

Пістунов І.М. **Нейронні сітки та їх застосування** навч. наоч. посіб. Дніпро : НТУ «ДП», 2024. 28 с.

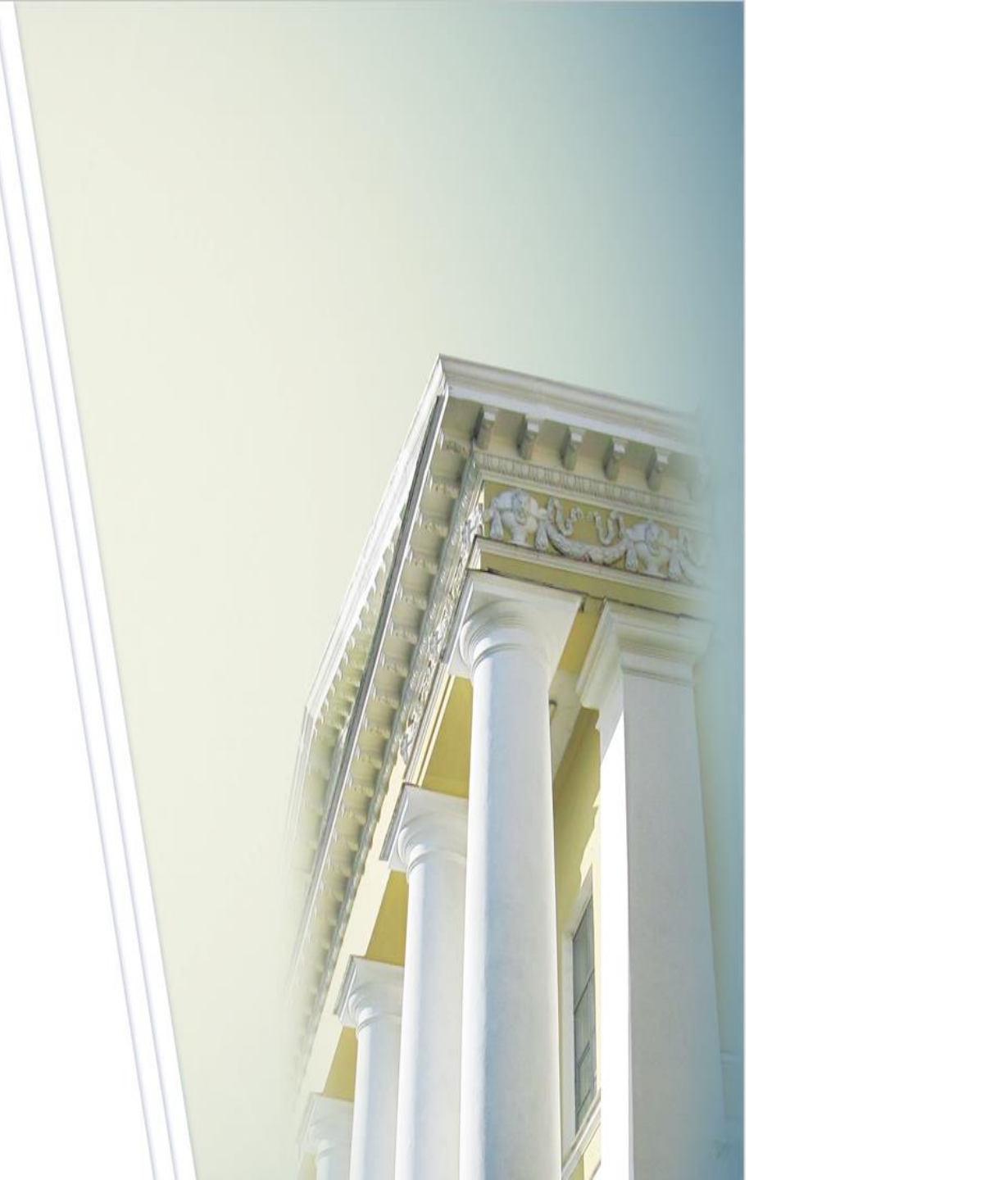
В посібнику розглядаються практичні питання стосовно використання програмного комплексу Statistica щодо побудови регресійних моделей, прогнозування часових рядів кластеризації об'єктів та їх класифікації.

Призначено для студентів спеціальності 051 «Економіка», 071 «Облік і оподаткування», 072 «Фінанси, банківська справа та страхування», 073 «Менеджмент», 292 «Міжнародні економічні відносини»

Рецензенти:

Васильєва Н.К., завідувач каф. інформаційних систем ДДАЕУ, проф.

Алексєєв М.О., зав каф. програмних засобів комп'ютерних систем НТУ «ДП», проф.



ЗМІСТ

ВСТУП	3
1. ОСНОВНІ ПОНЯТТЯ	4
2. МЕТОД ЗВОРОТНОГО ПОШИРЕННЯ ПОХИБКИ	6
3. МОДЕЛЮВАННЯ ЕКОНОМІЧНИХ ПОКАЗНИКІВ ІЗ ЗАСТОСУВАННЯМ ПРОГРАМИ STATISTICA	15
4. ПРОГНОЗУВАННЯ ЕКОНОМІЧНИХ ПОКАЗНИКІВ	18
5. КЛАСТЕРИЗАЦІЯ ЕКОНОМІЧНИХ ОБ'ЄКТІВ	19
6. КЛАСИФІКАЦІЯ ОБ'ЄКТІВ	27



ВСТУП

Нейронні сітки – схеми з'єднання однорідних елементів – нейронів. Ця назва стосується не тільки біологічних об'єктів, але й їх математичних моделей. Ці моделі можуть розв'язувати задачі: логічні, стохастичні, керування, моделювання, прогнозування. Таку характеристику на с. 147 – 148 другого тому дає перша у світі Енциклопедія кібернетики, видана у Києві в 1973 році.

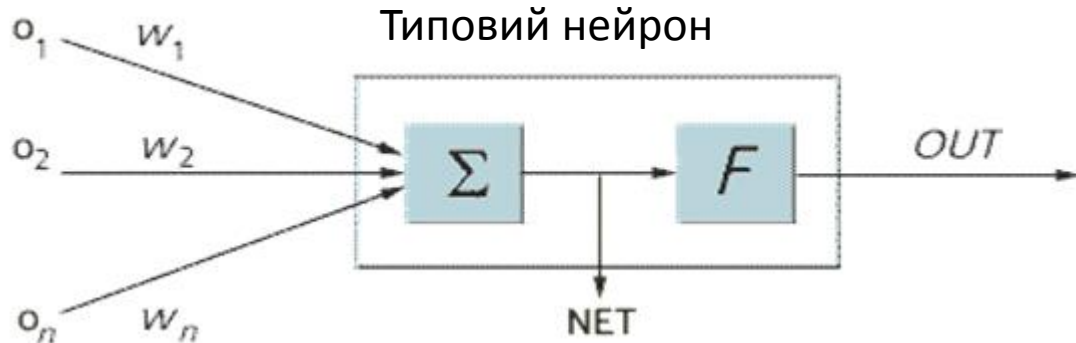
В наш час цей термін застосовується для позначення структури математичної моделі, яка, в свою чергу називається **нейронною мережею**.

До завдань, ефективно розв'язуються за допомогою ІС ШІ, в даний час відносяться наступні:

- апроксимація функцій (побудова моделей для залежностей, що погано формалізуються), наприклад, в економіці оцінка вартості нерухомості, контроль якості випущеної продукції;
- класифікація образів (визначення приналежності вхідного образу попередньо визначеним класам), наприклад, оцінки кредитоспроможності, оцінка ймовірності банкрутства;
- кластеризація (класифікація образів «без вчителя»), наприклад, аналіз діяльності конкуруючих фірм, стиснення інформації в інформаційних сховищах;
- прогнозування, наприклад, економічних параметрів і фондових індексів, обсягів продажу;
- асоціативна пам'ять (пам'ять, що адресується за змістом), наприклад, для створення сховищ даних, мультимедійних баз даних;
- оптимізація (знаходження рішення, що задовольняє системі обмежень і максимізує або мінімізує цільову функцію), наприклад, прибутковості роботи компанії, вирішує задачі ціноутворення;
- управління, наприклад, товарними запасами, динамічними об'єктами.



