчислення передбачуваного значення змінної з декількох відомих значень. Використовуючи регресійний аналіз, можна продовжити лінію тренда в діаграмі за межі реальних даних для пророкування майбутніх значень.

Опис послідовності дій при моделюванні:

Отримані в результаті експерименту дані залежності між величинами x і y можна представити у вигляді таблиці 1:

Таблиця 3.1- Експериментальні дані

x	x1	x2	x3	…	xn
y	y1	y2	y3	…	yn

Необхідно знайти емпіричну формулу $\mathbf{y = f(x)}$, що пов'язує між собою відповідні значення змінних так, щоб значення цієї функції при $\mathbf{x = x_i}$ можливо мало відрізнялися б від $\mathbf{y_i}$, отриманих з досвіду.

Цілі роботи:

1) побудувати математичну модель у вигляді емпіричної формули;
2) зробити оцінку параметрів моделі;
3) перевірити модель на адекватність.

Вибір загального вигляду емпіричної формули може бути проведений на основі теоретичних уявлень про характер досліджуваної залежності. В інших випадках доводиться підбирати формулу, порівнюючи криву, побудовану за даними спостережень з типовими графіками формул. Такими графіками можуть служити лінії тренда, які можна додати на діаграму Microsoft Excel.

Лінія тренда-це графічне представлення напрямку зміни ряду даних. Лінії тренда використовуються для аналізу помилок передбачення.

Точність апроксимації. Лінія тренда найбільшою мірою наближається до представленої на діаграмі залежності, якщо значення R-квадрат дорівнює або близько до 1. При апроксимації даних за допомогою лінії тренда значення R-квадрат розраховується автоматично. Отриманий результат можна вивести на діаграмі.

При цьому можна використовувати такі функціональні залежності:
Лінійна: Y = a + bx, де a –координата перетину осі абсцис і b –кут нахилу константи;

Логарифмічна: Y = ln x + b, де c і b – константи, ln – функція натурального логарифма.

Експоненціальна: Y = cebx, де c і b – константи, e – основа натурального логарифма.

Статечна: Y = cxb, де c і b-константи;

Поліноміальна: Y = b + c1x +c2x2 + c3x3+ ... + c6x6,де b і c1 ... c6 – константи.

Величина достовірності апроксимації – R. Число від 0 до 1, яке відображає близькість значень лінії тренда до фактичних даних. Лінія тренда найбільш відповідає дійсності, коли значення R в квадраті близько до 1. Воно також називається квадратом змішаної кореляції.

Класичний підхід до оцінювання параметрів лінійної регресії Y =a + bx заснований на методі найменших квадратів, який дозволяє отримати такі оцінки параметрів a і b, при яких сума квадратів відхилень фактичних значень результативної ознаки Y від розрахункових (теоретичних) f(x) буде мінімальна:

ОШ= □□□*f*(x)-Y □ □□□□min, $ОШ_{лин} = (f_{лин}\ (xi)-yi)^2$;

$ОШ_{эксп}= (f_{эксп}\ (xi)\ -yi)^2$; $ОШ_{лог}= (fлог\ (xi)-yi)^2$.

тобто з усього безлічі ліній регресії на графіку вибирається так, щоб сума квадратів відстаней по вертикалі між точками і цією лінією була б мінімальною.

Середня квадратична помилка обчислюється за формулою:

$$\delta = \sqrt{\frac{\sum(F(x) - Y)^2}{n}}$$

<u>Методика виконання роботи</u>

1. Оформити вихідні дані у вигляді зведеної таблиці Microsoft Excel.
2. З допомогою Майстра діаграм M. Excel побудувати графік залежності всього діапазону даних зведеної таблиці.
3. Побудувати лінію тренда.
4. Для отриманих математичних моделей зробити оцінку параметрів:
а) провести обчислення середньої квадратичної помилки δ;
б) порівняти δ з величиною достовірності апроксимації - R.
5. Перевірити модель на адекватність. Функція, якій відповідає мінімальне значення δ і максимальне значення R, є математичною моделлю, найбільш близько описує вихідні дані.

Приклад 1: Проведено дослідження залежності функціонально-технологічних властивостей (показник активної кислотності pH і лужності) водно-спиртових сумішей від варіації об'ємних часток спирту V1 і води V2 (V1+ V2 = 100 мл). В результаті експерименту були отримані наступні залежності (таблиця 2).

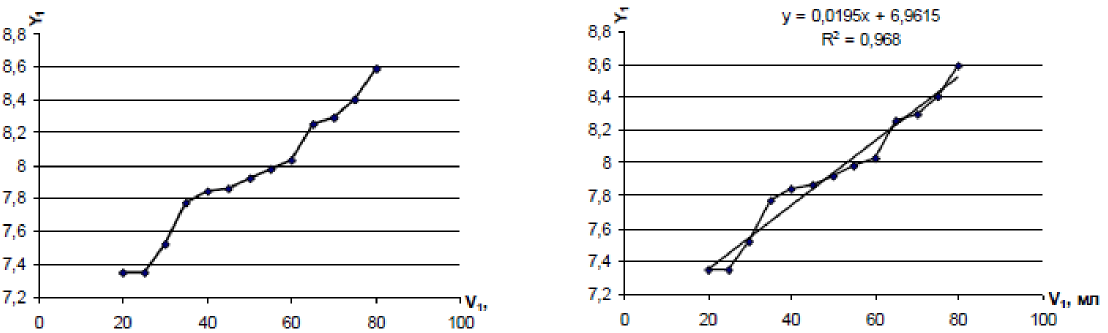
Таблиця 3.2 - Залежність лужності та показника активної кислотності pH від об'ємної частки спирту

Об'ємна частка спирту V1, мл	20	25	30	35	40	45	50	55	60	65	70	75	80
Об'ємна частка води V2, мл	80	75	70	65	60	55	50	45	40	35	30	25	20
pH Y1	7,35	7,35	7,52	7,77	7,84	7,86	7,92	7,98	8,03	8,25	8,29	8,4	8,6
Лужність Y2	3	2,7	2,6	2,4	2,1	1,9	1,8	1,6	1	1,3	1,2	1	0,9

Необхідно побудувати різні види залежностей pH і лужності спирту від об'ємної частки спирту у водно-спиртової суміші і вибрати рівняння лінії тренда найбільш відповідне дійсності для пророкування майбутніх значень.

Розв'язання задачі

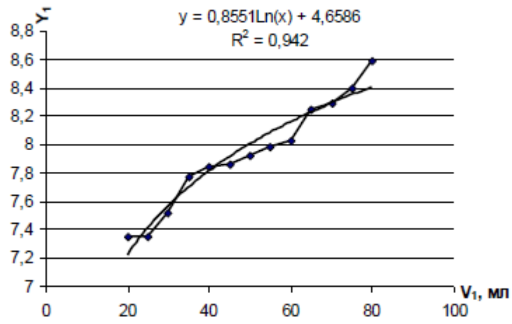
1. Здійснимо вибір прогнозної моделі, що дозволяє найбільш точно вказати залежність рівня pH водно-спиртової суміші від об'ємної частки спирту. Для цього побудуємо залежність величини Y1отV1(рис.1).



а б

Рисунок 3.1-Графік залежності pH від об'ємної частки спирту (x-об'ємна частка спирту, Y-рівень pH): а-без лінії тренда, б-з лінією лінійного тренда

Додамо до побудованого графіку лінію тренда, яка дозволяє однозначно визначити характер спостережуваної динаміки (рис. 2, 3).



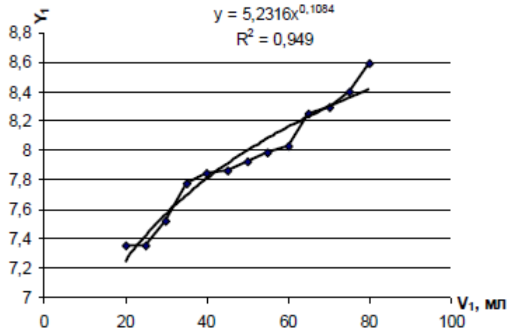


Рисунок 3.2-логарифмічний тренд Рисунок 3.3-статечний тренд

Отже, за значенням коефіцієнта детермінації R^2 (квадрата кореляції) найбільш значущою виявляється лінійна лінія тренду для даної моделі R2 приймає найбільше значення). Отримуємо математичні моделі:

$$f_{лин}(x) = 0,0195x +6,9615;$$

$$f_{эксп}(xi) \quad =5,2316x^{0,1084};$$

$$f_{лог}(xi) = 0,8551Ln(x)+4,6586.$$

2. Для отриманих моделей оцінимо параметри: а) проведемо обчислення середньої квадратичної помилки δ;

$ОШ_{лин} = (f_{лин}\ (xi)\text{-}yi)^2$; $ОШ_{эксп}= (f_{эксп}\ (xi)\ \text{-}yi)^2$; $ОШ_{лог}= (fлог\ (xi)\text{-}yi)^2$.

3. Порівняємо значення δ отриманих формул та величини достовірності апроксимації - R. Приходимо до висновку, що найкращим чином вихідні дані описує лінійна регресійна модель.

Для того щоб «поліпшити» побудовану емпіричну залежність побудуємо поліноміальну залежність 6-го ступеня. Залежність y = 2E-09x6 - 7E-07x5 + 8E-05x4 - 0,0053x3 + 0,1844x2 - 3,1966x + 28,928 має великий коефіцієнт детермінації R^2= 0,9958 (тобто частка варіації величини pH вже на 99,58% буде пояснюватися варіацією об'ємною часткою спирту у водноспиртової суміші).

Лабораторно – практична робота №4
Багатофакторна (множинна) регресія

Мета роботи: оволодіти навичками моделювання багатофакторної (множинної) регресії

Теоретичні відомості

При вирішенні практичних завдань дослідники стикаються з тим, що кореляційні зв'язки не обмежуються зв'язками між двома ознаками: результативною і факторною, насправді результативна ознака залежить від декількох факторних.

Наприклад. Інфляція тісно пов'язана з динамікою споживчих цін, роздрібним товарообігом, чисельністю безробітних, обсягом експорту та імпорту, курсом долара, кількістю грошей в обігу, обсягом промислового виробництва та іншими факторами.

Між факторами існують складні взаємозв'язки, тому їх вплив комплексний і його не можна розглядати як просту суму ізольованих впливів. Вивчення зв'язку між двома і більше пов'язаних між собою ознак носить назву множинної (багатофакторної) регресії.

Багатофакторний кореляційний і регресійний аналіз може бути використаний в економіко-статистичних дослідженнях:

- для приблизної оцінки фактичного і заданого рівней;
- в якості укрупненого нормативу (для цього достатньо в рівняння регресії підставити замість фактичних значень факторів їх середні значення);
- для виявлення резервів виробництва;
- для проведення міжзаводського порівняльного аналізу та виявлення на його основі прихованих можливостей підприємств;
- для короткострокового прогнозування розвитку виробництва.
- Математично завдання зводиться до знаходження аналітичного виразу, найкращим чином описує зв'язок факторних ознак з результативною, тобто до відшукання функції $_{1,2,...,k} = f(x_1, x_2, ..., x_k)$.

Побудова моделей множинної регресії включає кілька етапів.

1) вибір форми зв'язку (рівняння регресії).

Ускладнюється тим, що теоретична залежність між ознаками виражається великим числом різних функцій.

2) вибір типу рівняння.

Ускладнений тим, що для будь-якої форми залежності вибирається ряд рівнянь, які певною мірою будуть описувати ці зв'язки. Часто способом вибору типу рівняння є метод перебору різних рівнянь. Але практика побудови багатофакторних моделей взаємозв'язку показує, що всі реально існуючі залежності між соціально-економічними явищами можна описати, використовуючи такі типи моделей:

лінійна $$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + ... + a_k x_k \,;$$

степенева $$\hat{y} = a_0 x_1^{a_1} \cdot x_2^{a_2} + ... + x_k^{a_k} \,;$$

показова $$\hat{y} = e^{a_0 + a_1 x_1 + a_2 x_2 + ... + a_k x_k} \,;$$

параболічна $$\hat{y} = a_0 + a_1 x_1^2 + a_2 x_2^2 + ... + a_k x_k^k \,;$$

гіперболічна $$\hat{y} = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + ... + \frac{a_k}{x_k}$$

експонента: $$\tilde{y} = e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k}$$

3) Визначення розмірності моделі зв'язку, тобто визначення оптимального числа факторних ознак.