ОСНОВНІ ПОНЯТТЯ

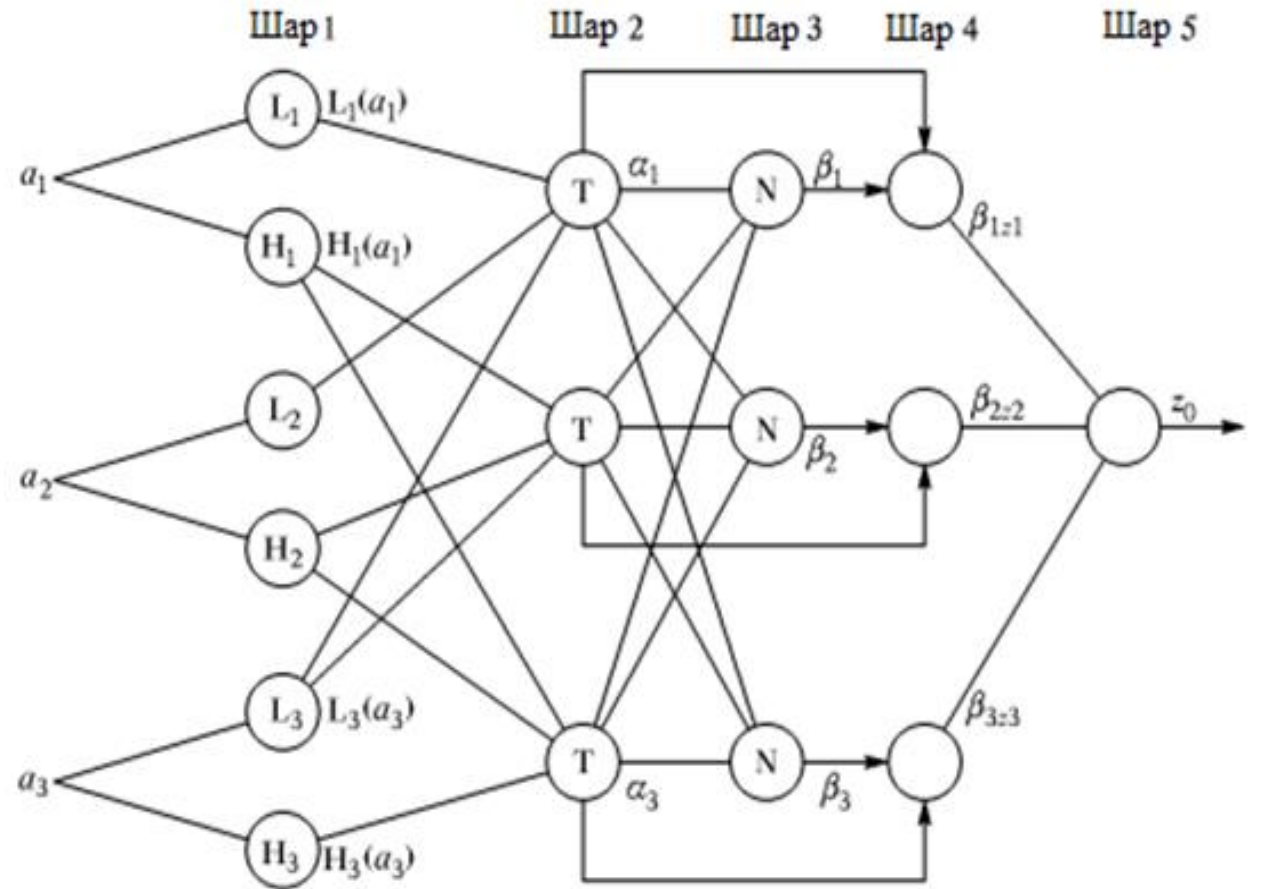
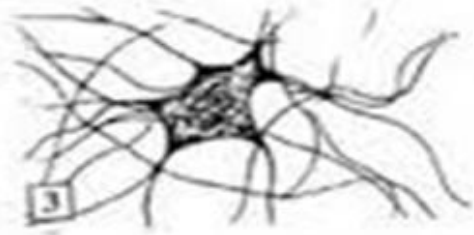


Експоненційна (радіальна)	$f(n) = e^{-n^2}$	
Трикутна	$f(n) = \begin{cases} 0, & n < -1 \\ 1 - n , & -1 \leq n \leq 1 \\ 0, & n > 1 \end{cases}$	
Логістична	$f(n) = \frac{1}{1 + e^{-n}}$	
Гіперболічна (тангенціальна)	$f(n) = \frac{2}{1 + e^{-2n}} - 1$	

Приклади функцій активації

Назва функції	Аналітичний вигляд	Геометричне зображення
Лінійна	$f(n) = n$	
Позитивна лінійна (напівлінійна)	$f(n) = \begin{cases} 0, & n < 0 \\ n, & n \geq 0 \end{cases}$	
Лінійна з обмеженням II	$f(n) = \begin{cases} 0, & n < 0 \\ n, & 0 \leq n \leq 1 \\ 1, & n > 1 \end{cases}$	
Симетрична лінійна з обмеженням II	$f(n) = \begin{cases} -1, & n < -1 \\ n, & -1 \leq n \leq 1 \\ 1, & n > 1 \end{cases}$	

МАШИННА АНАЛОГІЯ ЖИВІЙ НЕЙРОННІЙ МЕРЕЖІ



МЕТОД ЗВОРОТНОГО ПОШИРЕННЯ ПОХИБКИ

Метод зворотного поширення помилки ([англ. *backpropagation*](#)) — метод навчання [багатошарового перцептрону](#). Це ітеративний [градієнтний алгоритм](#), який використовується з метою [мінімізації](#) помилки роботи [багатошарового перцептрону](#) та отримання бажаного виходу. Основна ідея цього методу полягає в поширенні [сигналів](#) помилки від виходів мережі до її входів, в напрямку, зворотному прямому поширенню сигналів у звичайному режимі роботи. Барц і Охонін запропонували відразу загальний метод («принцип подвійності»), який можна застосувати до ширшого класу систем, включаючи [системи з запізненням](#), [розподілені системи](#), тощо ^[1]. Для можливості застосування методу зворотного поширення помилки [функція активації нейронів](#) повинна бути [диференційовною](#).

Алгоритм зворотного поширення помилки застосовується для [багатошарового перцептрону](#). У мережі є множина входів x_1, \dots, x_n , множина виходів *Outputs* і безліч внутрішніх вузлів. Перенумеруємо всі вузли (включаючи входи і виходи) числами від 1 до *N* (наскрізна нумерація, незалежно від топології шарів). Позначимо через $w_{i,j}$ вагу зв'язку, що з'єднує *i*-й і *j*-й вузли, а через o_i — вихід *i*-го вузла. Якщо нам відомий навчальний приклад (правильні відповіді мережі t_k , $k \in \text{Outputs}$), то функція помилки, отримана за [методом найменших квадратів](#), виглядає так:

$$E(\{w_{i,j}\}) = \frac{1}{2} \sum_{k \in \text{Outputs}} (t_k - o_k)^2.$$

Як модифікувати ваги? Ми будемо реалізовувати [стохастичний градієнтний спуск](#), тобто будемо підправляти ваги після кожного навчального прикладу *i*, таким чином, «рухатися» в багатовимірному просторі ваг. Щоб «добратися» до мінімуму помилки, нам потрібно «рухатися»



в сторону, протилежну **градієнту**, тобто, на підставі кожної групи правильних відповідей, додавати до кожної ваги $w_{i,j}$

$$\Delta w_{i,j} = -\eta \frac{\partial E}{\partial w_{i,j}},$$

де $0 < \eta < 1$ — множник, що задає швидкість «руху».

Похідна розраховується таким чином. Нехай спочатку $j \in \text{Outputs}$, тобто вага, яка нас цікавить, входить в нейрон останнього рівня. Спочатку зазначимо, що $w_{i,j}$ впливає на вихід мережі лише як частина суми $S_j = \sum_i w_{i,j} x_i$, де сума

береться по входах j -го вузла. Тому

$$\frac{\partial E}{\partial w_{i,j}} = \frac{\partial E}{\partial S_j} \frac{\partial S_j}{\partial w_{i,j}} = x_i \frac{\partial E}{\partial S_j}$$

Аналогічно, S_j впливає на загальну помилку тільки в рамках виходу j -го вузла o_j (нагадуємо, що це вихід всієї мережі).

Якщо ж j -й вузол — не на останньому рівні, то у нього є виходи; позначимо їх через $\text{Children}(j)$. У цьому випадку

$$\frac{\partial E}{\partial S_j} = \sum_{k \in \text{Children}(j)} \frac{\partial E}{\partial S_k} \frac{\partial S_k}{\partial S_j},$$

і

$$\frac{\partial S_k}{\partial S_j} = \frac{\partial S_k}{\partial o_j} \frac{\partial o_j}{\partial S_j} = w_{i,j} \frac{\partial o_j}{\partial S_j} = w_{i,j} o_j (1 - o_j).$$

А $\frac{\partial E}{\partial S_k}$ — це аналогічна поправка, але обчислена для вузла

наступного рівня (будемо позначати її через δ_k — від Δ_k вона відрізняється відсутністю множника $(-\eta x_{i,j})$). Оскільки ми навчилися обчислювати поправку для вузлів останнього рівня і виразити поправку для вузла нижчого рівня через поправки більш високого, можна вже створювати алгоритм навчання. Саме через цю особливість обчислення поправок цей алгоритм називається **алгоритмом зворотного поширення помилки** (англ. *backpropagation*).