Є однією з основних проблем побудови множинного рівняння регресії. Чим більше факторних ознак включено в рівняння, тим воно краще описує явище, але тим складніше реалізувати модель. Скорочення розмірності моделі за рахунок виключення другорядних факторів може призвести до того, що така модель буде неадекватна досліджуваним явищам і процесам.

4) забезпечення достатнього обсягу сукупності для отримання несмещенних оцінок.

Оцінка параметрів лінійної множинної регресії

1) *у натуральному масштабі,* тобто для рівняння $\tilde{y} = a + b_1 x_1 + b_2 x_2 + ... + b_k x_k$ система нормальних рівнянь має вигляд:

$$\begin{cases} \sum y = n \cdot a + b_1 \cdot \sum x_1 + b_2 \cdot \sum x_2 + ... + b_k \cdot \sum x_k \\ \sum y \cdot x_1 = a \cdot \sum x_1 + b_1 \cdot \sum x_1^2 + b_2 \cdot \sum x_1 \cdot x_2 + ... + b_k \cdot \sum x_1 \cdot x_k \\ ....\,.........\,.........\,.........\,.........\,.........\,.........\,.........\,....... \\ \sum y \cdot x_k = a \cdot \sum x_k + b_1 \cdot \sum x_k \cdot x_1 + b_2 \cdot \sum x_k \cdot x_2 + ... + b_k \cdot \sum x_k^2 \end{cases}$$

Її рішення може бути знайдено, наприклад, методом визначників.

Обчислення параметрів лінійної множинної регресії можна провести за допомогою інструменту *Сервіс/Аналіз даних/Регресія.*

2) у стандартизованому масштабі:

$$\overline{t_y} = \beta_1 \cdot t_{x1} + \beta_2 \cdot t_{x_2} + \ldots + \beta_k \cdot t_{x_k},$$

де $t_y, t_{x_1}, t_{x_2}, \ldots, t_{x_k}$ – стандартизовані змінні

$$t_y = \frac{y - \bar{y}}{\sigma_y};$$

$$t_{x_i} = \frac{x_i - \bar{x}_i}{\sigma_{x_i}},$$

β_i – стандартизовані коефіцієнти регресії. Вирішують систему нормальних рівнянь виду:

$$\begin{cases} r_{yx_1} = \beta_1 + \beta_2 \cdot r_{x_2x_1} + \beta_3 \cdot r_{x_3x_1} + \ldots + \beta_k \cdot r_{x_kx_1} \\ r_{yx_2} = \beta_1 \cdot r_{x_1x_2} + \beta_2 + \beta_3 \cdot r_{x_3x_2} + \ldots + \beta_k \cdot r_{x_kx_2} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ r_{yx_k} = \beta_1 \cdot r_{x_1x_k} + \beta_2 \cdot r_{x_2x_k} + \beta_3 \cdot r_{x_3x_k} + \ldots + \beta_k \end{cases}$$

Вирішуючи її методом визначників, знайдемо *β-коефіцієнти.*

Визначення β-коефіцієнтів:

1) Знаходимо матрицю парних коефіцієнтів кореляції. Для двофакторної лінійної регресії вона має вигляд:

	y	x_1	x_2
y	1		
x_1	r_{yx_1}	1	
x_2	r_{yx_2}	$r_{x_1x_2}$	1

Найзручніше знайти цю матрицю Excel, використовуючи інструмент аналізу даних *Кореляція*. Для цього в головному меню виберіть *Сервіс/Аналіз даних/Кореляція.*

2) для стандартизованого рівняння регресії

$$\bar{t}_y = \beta_1 \cdot t_{x1} + \beta_2 \cdot t_{x_2}$$

отримаємо

$$\beta_{x_1} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{1 - \left(r_{x_1x_2}\right)^2}; \quad \beta_{x_2} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{1 - \left(r_{x_1x_2}\right)^2}.$$

Коефіцієнти "чистої" регресії пов'язані з β-коефіцієнтами наступним чином:

$$\beta_i = \frac{\sigma_{x_i}}{\sigma_y} \cdot b_i.$$

Методика побудови рівняння регресії при двофакторному регресійному аналізі

$$\bar{y} = a + bx + cz$$

призводить до наступних формул для оцінки параметрів:

$$b = \frac{r_{yx} - r_{yz} r_{xz}}{1 - r_{xz}^2} \cdot \frac{\sigma_y}{\sigma_x}, \quad c = \frac{r_{yz} - r_{yx} r_{xz}}{1 - r_{xz}^2} \cdot \frac{\sigma_y}{\sigma_z}, \quad a = \bar{y} - b\bar{x} - c\bar{z}.$$

Методика побудови рівняння регресії у вигляді статечної функції

$$y = a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \ldots \cdot x_k^{b_k} \cdot \varepsilon$$

Перетворимо її в лінійний вигляд:

$$\lg y = \lg a + b_1 \cdot \lg x_1 + b_2 \cdot \lg x_2 + \ldots + b_k \cdot \lg x_k + \lg \varepsilon,$$

де змінні виражені в логарифмах. Далі процедура МНК така ж, що й описана вище: будується система нормальних рівнянь і визначаються параметри, які потім слід потенціювати.

Оцінка тісноти зв'язку і статистичної значущості у множинній регресії

1) **коефіцієнт множинної детермінації** R^2,

$$R^2 = 1 - \frac{\sigma_{ост}^2}{\sigma_y^2} = \frac{\sum(\tilde{y} - \bar{y})^2}{\sum(y - \bar{y})^2};$$

2) **індекс множинної кореляції** R;

3) **лінійний коефіцієнт множинної кореляції**

$$\tilde{y} = a + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k)$$

$$R = \sqrt{\sum(\beta_i \cdot r_{yx_i})};$$

4) у разі двофакторної лінійної моделі **індекс множинної кореляції** R може бути знайдений за формулою:

$$R_{yx_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 + 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}.$$

5) **Скоригований індекс** (коефіцієнт) кореляції:

$$R_{коррект}^2 = 1 - \left(1 - R^2\right) \cdot \frac{n-1}{n-k-1};$$

k – число параметрів при змінних.

В статистичних пакетах прикладних програм у процедурі множинної регресії зазвичай наводиться скоригований коефіцієнт (індекс) множинної кореляції (детермінації).

6) дельта-коефіцієнти Δ_i:

$$\Delta_i = r_{yx_i} \cdot \frac{\beta_i}{R^2},$$

де r_{yx_i} – коефіцієнт парної кореляції між y і x_i;

R^2 – множинний коефіцієнт детермінації.

7) приватні коефіцієнти еластичності:

$$Э_i = b_i \cdot \frac{\bar{x}_i}{\bar{y}},$$

де b_i – коефіцієнт чистої регресії при факторі x_i;

$\bar{y}$ – середнє значення результативної ознаки;

$\bar{x}_i$ – середнє значення ознаки x_i.

Значимість рівняння множинної регресії в цілому оцінюється за допомогою *F-критерію* Фішера:

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m},$$

де n – число спостережень, m – число параметрів при змінній x. Якщо розрахункове значення критерію з $k_1 = m$ и $k_2 = n-m-1$ ступенями свободи більше табличного при заданому рівні значущості, то модель вважається значущою.

Прогнозування за рівнянням лінійної множинної регресії

$$\tilde{y}_{точеч} - t_{табл} \cdot \mu_{\tilde{y}} \le \tilde{y}_{прогн} \le \tilde{y}_{точеч} + t_{табл} \cdot \mu_{\tilde{y}}$$

де $\mu_{\tilde{y}}$ – помилка прогнозного значення, що обчислюється за формулою

$$\mu_{\tilde{y}} = S\sqrt{1+(X_0)^T \cdot (X^T X)^{-1} \cdot X_0}$$

$$X_0 = \begin{pmatrix} 1 \\ x_{1\,прогн} \\ x_{2\,прогн} \end{pmatrix} \qquad X = \begin{pmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{pmatrix}$$

для двофакторної моделі.

Мірою для оцінки включення фактора в модель служить приватний F-критерій, тобто F_{x_i}. Так, якщо оцінюємо значимість впливу фактора x_p після включення в модель факторів $x_1, x_2, \dots, x_{p-1}$, то формула приватного F-критерію прийме вигляд:

$$F_{x_p} = \frac{R^2_{yx_1x_2\dots x_p} - R^2_{yx_1x_2\dots x_{p-1}}}{1-R^2_{yx_1x_2\dots x_p}} \cdot \frac{n-m-1}{1}.$$

Якщо фактичне значення критерію з $k_1 = 1$ и $k_2 = n - m - 1$ ступенями свободи більше табличного при заданому рівні значущості, то додаткове включення фактора x_p в модель статистично виправдано і коефіцієнт регресії при даному факторі статистично значущий.

Оцінка значущості коефіцієнтів "чистої" регресії

$$\tilde{y} = a + b_1 x_1 + b_2 x_2 + ... + b_k x_k$$

Для кожного фактора використовується формула

$$t_{b_i} = \frac{b_i}{m_{b_i}},$$

де b_i – коефіцієнт чистої регресії при факторі x_i; m_{b_i} – середня квадратична помилка коефіцієнта регресії b_i,

$$m_{b_i} = \frac{\sigma_y \cdot \sqrt{1 - R^2_{yx_1x_2...x_k}}}{\sigma_{x_i} \sqrt{1 - R^2_{x_ix_1x_2...x_k}}} \cdot \frac{1}{\sqrt{n - m - 1}},$$

де σ_y – середнє квадратичне відхилення для ознаки *y*;

$R^2_{yx_1x_2...x_k}$ – коефіцієнт детермінації для рівняння множинної регресії;

σ_{x_i} – середнє квадратичне відхилення для ознаки x_i;

$R^2_{x_ix_1x_2...x_k}$ – коефіцієнт детермінації для залежності фактора x_i з усіма іншими факторами рівняння множинної регресії.

Практичні рекомендації по виконанню розрахунків за допомогою табличного редактора MS Excel

Досліджується залежність продуктивності праці y (т/год) від рівня механізації робіт x_1 (%),середнього віку працівників x_2 (років) і енергоозброєності x_3 (кВт/100 працюючих) за даними 14 промислових підприємств.

x_1	32	30	36	40	41	47	56	54	60	55	61	67	69	76
x_2	33	31	41	39	46	43	34	38	42	35	39	44	40	41
x_3	300	290	350	400	400	480	500	520	590	540	600	700	700	750
y	20	24	28	30	31	33	34	37	38	40	41	43	45	48

Неминуче:

1. Розрахувати параметри лінійного рівняння множинної регресії з повним переліком факторів.

2. Оцінити значимість рівняння в цілому, використовуючи значення множинного коефіцієнта кореляції і загального F-критерію Фішера.

3. Оцінити статистичну значимість параметрів регресійної моделі за допомогою t-критерію.

4. Дослідити колінеарність між факторами. При наявності мульти-колінеарності виключити будь-який фактор з рівняння регресії.

5. Побудувати нове рівняння множинної регресії, провести всі необхідні дослідження, аналогічні проведеним вище.

6. На підставі результатів п. 5 знайти

а) середні коефіцієнти еластичності фактора y від незалежних факторів;

б) прогнозне значення результату при значенні найважливішої пояснювальної змінної, що дорівнює максимальному спостереженому значенню, збільшеному на 10% , і при значенні другої пояснювальної змінної, що дорівнює мінімальному спостереженому значенню, зменшеному на 15%.

в) інтервальне передбачення значення y з надійністю 0,95.

1. Отримання протоколу розрахунку. Операція проводиться за допомогою інструменту *Аналіз даних/Регресія.* Вона аналогічна розрахунку параметрів парної лінійної регресії, розглянутої вище, тільки на відміну від парної регресії при заповненні рядка вхідний інтервал X в діалоговому вікні слід вказати відразу всі стовпці значень факторних змінних.

Результати аналізу мають вигляд:

Підведення підсумків

Регресійна статистика	
Множинний R	0,97517313
R-квадрат	0,950962633

Нормований R-квадрат	0,936251423
Стандартна помилка	2,038864298
Спостереження	14

Дисперсійний аналіз

	df	*SS*	*MS*	*F*
Регресія	3	806,1446094	268,7148698	64,64204
Задишок	10	41,56967627	4,156967627	
Всього	13	847,7142857		

	Коефіцієнти	*Стандартна помилка*	*t-статистика*
Y-перетини	5,711742473	6,18918556	0,922858495
x1	0,148601283	0,340417689	0,436526326
x2	0,064880259	0,162051974	0,400366976
x3	0,037784221	0,033824423	1,11706919

2. Оцінюємо статистичну значимість в цілому. Вивчивши результати, відзначаємо, що в цілому отримане рівняння лінійної множинної регресії

$$\tilde{y} = 5{,}71 + 0{,}15x_1 + 0{,}06x_2 + 0{,}04x_3$$

є статистично значущим. Дійсно, $F_{расч} = 64{,}64$. Порівняємо це число з критичним значенням критерію Фішера, отриманим при числі ступенів свободи $k_1 = m$ и $k_2 = n - m - 1$, де n – число спостережень, m – число параметрів при змінній *x*. В нашому випадку $k_1 = 3$, $k_2 = 14 - 3 - 1 = 10$. Критичне значення дасть функція *FРАСПОБР.* $F_{крит} = 3{,}71$, що істотно менше розрахункового значення.

Про частку варіації результативної ознаки у, поясненої побудованим рівнянням множинної регресії найкраще судити за значенням нормованого коефіцієнта кореляції, в даному випадку він дорівнює 0,9363. Тобто побудоване рівняння пояснює майже 94% всієї варіації ознаки *y*.

3. Оцінюємо статистичну значимість за окремими параметрами. Щоб оцінити статистичну значимість параметрів регресійної моделі за допомогою *t-критерію*, знайдемо відповідне нашим параметрам критичне значення за допомогою функції *СТЬЮДРАСПОБР* при заданому рівні значущості 0,05 і числі ступенів свободи $n - m - 1$. Коефіцієнт визнається значущим, якщо виконується нерівність $t_{расч} > t_{крит}$.

Отримаємо

	x_1	x_2	x_3
$t_{расч}$	0,44	0,4	1,12
$t_{крит}$	2,2281		

Таким чином, жоден з факторів не має статистично значимого коефіцієнта регресії, і побудоване рівняння для прогнозування непридатне.

4. Досліджуємо колінеарність між факторами. Матрицю парних коефіцієнтів кореляції можна отримати, використовуючи інструмент *Аналіз даних/Кореляція*. Заповнивши діалогове вікно,

y	x1	x2	x3
20	32	33	300
24	30	31	290
28	36	41	350
30	40	39	400
31	41	46	400
33	47	43	480
34	56	34	500
37	54	38	520
38	60	42	590
40	55	35	540
41	61	39	600
43	67	44	700
45	69	40	700
48	76	41	750

отримаємо наступний результат:

y	x1	x2	x3
20	32	33	300
24	30	31	290
28	36	41	350
30	40	39	400
31	41	46	400
33	47	43	480
34	56	34	500
37	54	38	520
38	60	42	590
40	55	35	540
41	61	39	600
43	67	44	700
45	69	40	700
48	76	41	750

	y	x1	x2	x3
y	1			
x1	0,968685	1		
x2	0,425757	0,361976	1	
x3	0,974574	0,991322	0,422666	1

Для оцінки мультиколінеарності факторів обчислимо визначник матриці парних коефіцієнтів кореляції факторів.

$$Det|R| = \begin{vmatrix} r_{x_1x_1} & r_{x_2x_1} & r_{x_3x_1} \\ r_{x_1x_2} & r_{x_2x_2} & r_{x_3x_2} \\ r_{x_1x_3} & r_{x_2x_3} & r_{x_3x_3} \end{vmatrix} = \begin{vmatrix} 1 & 0{,}36 & 0{,}99 \\ 0{,}36 & 1 & 0{,}42 \\ 0{,}99 & 0{,}42 & 1 \end{vmatrix} \approx 0{,}01$$

.

Оскільки визначник матриці міжфакторної кореляції близький до нуля, маємо мультиколінеарність факторів і випливає звідси ненадійність результатів множинної регресії.

Оцінка значущості мультиколінеарності факторів може бути проведена методом випробування гіпотези про незалежність змінних, тобто $H_0 : Det|R| = 1$. Доведено, що величина $\theta = \left[n - 1 - \frac{1}{6}(2m+5) \cdot \lg Det|R| \right]$ має наближений розподіл χ^2 з числом ступенів свободи $\frac{1}{2}m(m-1)$. Якщо фактичне значення χ^2 перевершує табличне (критичне), то гіпотеза $H_0 : Det|R| = 1$ відхиляється, і мультиколінеарність вважається доведеною.

Маємо $\theta = 14 - 1 - \frac{1}{6}(2 \cdot 3 + 5) \cdot \lg 0{,}01 = 13 - \frac{11}{6} \cdot (-2) = 13 + \frac{11}{3} \approx 16{,}67$.

Критичне значення χ^2 можна знайти через статистичну функцію *ХИ2ОБР(*α; $n-1$), де α – рівень значущості (за умовою 0,05), а n – число ступенів свободи. У нашому випадку ступенів свободи $\frac{1}{2} \cdot 3 \cdot (3-1) = 3$.

Отримаємо $\chi^2_{критическое} = 7{,}8$. $\chi^2_{критическое} < \theta$. Мультиколінеарність факторів знехтувати не можна.

Особливо висока колінеарність факторів x_1 и x_3, $r_{x_1 x_3} = 0{,}99$. Один з цих факторів слід виключити з рівняння регресії. Логічно виключити той, який має менший коефіцієнт парної кореляції. Оскільки $r_{yx_1} = 0{,}9687$, а $r_{yx_3} = 0{,}9746$, виключаємо фактор x_1.

5. Побудуємо регресію на факторах x_2 і x_3.

Підведення підсумків

Регресійна статистика	
Множинний R	0,974693901
R-квадрат	0,950028201
Нормований R-квадрат	0,940942419
Стандартна похибка	1,962415214
Спостереження	14

Дисперсійний аналіз

	df	*SS*	*MS*	*F*
Регресія	2	805,3524775	402,6762388	104,5621
Залишок	11	42,3618082	3,851073473	
Всього	13	847,7142857		

	Коефіцієнт	*Стандартна помилка*	*t-статистика*
Y-перетин	7,265656067	4,873196972	1,490942416
x2	0,031021017	0,136948082	0,226516625

x3	0,052435862	0,004030875	13,00855684

Отримаємо результати:
$\tilde{y} = 7,27 + 0,03x_2 + 0,05x_3$, $\overline{R}_{нормир} = 0,94$, $F_{расч} = 104,56$, що багато більше, ніж $F_{крит} = 3,71$.

	x_2	x_3
$t_{расч}$	0,22	13
$t_{крит}$	2,2281	

Таким чином, при досить задовільною значимості рівняння регресії в цілому, ми домоглися значущості коефіцієнта регресії при змінній x_3.

6.

а) Знайдемо коефіцієнти еластичності:

$$Э_i = b_i \cdot \frac{\bar{x}_i}{\bar{y}},$$

де b_i – коефіцієнт чистої регресії при факторі x_i;
$\bar{y}$ – середнє значення результативної ознаки;
$\bar{x}_i$ – середнє значення ознаки x_i.

Маємо

	y	x_2	x_3
Середнє	35,14285714	39	508,5714286
Еластичність		$0,03 \cdot \frac{39}{35,14} \approx 0,03$	$0,05 \cdot \frac{508,57}{35,14} \approx 0,72$

Таким чином, при зміні фактора x_2 (середнього віку працівників) на 1%, продуктивність зростає незначно, на 0,03%; при зміні фактора x_3 (енергоозброєності) на 1%, продуктивність праці збільшується на 0,72%.

б) Виконаємо прогнозування. Максимальне спостережене значення фактора x_3 – 750. Мінімальне значення фактора x_2 – 31. Прогнозні значення факторів:

$$x_2 = 31 - 0,15 \cdot 31 = 0,85 \cdot 31 = 26,35;\quad x_3 = 750 + 0,1 \cdot 750 = 0,9 \cdot 750 = 675.$$

Тоді $\tilde{y} = 7,27 + 0,03 \cdot 26,35 + 0,05 \cdot 675 \approx 41,81$.

в) довірчий інтервал для даного прогнозного значення y можна знайти, знаючи граничну помилку прогнозу $\Delta_{y_p} = t_\alpha \cdot m_{\tilde{y}}$, де t_α – відповідне критичне значення критерію Стьюдента, а $m_{\tilde{y}}$ – помилка прогнозного значення. У нашому випадку $t_{\alpha=0,1} = 3,1693$.

Помилку прогнозного значення функції регресії отримаємо за формулою

$$m_{\tilde{y}} = S\sqrt{X_0^{\ T} \cdot (X^T \cdot X)^{-1} \cdot X_0}.$$

Крок 1. Параметр S-стандартна помилка регресії наведено в останній регресійній статистиці $S = 1{,}96$.

Крок 2. Матриця X_0 складається з чисел:

$$X_0 = \begin{pmatrix} 1 \\ x_{2\,прогноз} \\ x_{3\,прогноз} \end{pmatrix}.$$

Тобто

$$X_0 = \begin{pmatrix} 1 \\ 26{,}35 \\ 675 \end{pmatrix},$$

$$X_0^T = \begin{pmatrix} 1 & 26{,}35 & 675 \end{pmatrix}.$$

Крок 3. Матриця X складається з чисел

$$X = \begin{pmatrix} n & \sum x_2 & \sum x_3 \\ \sum x_2 & \sum x_2^2 & \sum x_2 x_3 \\ \sum x_3 & \sum x_2 x_3 & \sum x_3^2 \end{pmatrix}.$$

Складаємо допоміжну таблицю:

	x_2	x_3	$x_2 \cdot x_3$	x_2^2	x_3^2
	…..	…..	….	…..	…..
Сума					

В данному вмпадку,

$$X = \begin{pmatrix} 14 & 585 & 7628{,}571 \\ 585 & 21544 & 281270 \\ 7628{,}571 & 281270 & 3909600 \end{pmatrix}.$$

Крок 4. Транспонуємо матрицю X. оскільки вона симетрична, то

$$X^T = X.$$

Крок 5. Знайдемо твір матриць $X^T * X$. В Excel це можна зробити за допомогою функції **МУМНОЖ.**

	58537523,04	2158299716	29989312607
$X^T \cdot X$	2158299716	79577299061	1,10572E+12
	29989312607	1,10572E+12	1,53641E+13

Крок 6. Знайдемо зворотну матрицю до матриці твору $(X^T \cdot X)^{-1}$. В *Excel* це можна зробити за допомогою функції **МОБР.**

$(X^T \cdot X)^{-1}$	0,281568563	-0,007773123	9,81695E-06
	-0,007773123	0,000215175	-3,13231E-07
	9,81695E-06	-3,13231E-07	3,38079E-09

Крок 7. Знайдемо твір матриць $X_0{}^T \cdot (X^T \cdot X)^{-1}$ (розмірність матриці твору 1×3).

$X_0{}^T \cdot (X^T \cdot X)^{-1}$	0,083373216	-0,002314683	3,84533E-06

Крок 8. Знайдемо твір матриць $X_0{}^T \cdot (X^T \cdot X)^{-1} \cdot X_0$ (розмірність матриці добудку 1x1, тобто тільки одне число).

$$X_0{}^T \cdot (X^T \cdot X)^{-1} \cdot X_0 = 0{,}024976921.$$

Крок 9. $m_{\tilde{y}} = 1{,}96\sqrt{0{,}024976921} \approx 0{,}31$.

Крок 10. $\Delta_{y_p} = t_\alpha \cdot m_{\tilde{y}} = 3{,}1693 \cdot 0{,}31 = 0{,}982483$.

Крок 11. Таким чином, прогнозне значення результату буде з ймовірністю 95% знаходиться в інтервалі $41{,}81 \pm 0{,}98$.