Коротке викладення вищесказаного:

- Для вузла останнього рівня

$$\delta_j = -o_j(1 - o_j)(t_j - o_j)$$

- Для внутрішнього вузла мережі

$$\delta_j = -o_j(1 - o_j) \sum_{k \in \text{Outputs}(j)} \delta_k w_{j,k}$$

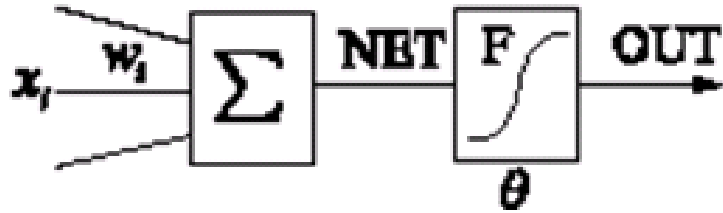
- Для всіх вузлів

$$\Delta w_{i,j} = -\eta \delta_j x_i$$

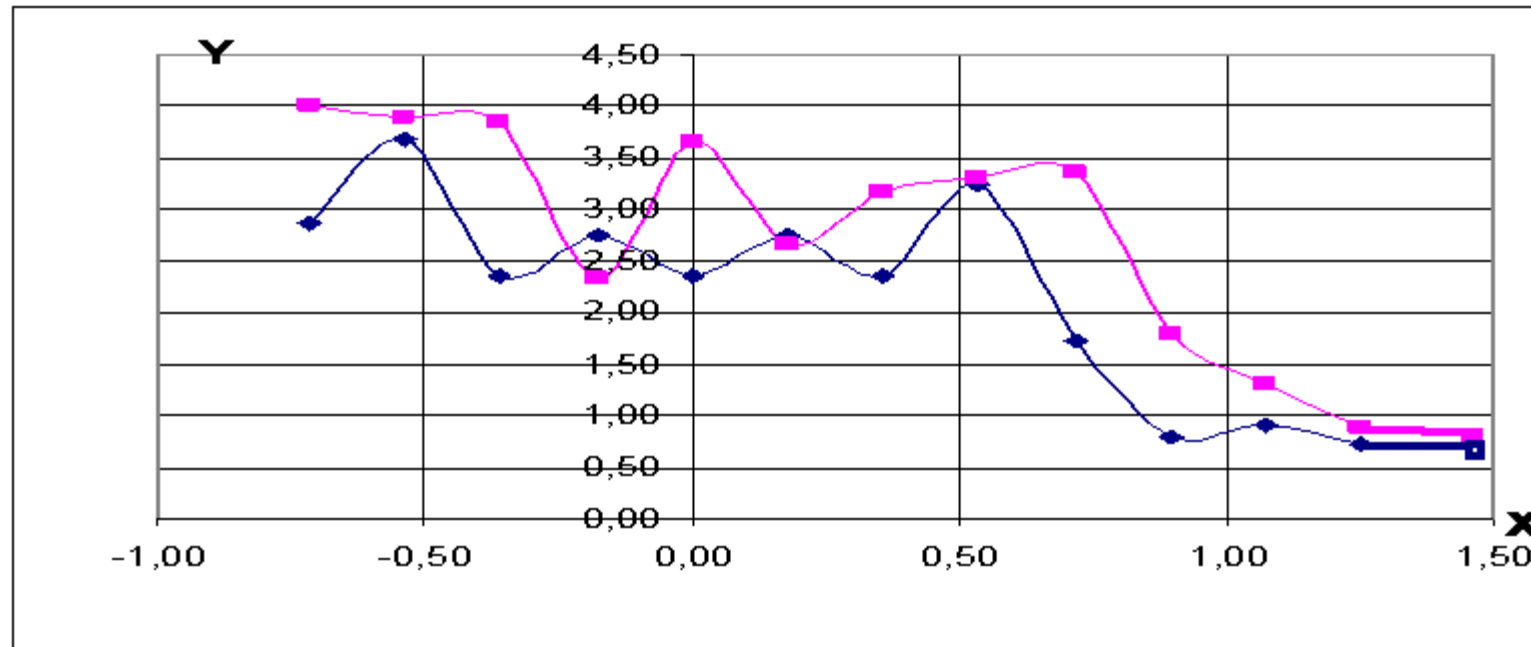


Приклади нейронних сіток

Загальний вигляд схеми перцептрона з одним нейроном та суматором на вході

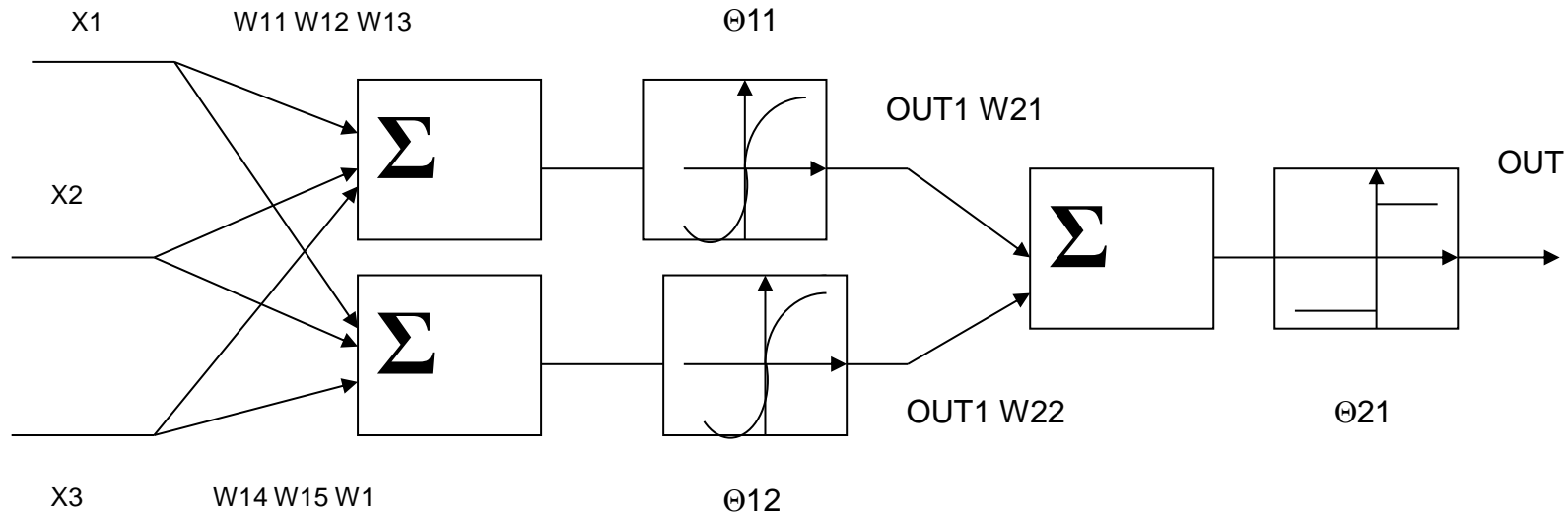


$$OUT = \frac{1}{1 + e^{-(w_1 x_i + w_2 x_{i+1} + w_3 x_{i+2})}}$$



Графік кількості викликів Y по годинам робочого дня X.
(♦ – експериментальна крива,
■ – розрахована крива)

Схема двошарового перцептрона з трьома входами на кожному нейроні



$$OUT = \text{Sign} \left[\frac{10,37669298}{1 + e^{(-1,443525248X_1 + 1,561436345X_2 + 1,582503396X_3 + 4,683681367)}} + \frac{7,940045065}{1 + e^{(-1,556474752X_1 + 4,561436345X_2 + 4,582503396X_3 + -3,683681367)}} + 2,883681367 \right]$$

Модель прогнозування інвестицій на вугільній шахті О.Ю. Чуріканової

$$OUT = \text{sigm}(NET) = \frac{1}{1 + \exp(\lambda_3 \left\{ \frac{1}{1 + \exp(\lambda_1 (\sum_{i=1}^{26} \sum_{j=1}^3 x_{ij}^1 w_{ij}^1))} + \frac{1}{1 + \exp(\lambda_2 (\sum_{i=1}^3 \sum_{j=1}^1 x_{ij}^2 w_{ij}^2))} \right\})}$$