Завдання для самостійної роботи

Варіант 1

x_1	32	30	36	40	41	47	56	54	60	55	61	67	69	76
x_2	33	31	41	39	46	43	34	38	42	35	39	44	40	41
x_3	30	29	35	40	40	48	50	52	59	54	60	70	70	75
y	20	24	28	30	31	33	34	37	38	40	41	43	45	48

Варіант 2

x_1	55	46	40	39	35	29	31	75	68	66	60	54	59	53
x_2	33	42	45	38	40	30	32	40	39	43	38	34	41	37
x_3	50	45	39	40	34	30	30	74	69	66	59	54	60	52
y	33	32	30	29	27	23	19	47	44	42	40	39	37	36

Варіант 3

x_1	48	57	55	61	56	62	68	70	77	42	41	37	31	33
x_2	44	35	39	43	36	40	45	41	42	47	40	42	32	34
x_3	47	56	54	62	56	62	67	70	76	42	40	37	30	32
y	34	35	38	39	41	42	44	46	49	32	31	29	25	21

Варіант 4

x_1	52	54	45	39	38	34	28	30	74	67	65	59	53	58
x_2	36	32	41	44	37	39	29	31	39	38	42	37	33	40

x_3	52	53	45	38	38	34	28	31	73	66	65	60	52	57
y	35	32	31	29	28	26	22	18	46	43	41	39	33	36

Варіант 5

x_1	43	49	58	56	62	57	63	69	71	78	34	32	38	42
x_2	48	45	36	40	44	37	41	46	42	43	35	33	43	41
x_3	42	48	58	55	61	56	62	70	70	78	35	32	38	41
y	33	35	36	39	40	42	43	45	47	50	22	26	30	32

Варіант 6

x_1	52	57	51	53	44	38	37	33	27	29	73	66	64	58
x_2	32	39	35	31	40	43	36	38	28	30	38	37	41	36
x_3	52	56	50	53	45	37	37	32	28	30	72	66	64	59
y	37	35	34	31	30	28	27	25	21	17	45	42	40	38

Варіант 7

x_1	39	43	44	50	59	57	63	58	64	70	72	79	35	33
x_2	44	42	49	46	37	41	45	38	42	47	43	44	36	34
x_3	45	42	50	46	38	40	45	39	41	48	43	44	35	34
y	31	33	34	36	37	40	41	43	44	46	48	51	23	27

Варіант 8

x_1	63	57	51	56	50	52	43	37	36	32	26	28	72	65
x_2	40	35	31	38	34	30	39	42	35	37	27	29	37	36
x_3	39	38	35	35	32	31	28	28	25	25	21	15	45	40
y	39	37	36	34	33	30	29	27	26	24	20	16	44	41

Варіант 9

x_1	64	59	65	71	73	80	36	34	40	44	45	51	60	58
x_2	46	39	43	48	44	45	37	35	45	43	50	47	38	42
x_3	50	40	50	55	50	60	35	34	42	41	48	49	50	50
y	42	44	45	47	49	52	24	28	32	34	35	37	38	41

Варіант 10

x_1	46	52	61	59	65	60	66	72	74	81	37	35	41	45
x_2	51	48	39	43	47	40	44	49	45	46	38	36	46	44
x_3	46	52	60	58	64	61	65	72	74	80	38	34	40	44
y	36	38	39	42	43	45	46	48	50	53	25	29	33	35

Лабораторно - практична робота №5
Розробка моделі перцептрона Розенблатта

Мета роботи: оволодіти навичками розробки моделі перцептрона Розенблатта.

Порядок роботи: Засобами доступного мови програмування або за допомогою Excel створити перцептрон з 4 або більше вхідними елементами (S-елементи), двома або більше асоціативними елементами (A-елементи) і двома результуючими (R-елементами). Налаштуйте пороги спрацьовування а-елементів і R-елементів.

Оцініть граничну кількість образів, що розпізнаються даними перцептроном і промоделюйте його роботу, подаючи на вхід вектори образів, в тому числі і з шумом. У найпростішому випадку перцептрон повинен розпізнавати два зразки.

Теоретичні положення

Рисунок 5.1 – Перцептрон Розенблатта

У найбільш простому вигляді перцептрон (Рис. 5.1) складається з сукупності чутливих (сенсорних) елементів (S-елементів), на які надходять вхідні сигнали. S-елементи випадковим чином пов'язані з сукупністю асоціативних елементів (елементів), вихід яких відрізняється від нуля лише тоді, коли порушено достатньо велике число S-елементів, що впливають на один A-елемент. A-елементи з'єднані з реагуючими елементами (R-елементами) зв'язками, коефіцієнти посилення (v) яких змінні і змінюються в процесі навчання. Зважені комбінації виходів R-елементів складають реакцію системи, яка вказує на приналежність розпізнаваного об'єкта певного образу. Якщо розпізнаються тільки два образи, то в перцептроне встановлюється тільки один R-елемент, який володіє двома реакціями — позитивною і негативною. Якщо образів більше двох, то для кожного способу встановлюють свій R-елемент, а вихід кожного такого елемента представляє лінійну комбінацію виходів А-елементів:

$$R_j = \Theta_j + \sum_{i=1}^{n} v_{ij} x_i ,$$

де R_j — реакція j-го R-елемента; x_i — реакція i-го А-елемента; v_{ij} — вага зв’язку від i-го А-елемента до j-му R елементу; Θ_j — поріг j-го R-елемента. Аналогічно записується рівняння i-го А-елемента:

$$x_i = \Theta_i + \sum_{k=1}^{S} y_k ,$$

Сигнал y_k може бути безперервним, але найчастіше він приймає тільки два значення: 0 або 1. Сигнали від S-елементів подаються на входи А-елементів з постійними вагами рівними одиниці, але кожен А-елемент пов'язаний тільки з групою випадково вибраних S-елементів. Припустимо, що потрібно навчити перцептрон розрізняти два образи V_1 і V_2. Будемо вважати, що в перцептроні існує два R-елемента, один з яких призначений образу V_1, а інший — образу V_2. Перцептрон буде навчений правильно, якщо вихід R_1 перевищує R_2, коли розпізнається об'єкт належить образу V_1, і навпаки. Поділ об'єктів на два образи можна провести і за допомогою тільки одного R-елемента. Тоді об'єкту образу V_1 повинна відповідати позитивна реакція R-елемента, а об'єктам образу V_2 — негативна.

Перцептрон навчається шляхом пред'явлення навчальної послідовності зображень об'єктів, що належать образам V_1 і V_2. У процесі навчання змінюються ваги vi А-елементів. Зокрема, якщо застосовується система підкріплення з корекцією помилок, перш за все враховується правильність рішення, прийнятого перцептроном. Якщо рішення правильно, то ваги зв'язків всіх спрацювали а-елементів, що ведуть до R-елементу, що видав правильне рішення, збільшуються, а ваги непрацювали а-елементів залишаються незмінними. Можна залишати незмінними ваги спрацювали а-елементів, але зменшувати ваги непрацювали. У деяких випадках ваги спрацювали зв'язків збільшують, а непрацювали — зменшують.

Приклад

S-elements	**A Barriers**	**A-Elements (x)**	**weigth (v)**	**v*x**	**R Thres hold**	**R-elements**	
0	-1,3	C2+B2+B4	0,1	D2*E2			

0	-1,3	C3+B3+B5	1	D3*E3	5	G3+SUM($F$2:$F$4)	R-Black
0	-1,3	C4+SUM($B$3:$B$5)	1	D4*E4	5	G4+SUM($F$3:$F$5)	R-White
0	-1,3	C5+SUM($B$2:$B$4)	1,1	D5*E5			
0-Black							
1-White							
image1=0000 – білий фон							
image2=1111 – чорний фон							

Перцептрон працює так, що вихід R - елемента в комірці H3 буде більше при сигналі 0000 і менше при сигналі 1111

Завдання для дослідження

Використовуючи приклад, досліджуйте поведінку перцептрона при подачі на вхід різних сигналів. «Переобучите» перцептрон так, щоб він надійно розрізняв сигнал 1000 і 0001

«Навчіть» перцептрон розрізняти парні і непарні числа в двійковому вигляді

Парні 0000, 0010, 0100, 0110, 1000

Непарні 0001, 0011, 0101, 0111, 1001

Можливо, що для цього вам буде потрібно збільшити число А – елементів. Навчити персептрон розпізнаванню 0 і 1

Образи 0, в тому числі з шумом

1	1	1			1				1	
1		1		1		1		1		
1		1		1		1		1		1
1	1	1			1				1	

Образи 1, в том числі з шумом

	1				1				1	
1	1				1					
	1				1				1	
	1				1				1	

Лабораторно-практична робота №6
Моделювання багатошарового нелінійного персептрона в середовищі MathLab

Мета роботи: оволодіння навичками моделювання багатошарового нелінійного персептрона в середовищі MathLab

Загальні відомості

Нейронні мережі – це розділ штучного інтелекту, в якому для обробки сигналів використовуються явища, аналогічні тим, що відбувається в нейронах живих істот. Найважливіша особливість мережі, яка свідчить про її широкі можливості і величезний потенціал, полягає в паралельній обробці інформації всіма ланками. При величезній кількості міжнейронних зв'язків це дозволяє значно прискорити процес обробки інформації. У багатьох випадках стає можливим перетворення сигналів в реальному часі. Крім того, при великому числі міжнейронних сполучень мережа набуває стійкість до помилок, що виникають на деяких лініях. Функції пошкоджених зв'язків беруть на себе справні лінії, в результаті чого діяльність мережі не зазнає суттєвих збурень.

Інша не менш важлива властивість-здатність до навчання і узагальнення накопичених знань. Нейронна мережа має риси штучного інтелекту. Натренована на обмеженій безлічі даних мережа здатна узагальнювати отриману інформацію і показувати хороші результати на даних, що не використовувалися в процесі навчання.

Характерна особливість мережі полягає також у можливості її реалізації із застосуванням технології надвеликої ступеня інтеграції. Різниця елементів мережі невелика, а їх повторюваність величезна. Це відкриває перспективу створення універсального процесора з однорідною структурою, здатного переробляти різноманітну інформацію.

У навчально-методичному посібнику в короткій формі наведено теоретичні відомості про основні парадигми НС та приклади їх реалізації в комп'ютерному середовищі математичного моделювання MATLAB для вирішення типових завдань.

Алгоритм зворотного поширення помилки

Нехай визначена тришарова нейронна мережа з n входами, m виходами і l прихованими між ними елементами, тоді необхідно розглянути і побудувати два шари ваг: від входів до прихованих елементів і до виходу, тобто (W_1, W_2).

Призначення алгоритму зворотного поширення помилки – настройка всіх шарів багатошарової структури. Розглянемо роботу алгоритму на прикладі мережі з одним прихованим шаром і одним виходом (рис. 6.1). Перетворення вхідних сигналів, що задаються нейронною мережею, визначаються наступними формулами:

$$F(<W, X>) = 1/(1+\exp(-W^T X));$$

$$O_1 = 1/(1+\exp(-W_1^T X));$$
$$O_2 = 1/(1+\exp(-W_2^T O_1)).$$

Загальна функція помилки залежить від ваг всіх шарів, в нашому випадку від вектора W_2 та від матриці W_1:

$$E(W_1, W_2) = 1/2(Y - 1/(1+\exp(-W_2^T O_1)))^2,$$

де Y – вихід, який заданий в навчальній вибірці.

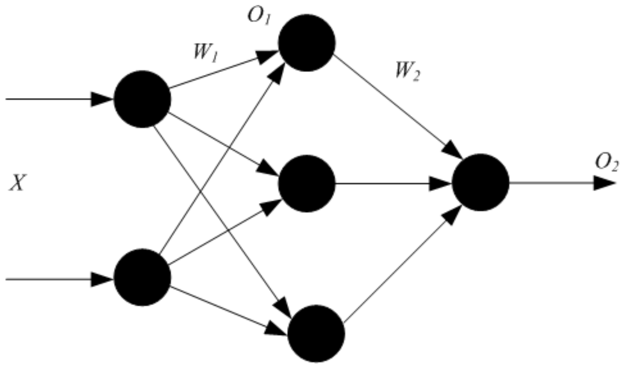


Рисунок 6.1 – Багатошаровий персептрон

Тепер необхідно визначити приріст кожної ваги за допомогою приватних похідних:

$$\frac{\partial E(W_2, W_1)}{\partial W_2};$$
$$\frac{\partial E(W_2, W_1)}{\partial W_1}.$$

Для багатошарової архітектури приватні похідні помилки по матриці ваг кожного шару визначаються за формулою похідної складної. У разі уніполярної сигмоїди правило зміни ваг буде наступним:

$$W_2 = W_2 + h \cdot (Y - O_2) \cdot O_2 \cdot (1 - O_2) \cdot O_1;$$
$$\partial = (Y - O_2) \cdot O_2 \cdot (1 - O_2);$$
$$W_1 = W_1 + h \cdot \partial \cdot W_1 \cdot (1 - O_1) \cdot O_1 \cdot X.$$

Таким чином, метод зворотного поширення помилки дозволяє змінювати ваги проміжних шарів, хоча бажані значення на проміжних шарах не задані.