де i – номер входу, j – номер нейрона у шарі, l – номер шару,
для $l=1: j=1,3; \underline{i}=1,26: x_1^1$ - інвестиції в підймальні установки; x_2^1 - інвестиції в
вентиляторні установки; x_3^1 - інвестиції в компресорні установки; x_4^1 -
інвестиції в водовідливні установки; x_5^1 - інвестиції в транспортні установки;
 x_6^1 - інвестиції в дегазацію; x_7^1 - інвестиції в теплоенергетичні установки; x_8^1 -
інвестиції в технічний комплекс поверхні шахти; x_9^1 - інвестиції в
електротехнічні установки; x_{10}^1 - інвестиції в споруди; x_{11}^1 - інвестиції в запаси
матеріалів; x_{12}^1 - інвестиції в акції; x_{13}^1 - інвестиції в цінні папери; x_{14}^1 - інвестиції
в депозити; x_{15}^1 - інвестиції в підготовку кадрів; x_{16}^1 - інвестиції в дослідження
та розробки; x_{17}^1 - інвестиції в рекламу; x_{18}^1 - середня зольність вугілля; x_{19}^1 -
середня вологість вугілля; x_{20}^1 - середня сірчистість вугілля; x_{21}^1 - залишкові
промислові запаси млн.т.; x_{22}^1 - середня геологічна потужність пластів м.; x_{23}^1 -
максимальна глибина розробки; x_{24}^1 - протяжність гірничих виробок; x_{25}^1 - кут
падіння пластів; x_{26}^1 - потужність (проектна) млн. тон на рік.

для $l=2: j=1,3; \underline{i}=1: x_1^2$ - прибуток. x_{ij}^l – i -ий вхідний сигнал j -го нейрона в шарі
 l , w_{ij}^l – ваговий коефіцієнт i -го входу нейрона номер j у шарі l NET_{lj} – сигнал
 NET j -го нейрона в шарі l , OUT_{li} – вхідний сигнал нейрона, $\lambda > 0$ – коефіцієнт
крутості безупинної функції $f(net)$ біля $net=0.5$.

Таблиця 1

Значення вагових коефіцієнтів w_{ij}^l першого шару

Номер нейрону	Номер входу													
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1,7	5,1	1,9	4,2	4,2	2,6	3,1	0,9	0,2	2,5	5,2	1,8	1,3	
2	5,2	1,1	5,9	2,9	6,2	-0,1	1,1	6,6	5,2	0,1	2,7	3,3	3,8	
3	-0,5	1,0	6,8	3,7	3,4	3,4	0,6	0,9	5,0	3,6	3,1	3,6	1,5	
Номер нейрону	Номер входу													
	14	15	16	17	18	19	20	21	22	23	24	25	26	
1	2,6	2,4	0,6	2,2	0,6	3,4	3,8	1,7	2,6	4,0	3,8	1,6	0,0	
2	1,9	0,2	0,1	0,2	5,5	0,4	2,4	0,7	1,1	0,9	0,1	2,9	1,1	
3	2,9	4,1	2,7	6,3	4,2	5,8	6,2	3,5	1,0	0,0	1,0	3,7	3,1	

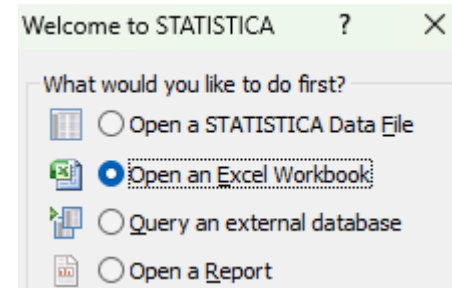
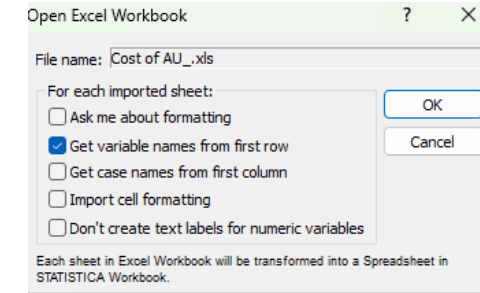
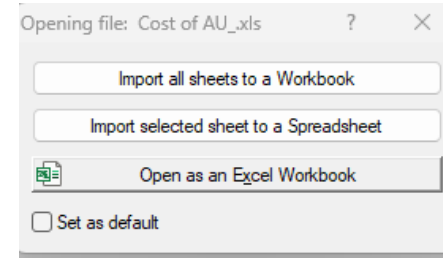
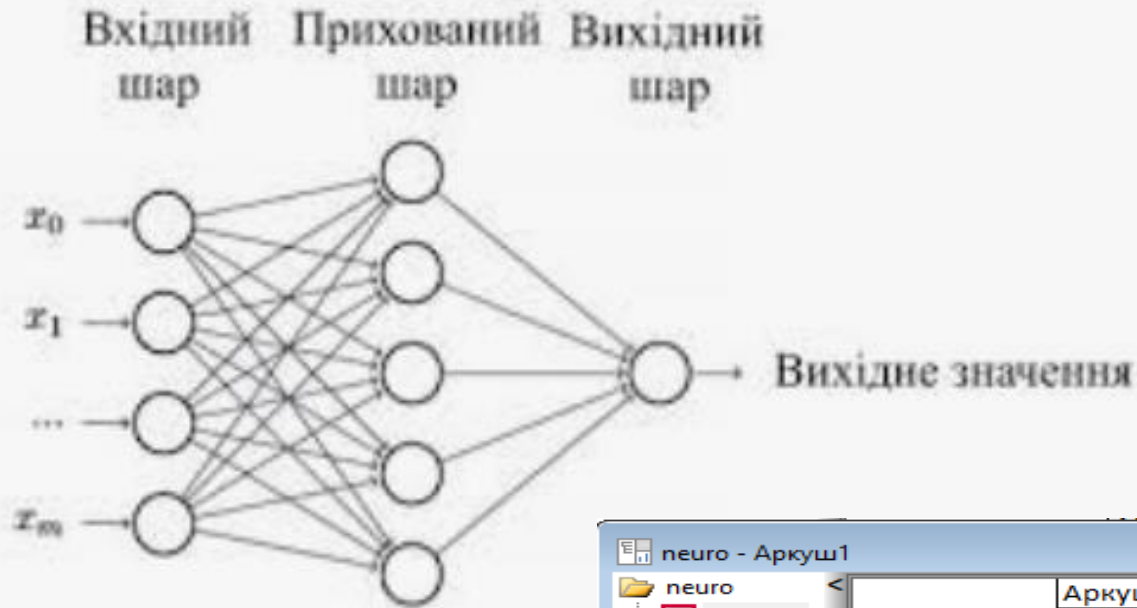
Таблиця 2

Значення вагових коефіцієнтів w_{ij}^l другого шару

Номер нейрону	Номер виходу
	1
1	1,1
2	0,1
3	4,2

МОДЕЛЮВАННЯ ЕКОНОМІЧНИХ ПОКАЗНИКІВ ІЗ ЗАСТОСУВАННЯМ ПРОГРАМИ STATISTICA

Типова структура нейронної сітки
у програмі STATISTICA



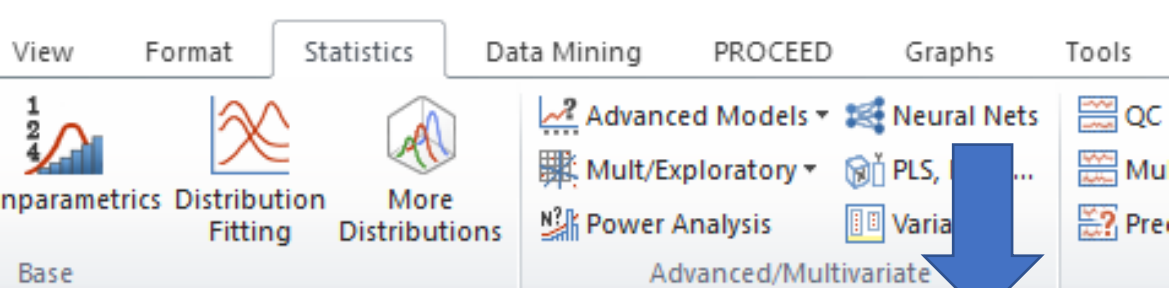
neuro - Аркуш1

neuro

Аркуш1

	X1	X2	X3	X4	X5	X6	X7	Y
1	55	86	93	89	59	52	53	68
2	74	55	86	93	89	59	52	53
3	98	74	55	86	93	89	59	52
4	47	98	74	55	86	93	89	59
5	90	47	98	74	55	86	93	89
6	52	90	47	98	74	55	86	93
7	63	52	90	47	98	74	55	86
8	91	63	52	90	47	98	74	55