Опис основних функцій

Функція *newff* створює нейронну мережу прямого поширення сигналу, що навчається за допомогою алгоритму зворотного поширення помилки:

net = newff(PR, [S1 S2 SNl], {TF1 TF2 TFNl), BTF, BLF, PF).

Розглянемо параметри функції *newff*: *PR* – матриця інтервалів значень *R* вхідних елементів, що задаються мінімальним і максимальним значеннями; *Si* – розмір *i*-го шару, для *N* верств*; TFi* – функція активації *i*-го шару, за замовчуванням використовується функція tansig – гіперболічний тангенс*; BTF* – функція навчання мережі методом зворотного поширення помилки, за замовчуванням використовується функція *traingdx; BLF* – функція зміни ваг при навчанні, за замовчуванням використовується *learngdm; PF* – функція вимірювання помилки, за замовчуванням *mse.* Функція *newff* повертає багатошарову нейронну мережу прямого і зворотного поширення сигналу та помилки відповідно. Функції активації можуть бути обрані з наступного переліку: гіперболічний тангенс *tansig*, логістична сигмоїда *logsig* або лінійна функція *purelin.*

ПОРЯДОК ВИКОНАННЯ РОБОТИ

Приклад 1 Створення та навчання нейронної мережі за допомогою алгоритму зворотного поширення помилки

Задамо за допомогою графіка вихідну функцію:

```
% входи НС
P = [0 1 2 3 4 5 6 7 8];
% бажані реакції НС
T = [0 0.44 0.88 0.11 -0.66 -0.95 -0.45 0.18 0.92];
% зображення апроксимованої функції
plot(P, T, 'o');
```

Використовуємо функцію *newff*, щоб створити двошарову мережу прямого поширення. Нехай мережа має входи з інтервалом значень від 0 до 8, перший шар з 10 нелінійними сигмоїдальними, другий - з одним лінійним нейронами. Використовуємо для навчання алгоритм зворотного поширення помилки (*backpropagation*) Левенберга – Марквардта (рис. 6).

```
% створення двошарової НС прямого поширення з інтервалом
% значень входів від 0 до 8, причому перший шар містить
% 10 нелінійних сигмоид, а другий — один лінійний нейрон.
% Для навчання використовується алгоритм зворотного поширення
% % помилки (back propagation).
net = newff([0 8], [10 1], {'tansig' 'purelin'},'trainlm');
% імітація роботи необученной НС
yl = sim (net, P);
% зображення результатів роботи необученої НС
plot(P, T, 'o', P, yl, 'x') ;
% Навчимо мережу на 100 епохах з цільовою помилкою 0.01:
% установка кількості проходів
```

net.trainParam.epochs = 50;
% установка цільового значення помилки
net.trainParam.goal = 0.01;
% навчання НС (рис. Шість)
net = train(net, P, T) ;
% імітація роботи навченої НС
y2 = sim(net, P);
% зображення результатів роботи НС (рис. Сім)
plot(P, T, 'o', P, yl, 'x', P, y2, '+');

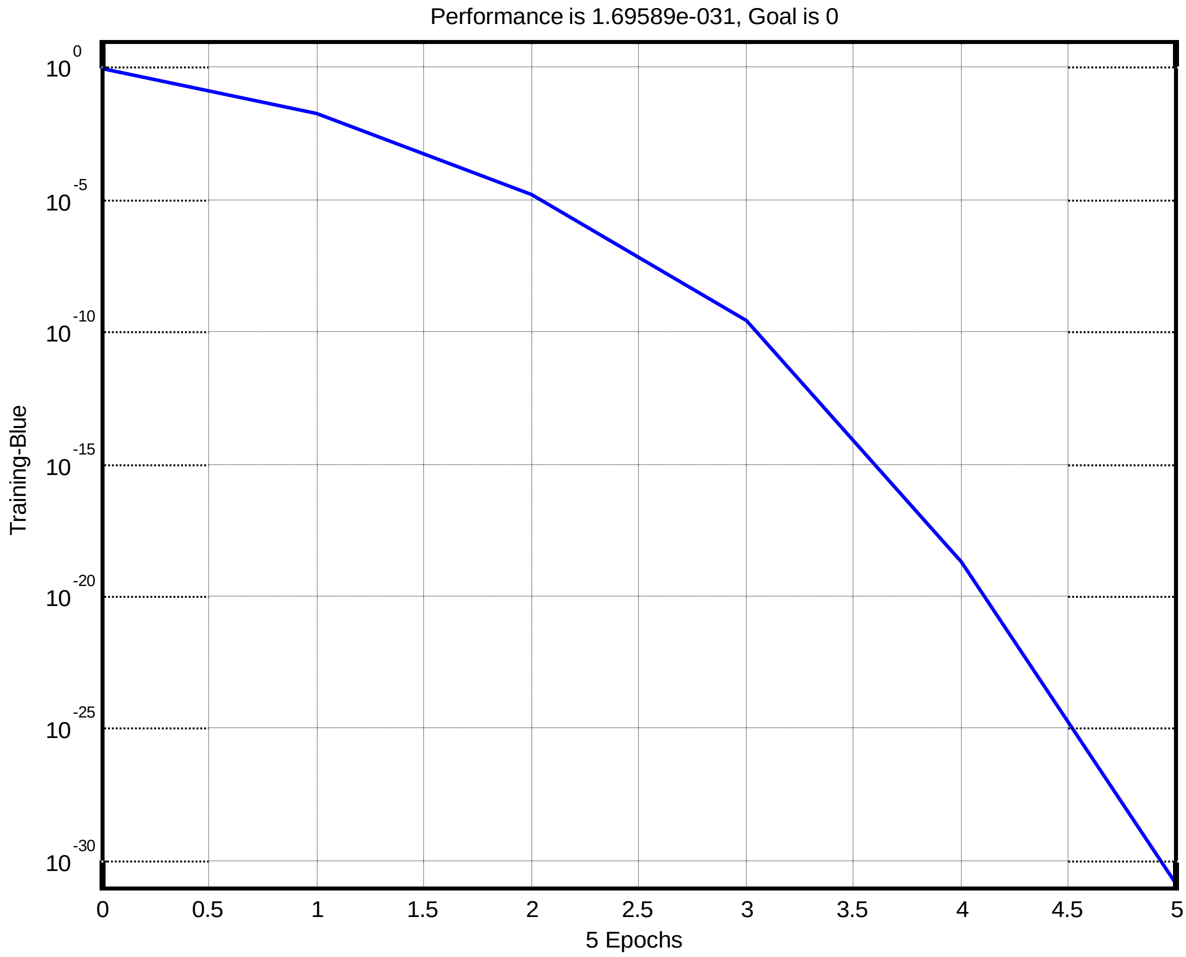


Рисунок 6.2 – Графік навчання двошарового персептрона

Для дослідження роботи алгоритму зворотного поширення помилки скористаємося прикладом, вбудованим в Matlab toolbox, набравши команду demo.

У діалоговому вікні необхідно послідовно вибирати пункти меню: Toolboxes->Neural Network->Other Demos->Other Neural Network Design textbook demos->Table of Contents->10-13->Backpropagation Calculation.

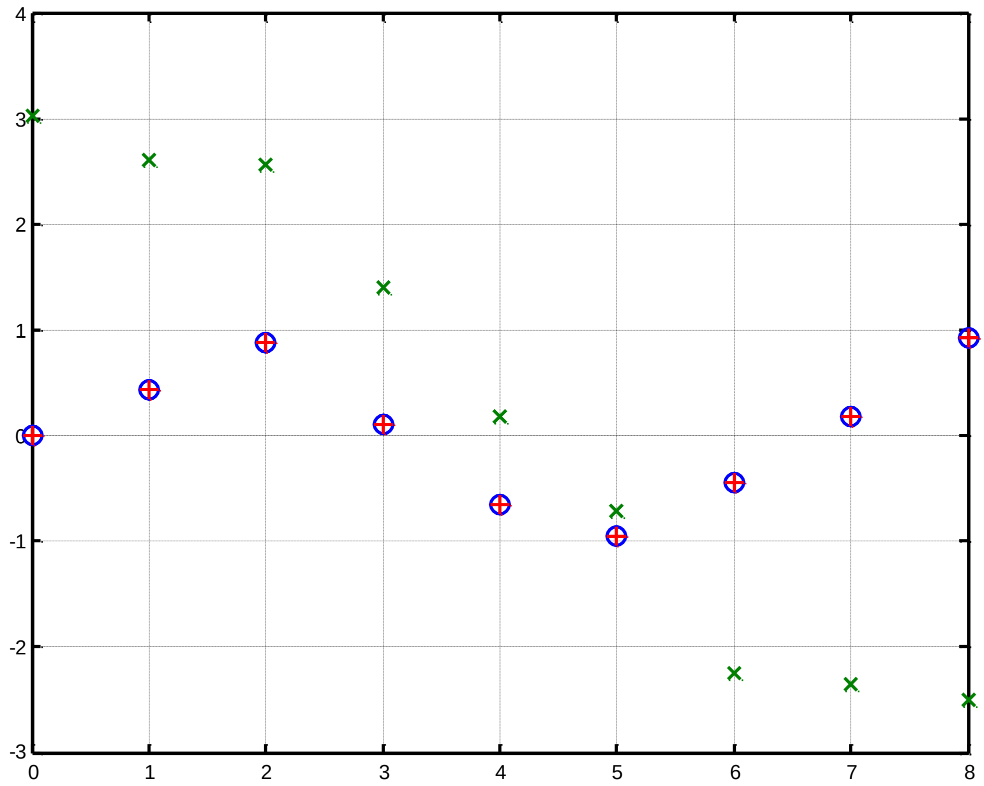


Рисунок 6.3 – Результат апроксимації векторів двошаровим персептроном

У прикладі використовується двошаровий персептрон з двома нелінійними нейронами в першому шарі і одним в другому. Дія алгоритму зворотного поширення помилки розбито на наступні кроки: призначення входу і бажаного виходу, прямий прохід вхідного сигналу до виходу, зворотне поширення помилки, зміна ваг. Змінні, що дозволяють простежити роботу алгоритму зворотного поширення помилки, позначені наступним чином:

P – вхідний сигнал;

W_1(i) – вектор ваг першого шару, W_1(1) – вага зв'язку, передає вихідний сигнал на перший нейрон, aW_1(2) – на другий;

W_2 (i) – вектор ваг другого шару, W_2 (1 – - Вага зв'язку, що передає вихідний сигнал з першого нейрона в другий шар, a_2 (2) - з другого;

B_1(i) – вектор порогових значень (*bias*) нейронів першого шару, *i* = 1, 2;

B_2-порогове значення (*bias*) нейрона другого шару;

N_1(i) – вектор виходів першого шару, *i* = 1, 2;

N_2 – вихід другого шару;

A_1(i) – вектор вихідних сигналів першого шару після виконання функції активації (сигмоид), *i* = 1, 2;

A_2 – вихід другого шару після виконання функції активації (лінійної);

LR-коефіцієнт навчання.

Нехай вхідний сигнал P = 1,0, тоді бажаний вихід $t = 1 + \sin(p * \pi / 4) = 1{,}707$.

Результати виконання етапів алгоритму представлені в табл. 6.1.

Таблиця 6.1 Результати поетапного виконання алгоритму зворотного поширення помилки

Етап	***Пряме поширення вхідного сигналу***	***Зворотне поширення помилки***	***Зміна ваг***
$A_1(1)$, $A_1(2)$	*$Logsig(W_1P+B_1)$ = = [0,321, 0,368]*	*Не виконується*	*Не виконується*
A_2	*$purelin(W_1P+B_1)$ = 0,446*	*Теж*	*Теж*
e	*$t - A_2 = 1{,}261$*	»	»
$N_1(1)$, $N_1(2)$	*Не виконується*	$\frac{\partial \log sim(N_1)}{\partial N_1 \cdot W_2 \cdot N_2} =$ $= [0{,}049,\ 0{,}100]$	»
N_2	*Теж*	$-2 \cdot \frac{\partial purelin(N_2)}{\partial N_2 \cdot e} = -2{,}522$	»
$W_1(1)$ $W_1(2)$	»	*Не виконується*	$W_1 = W_1 - \mathrm{lr} \cdot N_1 \times P =$ $= [-0{,}265, -0{,}420]$
$B_1(1)$, $B_1(2)$	»	*Теж*	$B_1 = B_1 - \mathrm{lr} \cdot N_1 =$ $= [-0{,}475, -0{,}140]$
B_2	»	»	$B_2 = B_2 - \mathrm{lr} \cdot N_2 = 0{,}732$
$W_2(2)$	»	»	$W_2 = W_2 - \mathrm{lr} \cdot N_2 \times N_1 =$ $= [0{,}171,\ 0{,}077]$

КОНТРОЛЬНІ ПИТАННЯ

1. Яким алгоритмом навчають багатошарові НС?
2. З яких основних етапів складається алгоритм зворотного поширення помилки?
3. Чому алгоритм зворотного поширення помилки відноситься до класу алгоритмів градієнтного спуску?
4. Як впливає функція приналежності на правило зміни ваг в зворотному алгоритмі поширення помилки?
5. Яка функція в середовищі MATLAB створює НС прямого поширення?
6. Які функції активації можуть бути призначені для нейронів НС прямого поширення?