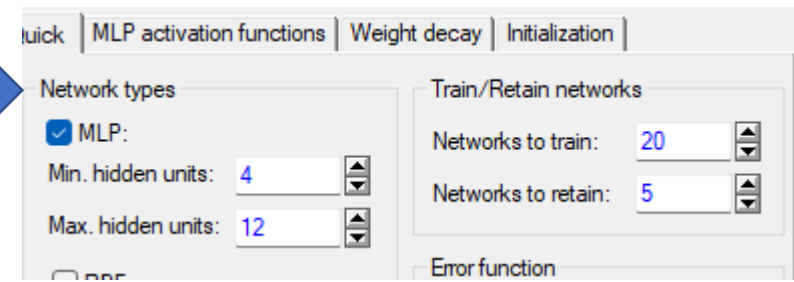
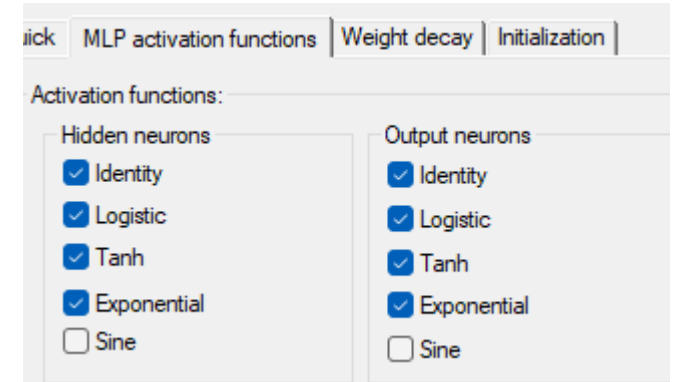
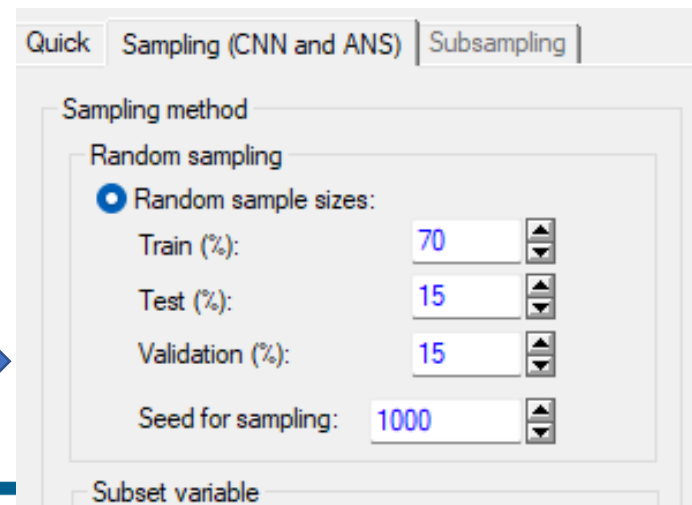
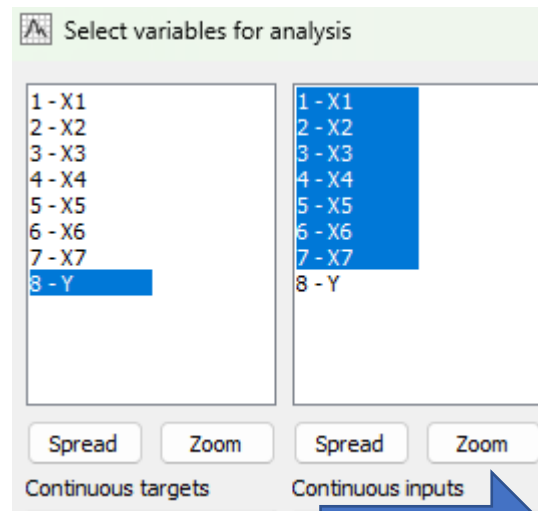
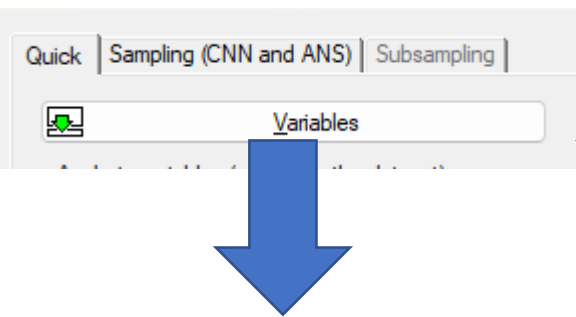
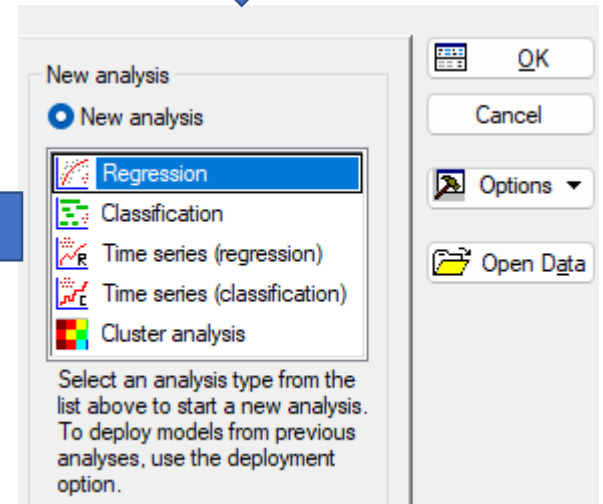
SANN - Results: Apkyw1 in neuro

Active neural networks

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct
1	MLP 7-6-1	0,533258	0,859875	0,473241	BFGS 2	SOS
2	MLP 7-7-1	0,548197	0,852883	0,390742	BFGS 2	SOS
3	MLP 7-11-1	0,491951	0,742211	0,573982	BFGS 4	SOS
4	MLP 7-10-1	0,985201	0,560917	0,556640	BFGS 0	SOS

Select/Deselect active networks Delete networks

Build models with CNN Build models with ANS Build models with Subsampling



Build models with CNN Build models with ANS Build models with Subsampling

Predictions | Graphs | Details | Custom predictions

Predictions spreadsheet

Predictions type

- Standalones
- Ensemble
- Standalones and ensemble

Include

- Inputs
- Targets
- Output
- Residuals
- Confidence intervals
- Absolute res.
- Square res.
- Standard res.
- Variables

Predictions

Predictions spreadsheet for Y (Аркыш1 in neuro)
Samples: Train

Case name	Y Target	Y - Output 1. MLP 7-6-1	Y - Output 2. MLP 7-7-1	Y - Output 3. MLP 7-11-1	Y - Output 4. MLP 7-10-1	Y - Output 5. MLP 7-6-1
1	68,00000	74,19032	79,68619	76,31714	66,92150	75,99683
2	53,00000	68,17614	64,26273	64,25772	47,00000	64,60367
4	59,00000	72,16313	73,43006	76,07604	55,28225	73,60958
5	89,00000	70,49875	70,53008	74,83130	89,84295	72,70037
7	86,00000	71,89489	73,15656	73,16711	90,99737	73,48834
8	55,00000	64,78077	57,85279	55,31192	54,40020	57,27018
9	74,00000	70,76004	70,42289	72,69594	73,32962	70,00080

SANN - Results: Аркыш1 in neuro

Custom inputs

Custom inputs

	X6	X1	X2	X3	X4	X5	X7
Case 1	47	55	99	67	24	18	88

Custom inputs Custom predictions

#	1.Y	2.Y	3.Y	4.Y	5.Y	X6
1	81,9290...	108,828...	83,2511...	90,9999...	82,4614...	47,00000

SANN - Results: Аркыш1 in neuro

Active neural networks

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct.
1	MLP 7-6-1	0,533258	0,859875	0,473241	BFGS 2	SOS
2	MLP 7-7-1	0,548197	0,852883	0,390742	BFGS 2	SOS
3	MLP 7-11-1	0,491951	0,742211	0,573982	BFGS 4	SOS
4	MLP 7-10-1	0,985201	0,560917	0,554610	BFGS 0	SOS