Лабораторно-практична робота №7
Прогнозування випадкових процесів емулятором нейронної мережі NeuroPro 0.25

Мета роботи – оволодіння з основними параметрами нейронної мережі та дослідження впливу їх зміни на час навчання нейронної мережі і процес розв'язання задачі.

Опис емулятора

Даний програмний продукт являє собою менеджер учнів штучних нейронних мереж, що працює в середовищі MS Windows 95 або MS Windows NT 4.0 і дозволяє робити наступні базові операції:

1) Створення нейропроекта;

2) Підключення до нейропроекту файла (бази) даних у форматі *dfb (dBase, FoxBase, FoxPro, Clipper)* або *db (Paradox);*

3) Редагування файлу даних — зміна існуючих значень і додавання нових записів в базу даних; збереження даних в іншому форматі;

4) Додавання в проект нейронної мережі шаруватої архітектури з числом шарів нейронів від 1 до 10, числом нейронів у шарі — до 100:

5) навчання нейронної мережі вирішенню завдання прогнозування або класифікації. Нейронна мережа може одночасно вирішувати кілька завдань прогнозування (прогнозування кількох чисел), так і декілька задач класифікації, а також одночасно завдань і прогнозування та класифікації.

6) тестування нейронної мережі на файлі даних, отримання статистичної інформації про точність вирішення завдання;

7) Обчислення показників значущості вхідних сигналів мережі, збереження значень показників значущості в текстовому файлі на диску;

8) Спрощення нейронної мережі;

9) Генерація і візуалізація вербального опису нейронної мережі, збереження вербального опису в текстовому файлі на диску;

10) Вибір алгоритму навчання, призначення необхідної точності прогнозу, налаштування нейронної мережі.

Від наявних в даний час нейромережевих програмних продуктів даний продукт відрізняє наявність можливостей цілеспрямованого спрощення нейронної мережі для подальшої генерації вербального опису. Саме наявність розвинених можливостей щодо спрощення мережі у сукупності з побудовою її вербального опису і надає пропонованого продукту нові споживчі властивості — можливість породження знань з даних таблиці. Під знаннями тут розуміється текст, що пояснює процес вирішення нейронною мережею завдання. А оскільки одним з переваг нейронних мереж є можливість вирішення неформалізованих задач класифікації і прогнозу (тих завдань, явний алгоритм розв'язання яких не відомий), то даний текст запропонує один з алгоритмів розв'язання такої задачі.

Застосування даного програмного продукту можливо в тих областях, де нейронні мережі традиційно і з успіхом застосовуються, а саме, в медицині, екології кліматології метеорології, при побудові моделей технічних об'єктів та їх ідентифікації, в економіці (прогнозування курсів валют, акцій і т. д.) і взагалі для вирішення будь-якої задачі класифікації чи прогнозу, яка вирішується при наявності вибірки даних і для вирішення якої раніше використовувалися традиційні математичні методи (регресійний аналіз, непараметрична статистика та інші), однак не була досягнута необхідна точність прогнозу. Оскільки' на основі однієї таблиці даних може бути отримано декількох напівемпіричних теорій (кількох нейронних мереж мінімальної структури, правильно вирішують одну і ту ж задачу), то можливе рішення деяких завдань когнітології і планування оптимізуючого експерименту.

Саме нейромережева обробка даних в даний час і сформувала нову хвилю в розвиненій штучного інтелекту. Роботи по вербалізації структури нейронних мереж наближають нейромережеве напрямок штучного інтелекту до класичних експертним системам, роблячи можливим розуміння і пояснення процесу прийняття рішення нейронною мережею.

Опис лабораторної роботи

У лабораторній роботі використовуються емулятор NeuroPro 0.25, в емуляторі як приклад наводиться можливість прогнозу результатів виборів. У ньому також є можливість цілеспрямованого спрощення нейронної мережі для подальшої генерації словесного опису.

При спрощенні нейронної мережі можливе виконання наступних операцій:

- скорочення числа вхідних сигналів нейронної мережі шляхом видалення вхідних сигналів, найменш значущих для прийняття рішення мережею;
- скорочення числа нейронів мережі шляхом видалення нейронів, найменш значущих для прийняття рішення мережею;
- комплексне, рівномірне спрощення нейронної мережі. Для кожного нейрона мережі виконується скорочення числа приходять на нього сигналів до максимально можливого числа, що задається користувачем;
- скорочення числа зв'язків в нейронній мережі шляхом видалення зв'язків, найменш значущих для прийняття мережею рішення;
- бінаризація зв'язків в нейронній мережі-приведення ваг синапсів до значень-1 і 1 або значень з більш широкого набору виділених значень.

Наявність можливостей щодо спрощення мережі в сукупності з побудовою її словесного опису дає можливість відкинути фактори моделі, що незначно впливають на результат, і спростити вихідну модель.

Оскільки одним з переваг нейронних мереж є можливість вирішення неформалізованих задач класифікації і прогнозу тих завдань, явний алгоритм розв'язання яких не відомий, даний текст запропонує один з можливих алгоритмів розв'язання такої задачі.

Емулятор нейронних мереж NeuroPro можна використовувати як систему для соціального прогнозу, що пророкує результат виборів. Можливості емулятора розглядаються на прикладі виборів президента США.

Приклад:

Яка партія переможе на чергових виборах у США-правляча чи опозиційна? На перший погляд здається, що це залежить від особистостей кандидатів і від їх програм.

Виявляється, що якщо передвиборчі кампанії всіх кандидатів відпрацьовані сумлінно і всі учасники зробили все можливе, то вибір практично зумовлюється об'єктивними ознаками ситуації і не залежить ні від програм, ні від особистостей, ні від назв партій, а тільки від того, до якої партії належить правлячий президент. В одних ситуаціях перемагає правляча партія, в інших – опозиція.

Результати виборів можна передбачити на підставі відповідей на 12 запитань:

1) Правляча партія була при владі більше одного терміну?
2) Правляча партія отримала більше 50% на минулих виборах?
3) У рік виборів була активна третя партія?
4) Була серйозна конкуренція при висуненні від правлячої партії?
5) Кандидат від правлячої партії був президентом у рік виборів?
6) Рік виборів був часом спаду або депресії?
7) зростання середнього національного валового продукту на душу населення більше 2,1%?
8) правлячий президент зробив істотні зміни в політиці?
9) Під час правління були суттєві соціальні хвилювання?
10) Адміністрація правлячої партії винна в серйозній помилці або скандалі?
11) Кандидат правлячої партії – національний герой?
12) Кандидат опозиційної партії – національний герой?

Емулятору NeuroPro для навчання надається інформація про підсумки виборів в США за 100 років. Для всіх виборів відомі відповіді на 12 питань і яка партія, коли перемогла. Після навчання нейронна мережа дає свій варіант для додаткового набору відповідей, тобто пророкує результати виборів.

Дані для емулятора представлені у вигляді таблиці Excel, де в якості полів представлені 12 питань, описані вище, а в якості записів – відповіді на всі ці питання по всім рокам проведення виборів, починаючи з 1860 і закінчуючи 1980. Нижче представлено опис абревіатур всіх полів:

MORE1 – правляча партія була при владі більше одного терміну;

MORE5 – правляча партія отримала більше 50% голосів виборців на минулих виборах;

THIRD – у рік виборів була активна третя партія;

CONC – була серйозна конкуренція при висуненні від правлячої партії;

PREZ – кандидат від правлячої партії був президентом у рік виборів;

DEPR – рік виборів був часом спаду або депресії;

VAL2_1-зростання середнього національного валового продукту на душу населення більше 2,1%;

CHANG-правлячий президент справив істотні зміни в політиці;

WAVE - під час правління були істотні соціальні хвилювання;

MIST-адміністрація правлячої партії винна в серйозній помилці або скандалі;

R_HERO – кандидат правлячої партії – національний герой;

O_HERO – кандидат опозиційної партії – національний герой.

Відповіді на питання представлені у вигляді 0 і 1. Одиниця означає ствердну відповідь на питання, нуль-негативний.

Вихідне поле позначається як I__, де результатом є значення змінної, відповідне кодом партії, що виграла (1 - правляча партія, 2-опозиційна партія).

Робота з нейронними мережами можлива тільки в рамках деякого нейропроекта. Для того щоб створити нейропроект, необхідно вибрати пункт меню "Файл/Создать або натиснути кнопку "Створити" на панелі кнопок, при цьому з'являється вікно наступного вигляду (рис.7.1).

Рисунок 7.1. - Вид вікна після вибору "Файл/Створити"

Після створення нейропроекта в нього можна вставляти нейронні мережі за допомогою кнопки "Відкрити файл даних і працювати з ними. Створений нейропроект також може бути збережений за допомогою команди меню "Файл/Сохранить", "Файл/Зберегти як ..." або натисненням на кнопку "Зберегти" на панелі інструментів.

Надалі можлива робота зі збереженими файлами нейропроекту. Для цього необхідно вибрати пункт меню "Файл-Відкрити" або натиснути кнопку "Відкрити" і вибрати в діалоговому вікні ім'я бажаного проекту.

Більшість операцій з нейронними мережами вимагають присутності підключеного до нейропроекту файлу даних.

Рисунок 7.2 - Вікно для створення нейронної мережі

Для підключення файлу даних або його заміни необхідно натиснути кнопку "Відкрити файл даних" у вікні нейропроекту і далі вибрати ім'я необхідного файлу даних. Відкритий файл відображається у власному вікні, де надається можливість його редагування. При приєднаному файлі даних можна проводити операції створення нових мереж, їх навчання, тестування і спрощення.

Для створення нової нейронної мережі необхідно натиснути кнопку "нова мережа" у вікні нейропроекту і заповнити вікно для створення неронної мережі (рис. 7.2).

Для нашого прикладу поле I _ _ _ необхідно позначити як вихідне, всі інші поля будуть вхідними. Після натискання кнопки "Створити" створюється нейронна мережа з наступними параметрами:

число вхідних полів: 12;
число входів мережі: 12;
число вихідних полів: 1;
число виходів мережі: 1;
шар 1: 10 нейронів;
шар 2: 10 нейронів;
шар 3: 10 нейронів.

Створену нейронну мережу можна далі навчати, тестувати, спрощувати і зберігати на диску разом з нейропроектом.

Для навчання активної в даний момент в нейропроекте нейронної мережі необхідно вибрати пункт меню "Нейромережа-Навчання". Якщо у файлі даних є всі необхідні поля і він не порожній, то запускається процес навчання мережі. При цьому на екран виводиться Вікно "навчання", де користувач має можливість спостерігати процес навчання і при необхідності самостійно завершити навчання натисканням кнопки "Завершити" (рис. 7.3).

Рисунок 7.3 – Відображення процесу навчання мережі

Навчання припиняється при досягненні нульового значення середньої оцінки на задачнику, у разі неможливості подальшого поліпшення оцінки або при аварійних ситуаціях (нульовий або нескінченний крок у напрямку оптимізації).

Маючи нейронну мережу можна подивитися, наскільки точно вона прогнозує значення вихідних полів у файлі даних. Для тестування нейронної мережі вибираємо пункт меню "Нейромережа/Тестування". Результат тестування мережі виводиться в "вікно тестування мережі" (рис. 7.4).

У вікні представлений результат прогнозу мережі по всіх роках, середня і максимальна помилки при прогнозуванні.

Рисунок 7.4 - Результат тестування мережі

Можливе тестування мережі на іншому файлі даних. Для цього необхідно спочатку підключити до проекту інший файл даних, а потім протестувати мережу. Результат тестування можна зберегти в текстовому файлі на диску. Далі цей файл можна обробляти в іншій програмі.

Не всі вхідні сигнали мережі і синапси нейронів необхідні для правильного вирішення мережею завдання. Часто можна досить сильно спростити мережу без погіршення точності рішення задачі.

При проведенні процесу спрощення мережі скорочується число вхідних сигналів мережі. У тих випадках, коли можна правильно вирішити завдання на основі меншого набору вхідних даних, це може надалі скоротити часові і матеріальні витрати на збір інформації.

Після спрощення нейронна мережа може придбати логічно прозору структуру і її можливо буде більш просто реалізувати на апаратній платформі.

Відомо, що майже неможливо зрозуміти, як навчена нейронна мережа вирішує завдання. Після спрощення нейронна мережа стає досить осяжній і можна спробувати побудувати алгоритм розв'язання задачі мережею на основі графічного подання або словесного опису структури мережі.

Для спрощення нейронної мережі є наступні операції в меню “Нейромережа”:

Скорочення числа вхідних сигналів – видалення найменш значущих вхідних сигналів (рис. 7.5)

Рисунок 7.5 - Зміна параметрів мережі при її спрощенні

При зменшенні кількості входів з 12 до 6 і зменшенні кількості проміжних зв'язків (синапсів) зменшується кількість правильно вирішених прикладів, що відповідають заданому рівню помилки, 30 замість 32.

Спрощення нейронної мережі можна проводити за різними параметрами.

Скорочення числа нейронів – видалення найменш значущих нейронів мережі.

Скорочення числа синапсів – видалення найменш значущих синапсів мережі.

Скорочення числа неоднорідних входів – видалення найменш значущих неоднорідних входів нейронів мережі.

Рівномірне спрощення мережі-скорочення числа приходять на нейрони мережі сигналів до кількості, що задається користувачем.

Бінаризація синапсів мережі – приведення значень ваг синапсів і неоднорідних входів нейронів до вибраних значень.

Після спрощення мережі навчена мережа мінімізується за кількістю вхідних параметрів і зв'язків. При використанні емулятора NeuroPro як системи для соціального прогнозу, що пророкує результат виборів в США виявилося, що для надійного передбачення результату виборів в США досить знати відповіді всього на п'ять питань, наведених нижче в порядку значимості:

1. Була серйозна конкуренція при висуненні від правлячої партії?
2. Під час правління були суттєві соціальні хвилювання?
3. Рік виборів був часом спаду або депресії?
4. Правлячий президент зробив значні зміни в політиці?
5. У рік виборів була активна третя партія?

Інші ознаки слабо пов'язані з підсумками виборів, тобто вони мало впливають на результат, і їх можна не використовувати при моделюванні. Система дозволяє побудувати багатофакторну модель і прибрати з неї фактори, що мало впливають на результат, тобто істотно спростити вихідну модель для даної задачі.

На прикладі представленої лабораторної роботи можна подивитися, як вирішується задача прогнозування за допомогою нейронних мереж. При вирішенні якого-небудь завдання прогнозування можна також спростити мережу без істотного погіршення точності вирішення завдання.

Порядок виконання роботи

1. Вивчити програмні характеристики і можливості емулятора NeuroPro 0.25;

2. Провести прогнозування нейронної мережі за допомогою емулятора NeuroPro 0.25, згідно з прикладом;

3. Провести навчання і тестування нейронної мережі, використовуючи файл даних Election.db;

4. Виконати зміну параметрів мережі при її спрощенні.

5. Провести аналіз і зробити висновки

Лабораторно – практична робота №8
Моделювання систем класифікації та прогнозу з використанням нейронних мереж

Мета роботи: оволодіти навичками моделювання систем класифікації та прогнозу з використанням нейронних мереж

Теоретична частина

Нейромережі оперують з інформацією, представленою тільки у вигляді чисел. Числа подаються на вхідні синапси нейромережі; відповіді, що знімаються з вихідних нейронів, також являють собою числа, тому для оцінки прикладу заздалегідь відомий відповідь також повинен бути представлений у вигляді числа (чисел). Інформація, на підставі якої нейромережа має давати відповідь, може бути са-мого різноманітного види: терміни, що описують які-небудь си-туації, числа різного виду і величини, динамічні ряди, масиви, графіки, динамічні криві, двох - і тривимірні зображення і т. д. Тому виникає необхідність коректного подання цієї інформації у вигляді чисел, що зберігають сенс і внутрішні взаємозв'язки в даних / 9/.

Існує величезна кількість способів представлення інформації для різних цілей / 10-12/.

Для роботи нейронних мереж з медико-біологічними дан-вими можна запропонувати наступну класифікацію даних, з якими може зіткнутися творець медичних експертних систем, і оптимальні способи їх подання у чисельному вигляді, що враховують специфіку роботи нейромережевих систем.

а) Число з "плаваючою точкою".

Один з найпоширеніших типів даних. Дані такого типу можуть приймати будь-які значення, дробові або цілі. Ча-ще всього вони позитивні, але можуть бути і негативними. Не менш важливо, що для роботи нейронних мереж практично не має значення, підпорядковується чи ні варіаційний ряд цих даних за-кону нормального розподілу. Як правило, дані такого виду розташовуються на будь-якому інтервалі з нечіткими кордонами.

Прикладом може послужити більшість даних лабораторних аналізів. Числову інформацію, приготовлену для нейромережевої обробки, бажано масштабувати, тобто вирівняти діапазони зміни величин, наприклад, обмеживши їх інтервалом [0,1] або [-1,1]. Зробити це можна за допомогою найпростішого лінійного пре-освіти:

$$\tilde{x}_n = \frac{x_n - x_{n\min}}{x_{n\max} - x_{n\min}}(b - a) + a,$$

де xn $\widetilde{x}_n$ – значення вихідного і масштабованого n-го параметра предметної області, що подається на n-й вхідний нейрон нейромережі; $[x_{nmin}, x_{nmax}]$ – реальний діапазон зміни n-го параметра; $[a, b]$ – прийнятний діапазон зміни вхідних сигналів.

Бажані вихідні сигнали персептрона повинні бути також закодовані в прийнятній формі і масштабовані в прийом-лемом діапазоні [a, b]. Це означає, що при формуванні навчаючого вектору D слід застосувати формулу масштабування:

$$\widetilde{d}_m = \frac{d_m - d_{m\min}}{d_{m\max} - d_{mn\min}}(b-a)+a$$

де dm $\widetilde{d}_m$ – задане та масштабоване значення m-ї компоненти вектора **D**. навчений на такій вибірці персептрон формуватиме вихідний вектор **Ỹ**, що містить значення, наведені до діапазону $[a, b]$. Тому до них має бути застосоване зворотне перетворення:

$$y_m = \frac{\widetilde{y}_m - a}{b-a}(d_{m\max} - d_{m\min}) + d_{m\min}.$$

Таким чином, персептрон можна застосовувати для моделювання предметної області, що описується числовими параметрами будь-якого діапазону. При формуванні навчальної вибірки вхідні і вихідні параметри бажано масштабувати-преобразувати до прийнятного діапазону [a, b]. Природно, що відповіді персептрона після цього слід інтерпретувати шляхом застосування перетворення, зворотного масштабуванню.

б) Взаємовиключні варіанти.

Один з найбільш складних типів даних, що вимагають проду - манного уявлення. Інформація при цьому представлена у вигляді одного і тільки одного варіанту із заздалегідь відомого і обмеженого набору варіантів і не може приймати вид дробового чис-ла. Найпростішим прикладів може служити стать людини-чоловіча або жіноча. Така інформація вимагає чисельного кодування. Можна закодувати чоловічу стать як 1, Жіночий-як 2 або на-оборот. Однак далеко не завжди можна зробити таке просте і довільне кодування. Для навчання нейромереж подібну інформацію логічно поділяти на 3 основних підтипу:

1) невпорядковані варіанти – наведений приклад з підлогою людини. Їх можна кодувати довільним способом. Час-то до цього типу відносяться дані, що представляються всього двома варіантами (так - ні, згоден - не згоден, хворів - не хворів і т. д.).

2) упорядковані варіанти. Такі дані знаходяться в оп-певних взаємозв'язках один з одним, які можуть бути вира-дружини відносинами типу "більше", "менше". Прикладом може служити ступінь тяжкості захворювання (I,II,III). У будь-якому випадку ва-ріанти розташовуються в певному порядку-як правило, зростання або убування.

3) частково впорядковані варіанти. Спосіб упорядкування не очевидний, проте його можна знайти, якщо це необхідно для вирішення завдання.

в) Сумісні варіанти.

Інформація може бути представлена одним або одночасним-але декількома варіантами з відомого і обмеженого набору варіантів. Прикладом може бути наявність у обстежуваного будь-яких захворювань або захворювань, перенесених в дитинстві. У таких випадках є два різних підходи до кодування даних:

1) якщо варіанти невпорядковані, найкращий (але, на жаль, не завжди зручний) спосіб полягає в тому, щоб розбити ознака на кілька ознак, кількість яких дорівнює кількості якості варіантів і кожен з них кодувати окремо. Кожен під ознаку в такому випадку робиться самостійною ознакою і подається на окремий вхідний синапс нейромережі.

2) якщо варіанти впорядковані, можна застосувати принцип бітової маски.

г) Дата, час.

Дуже часто медичні дані містять дати різних подій в житті обстежуваних. Для чисельного представлення тимчасових точок необхідно в кожній ознаці вибирати відповідну точку відліку і, відштовхуючись від неї, виражати часовий інтервал в зручних одиницях (секунди, години, добу, роки). Наприклад, якщо у хворого вказана дата виникнення будь-якого захворювання, зручніше перевести її в вік, який відповідав цій події. Зробивши це для всіх обстежуваних, фахівець призведе показники до єдиної шкалою. Якщо потрібен інший підхід, можна порахувати, скільки часу пройшло з моменту виникнення захворювання до моменту даного обстеження. У кожному випадку предметний фахівець повинен вибрати спосіб подання, найбільш добре відображає сенс параметра.

Програма NeuroPro 0.25 є вільно поширюваною бета-версією програмного продукту для роботи з штучними нейронними мережами і виробництва знань з таблиць даних з по-міццю нейронних мереж.

Можливості програми:

1. Читання, запис, редагування, конвертація файлів даних, представлених у форматах dbf (СУБД dBase, FoxPro, Clipper) і db (СУБД Paradox).

2. Створення шаруватих нейронних мереж (персептронів) для вирішення завдань прогнозування та класифікації:

– число шарів нейронів – до 10;

- число нейронів в шарі-до 100. Число нейронів в шарі може задаватися окремо для кожного шару нейронів;

– нейрони – сигмоидные з нелінійною функцією f(A)=A/(|A|+c), крутизна сигмоиды може задаватися окремо для кожного шару нейронів;

– робота з кількісними (безперервними) і якісними (дискретнозначными, від 2 до 20 дискретних станів для ознаки) вхідними ознаками;

– - рішення задач прогнозування (передбачення значень кількісних вихідних ознак) і класифікації (передбачення станів якісних вихідних ознак);

– нейромережа може мати кілька вихідних сигналів (вирішувати одночасно кілька завдань прогнозування та класифікації); для кожного з вихідних сигналів можуть бути встановлені свої вимоги до точності рішення задачі.

3. Навчання нейронної мережі із застосуванням одного з таких методів градієнтної оптимізації (градієнт обчислюється за принципом подвійного функціонування):

– градієнтний спуск.

– модифікований ParTan-метод.

- метод пов'язаних градієнтів.

– квазиньютоновский BFGS-метод.

4. Тестування нейронної мережі, отримання статистичної інформації про точність розв'язання задачі.

5. Обчислення та відображення значимості вхідних сигналів мережі, збереження значень показників значущості у файлі на диску.

6. Внесення випадкових збурень у ваги синапсів мережі.

7. Спрощення (контрастування) нейронної мережі:

– скорочення числа вхідних сигналів мережі;

– скорочення числа нейронів мережі;

- рівномірне проріджування структури синапсів мережі;

– скорочення числа синапсів мережі;

- скорочення числа неоднорідних входів (порогів) нейронів мережі;

– бінаризація ваг синапсів мережі (приведення ваг синапсів і порогових входів до кінцевого набору виділених значень). Можливий вибір з 4-х наборів виділених значень.