Select/Deselect active networks Delete networks

Build models with CNN Build models with ANS Build models with Subsampling

Predictions | Graphs | Details | Custom predictions

Number of cases to predict: 1 Clear previous predictions

Custom inputs Custom predictions

#	1.Y	2.Y	3.Y	4.Y	5.Y	X6
1	81,9290...	108,828...	83,2511...	90,9999...	82,4614...	47,00000

Summary

Save networks

Cancel

Options

Samples

- Train
- Test
- Validation
- Missing

- PMML
- C/C++
- C#
- Java
- SAS
- SQL stored procedure in C#
- SQL User Defined Function in C#
- Teradata
- Deployment to STATISTICA Enterprise



ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

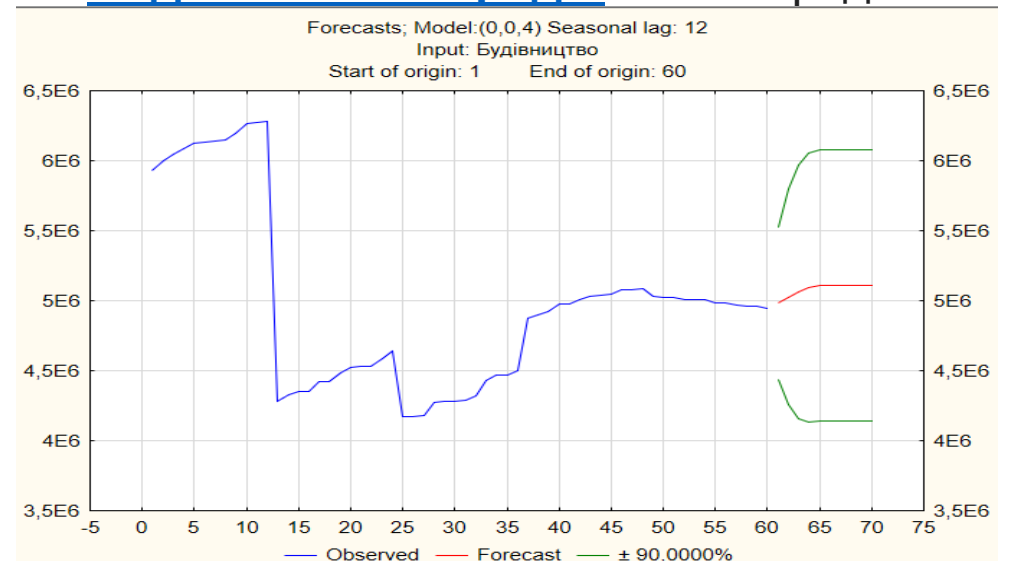
Прогнозування — процес передбачення майбутнього стану предмета чи явища на основі аналізу його минулого і сучасного, систематично оцінювана інформація про якісні й кількісні характеристики розвитку обраного предмета чи явища в перспективі. Результатом прогнозування є прогноз — знання про майбутнє і про ймовірний розвиток сьогочасних тенденцій конкретного явища-об'єкту в подальшому існуванні.

Існують два підходи до прогнозування: **якісний** та **кількісний**.

Кількісний підхід базується на математичних моделях й історичних даних. Якісний підхід покладається на освічену думку, інтуїцію й досвід професіоналів. Серед його різновидів є консенсус керівництва, Делфі-метод, оцінка торговими працівниками — кожного за своїм регіоном, опитування клієнтів.

Кількісні методи діляться на два види: **причинно-наслідкові** й моделі часових рядів. Часові ряди діляться на:

- моделі з декомпозицією: виділення сезонності й тренду;
- моделі згладжування:
 - середнє арифметичне;
 - ковзне середнє арифметичне;
 - середнє зважене;
 - ковзне середнє зважене;
 - експоненційне згладжування



ПРОГНОЗУВАННЯ ЕКОНОМІЧНИХ ПОКАЗНИКІВ

108 14.08.1974

Cost of AU_ - Аркуш1

Аркуш1		
	Дата	Вартість унції золота \$
1	29.04.1974	38,95
2	30.04.1974	38,82
3	01.05.1974	38,9
4	02.05.1974	
5	03.05.1974	
6	04.05.1974	
7	05.05.1974	
8	06.05.1974	

New analysis

New analysis

- Regression
- Classification
- Time series (regression)**

OK Cancel Options Open Data

Custom inputs Custom predictions

	1.Варті...	2.Варті...	3.Варті...	4.Варті...	5.Варті...	Дата_
	43.3955...	43.3108...	43.3836...	43.3426...	43.3641...	27259

	Algorithr				
2	MLP 1-4-1	0,986972	0,930855	0,977272	BFGS 4
3	MLP 1-7-1	0,987358	0,930530	0,978074	BFGS 5
4	MLP 1-4-1	0,986972	0,930855	0,977272	BFGS 3
					BFGS 7

SANN - Data selection: Аркуш1 in Cost of AU_

Quick Sampling (CNN and ANS) Subsampling Time series

Variables

Analysis variables (present in the dataset)

Select variables for analysis

1 - Дата	1 - Дата
2 - Вартість унції золота	2 - Вартість унції золота
3 - Var3	3 - Var3

Spread Zoom Spread Zoom

Continuous targets Continuous inputs

Quick Sampling (CNN and ANS) Subsampling Time series

Sampling method

Random sampling

Random sample sizes:

Train (%): 70

Test (%): 15

Validation (%): 15

Seed for sampling: 1000

Quick MLP activation functions Weight decay Initialization

Network types

MLP:

Min. hidden units: 4

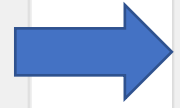
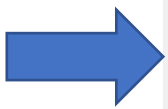
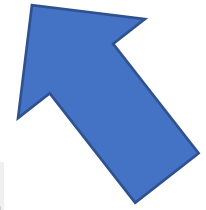
Max. hidden units: 12