8. Генерація вербального опису нейронної мережі. Вербальне опис може редагуватися і зберігатися у файлі на диску.

Порядок роботи з програмою **NeuroPro 0.25**

1 До початку роботи необхідно створити файл даних, на основі якого буде сформована структура входів нейронної мережі, і на якому буде вироблятися її навчання. Файл у форматі **.dbf** повинен містити стовпці вихідних параметрів і стовпець з результатами роботи нейромережі.

2 Після запуску програми потрібно створити новий файл НС **(Файл - Создать),** потім натиснути на кнопку Відкрити файл даних.

У вікні потрібно вказати шлях до файлу вихідних даних. Після цього стають доступними опції роботи з нейронною мережею.

3 Потрібно натиснути на кнопку Нова мережа, і у вікні Входи і виходи вказати, які з полів файлу вихідних даних є вхідними, які вихідними, і який тип представлених даних (кількісний (безперервний) або якісний (дискретний)).

4 На вкладці Структура мережі потрібно вибрати число шарів (прихованих шарів персептрона без обліку вхідного і вихідного) і число нейронів у кожному шарі.

5 Провести навчання нейромережі на файлі вихідних даних: **Нейромережа – Навчання.**

6 Після навчання можна:
– переглянути отримані синаптичні ваги: **Нейромережа – Вербалізація;**
– переглянути діаграму значимості вхідних сигналів нейромережі;
– спростити мережу в автоматичному режимі шляхом скорочення числа найменш значущих вхідних сигналів, нейронів або синапсів.

Для тестування навченої нейронної мережі на новій вибірці даних потрібно завантажити новий файл даних **(Відкрити файл даних)** і провести тестування **(Нейромережа – Тестування).** Результати будуть представлені у вигляді таблиці із зазначенням результа-тів класифікації і похибок.

Питання для самоперевірки

1) класи завдань, що вирішуються за допомогою нейронних мереж.
2) Структура формального нейрона.
3) Види активаційних функцій.
4) Властивості нейронних мереж.
5) Види структур нейронних мереж.
6) Структура і властивості багатошарового персептрона.
7) Підготовка вихідних даних для нейромережевої обробки.

Порядок виконання роботи

Створення і навчання системи класифікації цифр на ч з'єднані з ними і безповоротні.

Потрібно створити нейронну мережу, вхідними даними для якої є стан кожного з 9 елементів індикатора, ото-бражающего цифри від 0 до 9 (рис. 15).

Входи нейромережі для кожної цифри кодуються набором з 9 двійкових значень: 0, якщо відповідний елемент індикатора «не горить», і 1 для активного елемента. Подання цифр на індикаторі показано на рис. 8.2. Виходом нейромережі також є двійкове число: 0 для парних і 1 для непарних чисел.

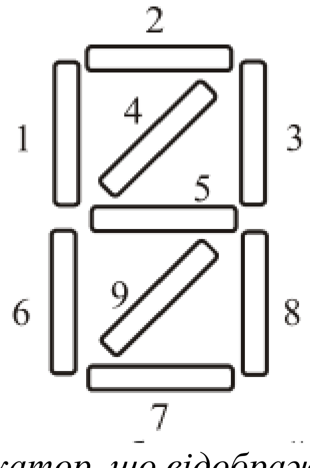


Рисунок 8.1 – Індикатор, що відображає цифри від 0 до 9

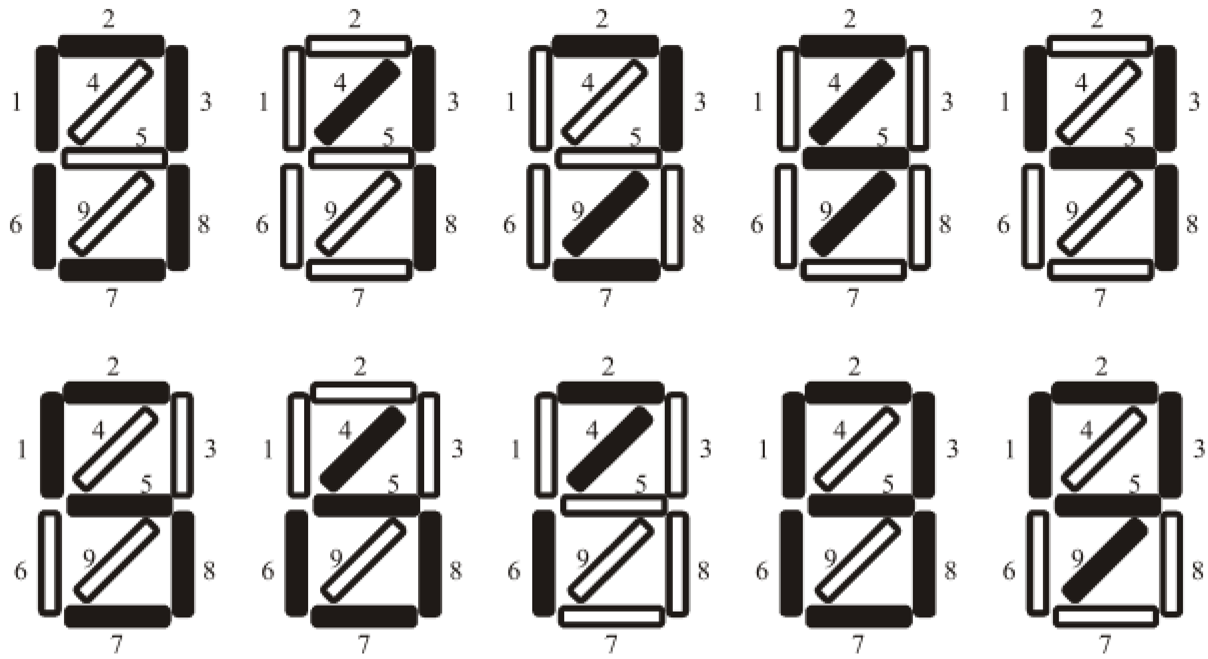


Рисунок 8.2 – Представлення цифр на індикаторі

1 Попередня підготовка даних.

Нейромережа здатна сприймати тільки числові значення. З допомогою програми MS EXCEL створіть навчальний файл: показання кожного елемента індикатора і значення парності числа для всіх десяти цифр (рис. 17).

Збережіть дані у форматі баз даних **.dbf** (***Файл*** , попередньо виділивши область збережених значень).

2 Запустіть NeuroPro 0.25, створіть новий файл НС, завантажте файл даних (кнопка).

3 представлене завдання класифікації може бути вирішена найпростішою НС. Створіть одношаровий персептрон з одно-го нейрона (число шарів – 1, число нейронів - 1). Входи і вихід задайте дискретними.

Рисунок 8.3 – Зовнішній вигляд навчального файлу

4 Навчіть нейронну мережу, перегляньте е ст. вербальний опис, результати тестування. Перегляньте діаграму значи-мости вхідних сигналів, що показує, якою мірою значення кожного з індикаторів враховується при прийнятті рішення.

5 Зробіть автоматичне скорочення числа вхідних сигналів. Перегляньте діаграму значущості.

Створення і навчання системи розпізнавання букв алфавіту.

Потрібно розробити Класифікатор, що пізнає букви російського алфавіту. Число виходів нейронної мережі має бути одно числу розпізнаваних символів, при появі певної літери на відповідному виході ІНС повинна з'явитися «1», на всіх ос-талевих виходах – «0».

1 Запропонуйте систему кодування букв російського алфавіту (за принципом, схожим з побудовою 9-елементного індикатора для цифр).

2 виберіть 5 букв з вашого прізвища і за допомогою про-грами MS EXCEL створіть навчальний файл.

3 запустіть NeuroPro 0.25, створіть новий файл НС, завантажте файл даних. Виберіть і змоделюйте нейромережеву модель, достатню для вирішення даної задачі класифікації. Навчіть і протестуйте отриману модель.

Створення нейронної мережі для прогнозування розвитку захворювання Потрібно створити нейронну мережу для прогнозування течії вперше виникла стенокардії. Вхідними даними є набір клініко-інструментальних факторів, що характеризують пацієнта на момент виникнення нападу. Дані закодовані в число-ву форму, придатну для обробки нейронною мережею:

1. Вік у роках (поле AGE)
2. Куріння (поле SMOKE)

Характеристика	Код НС
Не курити	0
Курив, але кинув; курить <10 сиг/день	1
курить 10-30 років, 10-20 сиг/день	2
Курить>30 курить	3

3. Артеріальна гіпертонія (поле HYPERTONY)

Характеристика	код НС
Ні	0
ГБ 1 ст.	1
ГБ 2 ст.	2
ГБ 3 ст.	3

4. Верхня межа пекло (поле AD_TOP)
5. Нижня межа пекло (поле AD_BOTTOM)
6. Рівень холестерину (поле HOLESTERIN)
7. Професія (поле WORK)

Професія	код НС
Помірна праця	1
Важка фізична праця	2
Керівна посада	3
Велика керівна посада	4

8. Спадковість (наявність серцевих захворювань) (поле NASLED)

Характеристика	код НС
Ніт	1
Батько або мати	2
Мати і батько	3
Мати, батько та ін. родичі	4

9. Стан при виникненні першого нападу (поле VOZNIKN)

Характеристика	код НС
Спокій	4
Незначні фізичні навантаження	3
Помірні фізичні навантаження	2
Важкі навантаження	1

10. Збереження нападів (поле SOHRAN)

Характеристика	код НС
ні	1
1-3 дня	2
3-5 днів	3
>5 днів	4

11. Збільшення лівого шлуночка (поле UVELICH_LG): (немає -0, є – 1)

12. Ефект нітрогліцерину (поле NITRO)

Характеристика	Код НДС
Не застосовувався 0	0
Хороший ефект 1	1
Без ефекту 2	2

13. Вживання алкоголю (поле ALCO)

Характеристика	код НС
Ні	1
Побутове вживання	2
Зловживання	3

Виходом нейронної мережі є прогноз, відповідно до якого пацієнт відноситься до одного з трьох класів:

Клас 1: одужання, л гостра форма стенокардії

Клас 2: середня тяжкість стенокардії

Клас 3: можливе виникнення інфаркту міокарда або летальний результат

Вибірка даних про пацієнтів, на якій повинно проводитися навчання, представлена у файлі **Навчальний.dbf.** У ньому задані 13 стовпців вхідних параметрів і 3 стовпець вихідних ознак, з-відних трьома класами прогнозу.

1 Створіть новий файл в NeuroPro 0.25. Відкрийте файл даних **Навчальна.dbf.** Створіть нову мережу.

2 На вкладці Входи і виходи визначте поля *Class 1, Class 2 і Class 3* вихідні – це стовпці, відповідні ймовірності кожного класу прогнозу. На вкладці Структура мережі за-дайте двошаровий персептрон з 13 нейронами в кожному шарі.
3 Навчіть мережу. Подивіться діаграму значущості вхідних даних.
4 протестуйте мережу на вибірці пацієнтів, записи яких не входили в навчальну вибірку. Для цього відкрийте файл **Тестова.dbf** і запустіть процедуру тестування. Подивіться ре-зультат.
5 зробіть скорочення надлишкового числа синапсів і нейронів мережі.
6 В тому ж файлі на основі вибірки **Навчальна.dbf** соз-дайте одношарову мережу з 20 нейронів, навчіть і протестуйте.

Звіт з практичного заняття повинен містити відповіді на питання для самоперевірки, лістинг робочого докумен-ту NeuroPro з отриманими результатами моделювання і таблиця-ми.

СПИСОК ЛІТЕРАТУРИ

Навчальне видання

ОСНОВИ СУЧАСНИХ ТЕОРІЙ МОДЕЛЮВАННЯ ПРОЦЕСІВ

Методичні вказівки до лабораторно – практичних робіт

для студентів спеціальності 131 «Прикладна механіка» всіх форм навчання

Укладачі: С.В. Ковалевський, проф. д.т.н.
В.В. Ємець, аспірант (PhD)

Редагування

Комп'ютерне верстання

10/2012. Формат 60 х 84/16. Ум. друк. арк. 3,49.
Обл.-вид. арк. 2,37. Тираж пр. Зам. №

Видавець і виготівник
Донбаська державна машинобудівна академія
84313, м. Краматорськ, вул. Академічна, 72.
Свідоцтво суб'єкта видавничої справи
ДК №1633 від 24.12.2003