Train/Retain networks

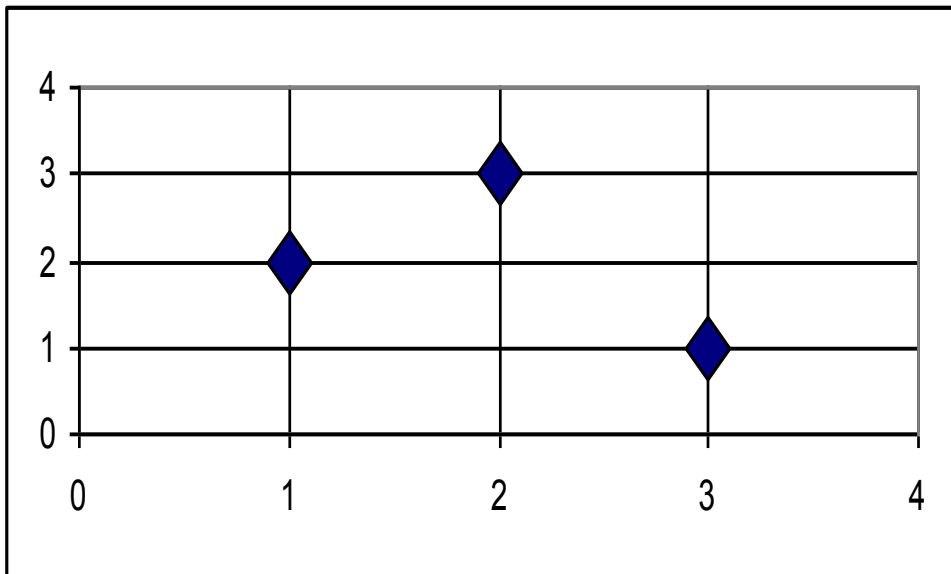
Networks to train: 20

Networks to retain: 5

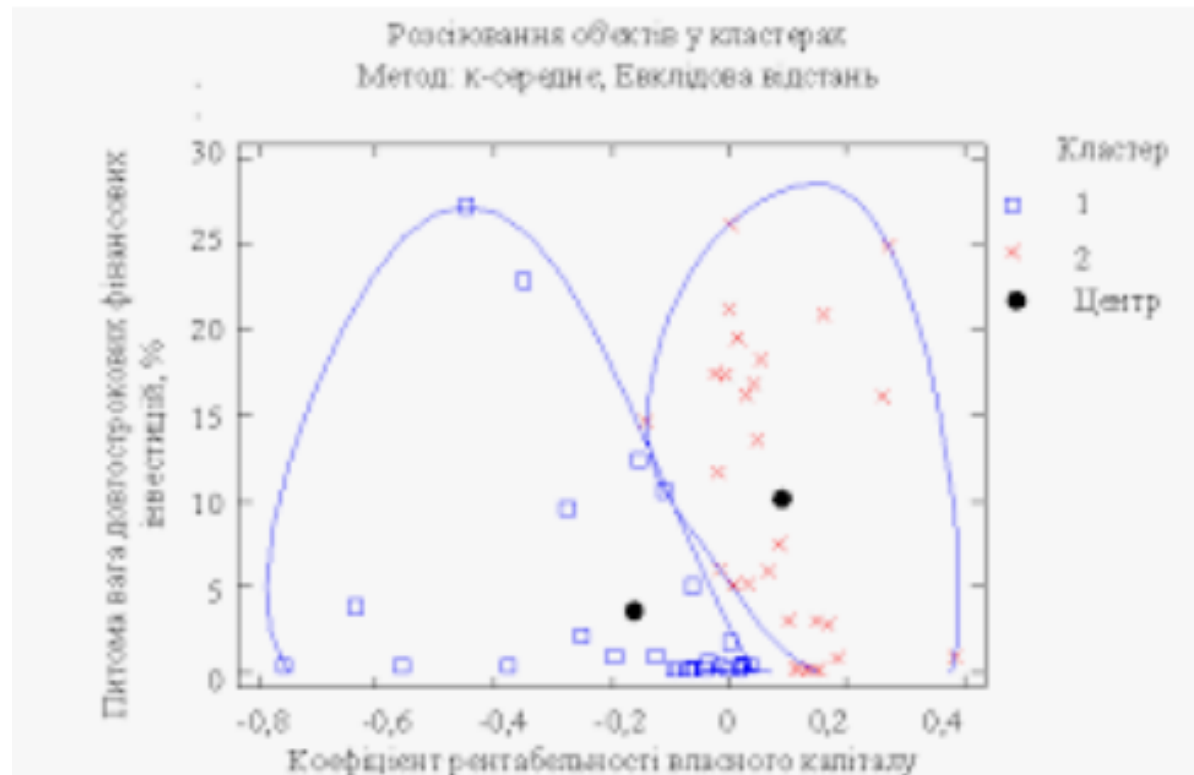
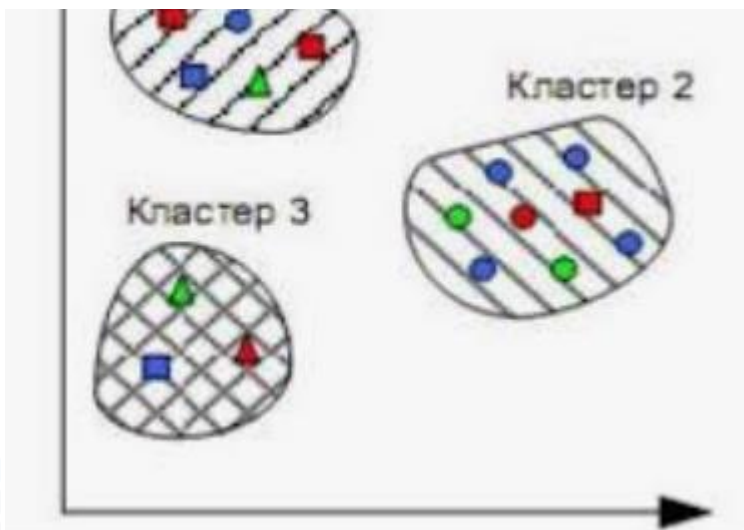
Error function



КЛАСТЕРИЗАЦІЯ ЕКОНОМІЧНИХ ОБ'ЄКТІВ



Представлення трьох об'єктів, як точок на площині



ВІДСТАНІ МІЖ ОБ'ЄКТАМИ

$$d_S(x_i; y_i) = \left(\sum_{i=1}^{Nf} |x_i - y_i|^p \right)^{\frac{1}{r}}$$

$$d_M(x_i; y_i) = \sqrt{\sum_{i=1}^{Nf} (\sqrt{x_i} - \sqrt{y_i})^2}$$

$$d_E(x_i; y_i) = \sqrt{\sum_{i=1}^{Nf} (x_i - y_i)^2}$$

$$d_{XEM}(x_i; y_i) = \sum_{i=1}^{Nf} (x_i - y_i)$$

$$d_{SUP}(x_i; y_i) = SUP|x_i - y_i|$$

$$d_L(x_i; y_i) = \sum_{i=1}^{Nf} |x_i - y_i|$$



Кластеризація повним перебором об'єктів

$$Z = \sum_{i=1}^{N_0} \sum_{j=1}^{N_0} q_{ij} d_{ij} \rightarrow \max$$

$$\sum_{i=1}^{N_0} q_{ij} \leq N_0$$

$$\sum_{j=1}^{N_0} q_{ij} = 1$$

$$\sum_{i=1}^{N_0} \sum_{j=1}^{N_0} q_{ij} = N_0$$

	1	2	3	4	5
1	0	0,89	0,41	0,74	0,46
2	0,89	0	0,86	0,87	0,61
3	0,41	0,86	0	0,96	0,66
4	0,74	0,87	0,96	0	0,62
5	0,46	0,61	0,66	0,62	0

		Об'єкти					Сума по клас-терам
		1	2	3	4	5	
Клас-тери	1	0	1	0	0	0	1
	2	1	0	0	0	0	1
	3	0	0	0	1	1	2
	4	0	0	1	0	0	1
	5	0	0	0	0	0	0
Сума по стовпцям		1	1	1	1	1	5

ПОЧАТКОВІ ДАНІ У ПРОГРАМІ EXCEL

Шахта	Общий объем добычи рядового угля, тыс.т./год	Зольность добытого рядового угля, %	Объем производства товарной угольной продукции, тыс.т./год	Зольность опгуженной товарной угольной продукции, %	Средней ствующее кол-во очистных забоев, заб	Средней ствующая линия очистных забоев, м	Средне-суточная нагрузка на очистной забой, т/сут	Средне-месячное подвигание линии очистных забоев, м/мес	Темпы проведения вскрывающихся и подготовливающих горных выработок, м/мес
Родинская	224,3	40,3	188,0	20,6	1,0	204,7	769,0	50,3	1440,0
Центральная	285,0	47,4	170,8	16,9	2,3	459,6	303,8	23,7	959,6
им. Г.М. Димитрова	371,5	44,2	312,3	21,2	1,2	291,4	1118,0	35,4	1048,5
им. А.Г. Стаханова	1224,8	44,7	721,0	11,9	3,6	763,3	725,6	37,4	5622,5
ГП "Красноармейскуголь"	2105,8	44,5	1313,0	15,5	8,0	1689,2	622,7	33,6	8648,7

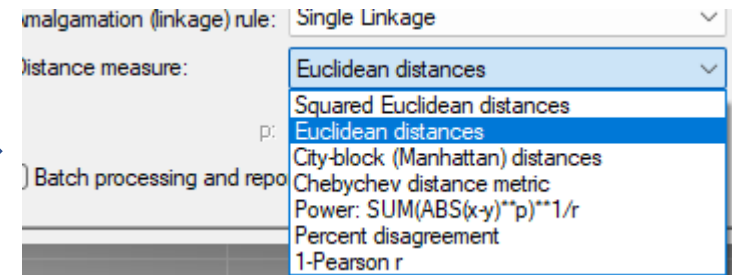
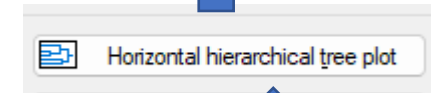
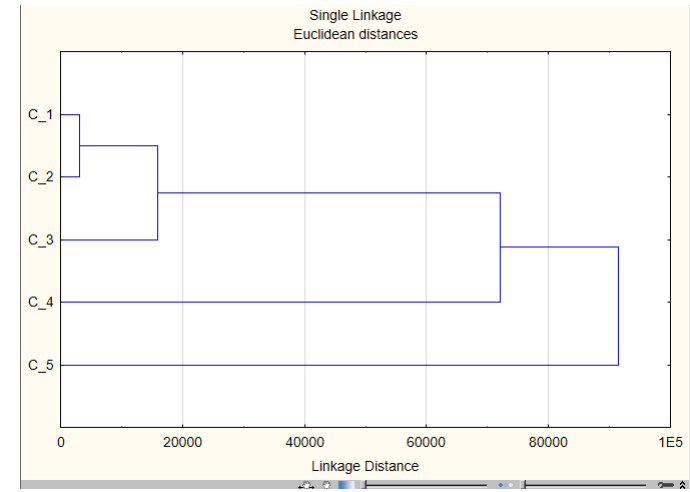
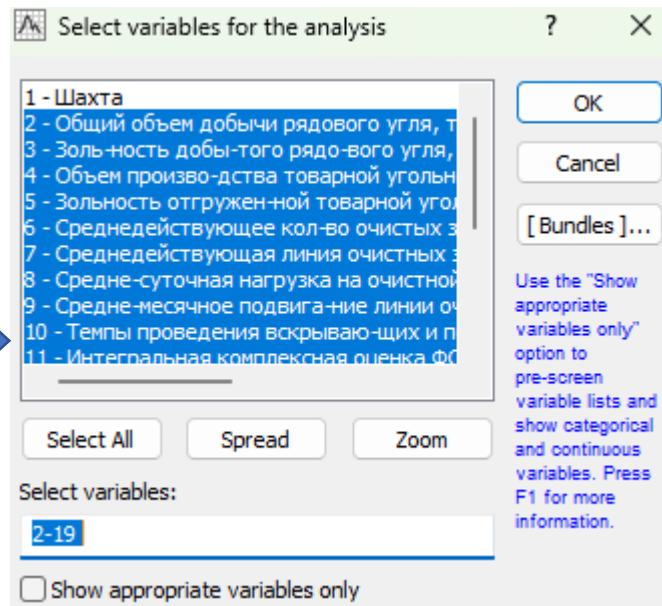
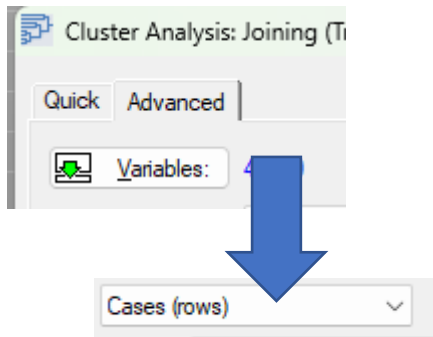
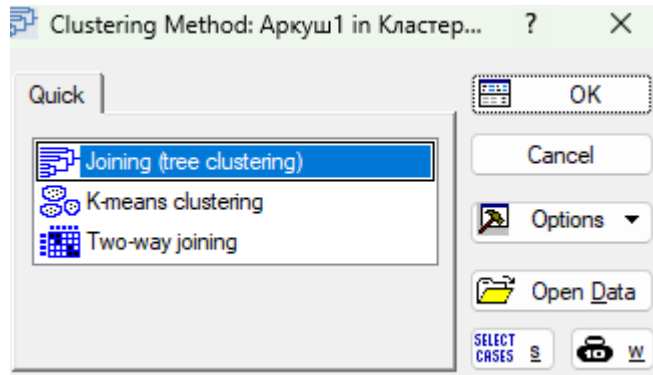
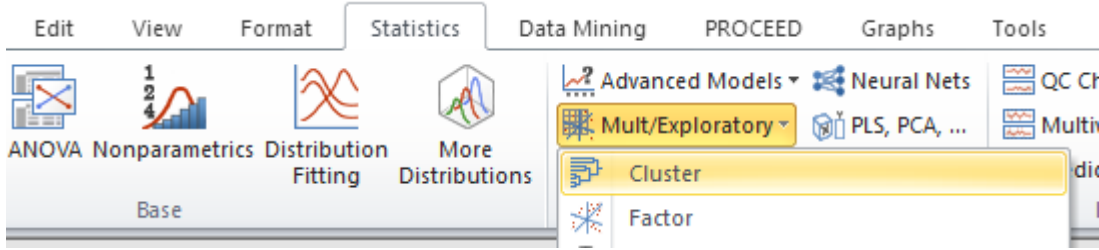
Після введення у програму Statistica

The screenshot shows the Statistica software interface. On the left, there is a navigation pane with 'Кластеризация' and 'Аркуш1'. The main window displays a table with the following data:

Шахта	Общий объем добычи рядового угля, тыс.т./год	Зольность добытого рядового угля, %	Объем производства товарной угольной продукции, тыс.т./год
1 Родинская	224,3	40,2801694	188,031797
2 Центральная	285	47,362407	170,81186
3 им. Г.М. Димитрова	371,52	44,1509851	312,253854
4 им. А.Г. Стаханова	1224,82	44,6887673	721,002869
5 ГП "Красноармейскуголь"	2105,76	44,4909809	1313,00535



Метод 1



МЕТОД 2

New analysis

- Regression
- Classification
- Time series (regression)
- Time series (classification)
- Cluster analysis**

Select an analysis type from the list above to start a new analysis. To deploy models from previous analyses, use the deployment option.

Buttons: OK, Cancel, Options, Open Data

Select variables for analysis

1 - Шахта	10 - Темпы проведения вскрытия
2 - Общий объем добычи рядово	11 - Интегральная комплексная
3 - Зольность добытого рядово	12 - Средне-суточная производи
4 - Объем производства товарн	13 - Объем реализованной товар
5 - Зольность отгруженной това	14 - Себестоимость 1 тонны доб
6 - Среднействующее кол-во	15 - Выработка товарной продук
7 - Среднействующая линия о	16 - Использование производств
8 - Средне-суточная нагрузка на	17 - Численность трудящихся (в
9 - Средне-месячное подвигани	18 - Цена 1т (без НДС), грн.
10 - Темпы проведения вскрытия	19 - План добычи, тыс.т/год

Buttons: Select All, Spread, Zoom

Continuous inputs (predictor): 1

Categorical inputs (predictor): 2-19

Buttons: OK, Cancel, [Bundles]...

Use the "Show appropriate variables only" option to pre-screen variable lists and show categorical and continuous variables. Press F1 for more information.

Select variables for analysis

10 - Темпы проведения вскрытия	1 - Шахта
11 - Интегральная комплексная	2 - Общий объем добычи рядово
12 - Средне-суточная производи	3 - Зольность добытого рядово
13 - Объем реализованной товар	4 - Объем производства товарн
14 - Себестоимость 1 тонны доб	5 - Зольность отгруженной това
15 - Выработка товарной продук	6 - Среднействующее кол-во
16 - Использование производств	7 - Среднействующая линия о
17 - Численность трудящихся (в	8 - Средне-суточная нагрузка на
18 - Цена 1т (без НДС), грн.	9 - Средне-месячное подвигани
19 - План добычи, тыс.т/год	10 - Темпы проведения вскрытия

Buttons: Select All, Spread, Zoom

Continuous inputs (predictor): 2-19

Categorical inputs (predictor): 1

Buttons: OK, Cancel, [Bundles]...

Use the "Show appropriate variables only" option to pre-screen variable lists and show categorical and continuous variables. Press F1 for more information.

Select variables for analysis

2 - Общий объем добычи рядово	1 - Шахта
3 - Зольность добытого рядово	
4 - Объем производства товарн	
5 - Зольность отгруженной това	
6 - Среднействующее кол-во	
7 - Среднействующая линия о	
8 - Средне-суточная нагрузка на	
9 - Средне-месячное подвигани	
10 - Темпы проведения вскрытия	
11 - Интегральная комплексная	
12 - Средне-суточная производи	

Buttons: Select All, Spread, Zoom

Continuous inputs (predictor): 2-19

Categorical inputs (predictor): 1

Buttons: OK, Cancel, [Bundles]...

Use the "Show appropriate variables only" option to pre-screen variable lists and show categorical and continuous variables. Press F1 for more information.

Show appropriate variables only

SANN - Data selection: Аркуш1 in Кластеризация

Quick | Sampling (CNN and ANS)

Buttons: OK, Cancel, Options

Set the test and/or validation sample fields to 0 to exclude these samples from the analysis.

MD handling (inputs)

- Casewise
- Mean substitution

SELECT CASES

- Case selection
- Case weights

Sampling method

Random sampling

- Random sample sizes:

Train (%): 70

Test (%): 15

Validation (%): 15

Seed for sampling: 1000

Subset variable

- Sampling variable:

Training sample	none
Testing sample	none
Validation sample	none



Kohonen graph

SANN - Results: Spreadsheet1

Net. ID	Net. name	Algorithm
3	SOFM 8-2	Kohonen 1000

Active neural networks

Predictions (Kohonen) | Graphs | Kohonen graph | Custom predictions |

Include

- Inputs
- Winning neuron position
- Winning neuron activation
- Variables

Weights

Data statistics

Single predictions

Dataset: 1

Activations

Options

Sample

- Train
- Test
- Validation

Summary

Save

Cancel

Options

0

0

none

1 2

SANN - Custom Neural Networks Spreadsheet1

Net. ID	Net. name	Algorithm
2	SOFM 8-6	Kohonen 1000

Active neural networks

Quick (Kohonen) | Kohonen Training | Initialization | Real time training graph |

Dimensions

Topological height: 2

Topological width: 3

Train

Go to results

Save networks

Data statistics

Summary

Cancel

Options

2 3

SANN - Results: Spreadsheet1

Net. ID	Net. name	Algorithm
2	SOFM 8-6	Kohonen 1000

Active neural networks

Predictions (Kohonen) | Graphs | Kohonen graph | Custom predictions |

Summary

Save

Cancel

Options

Sample

- Train
- Test
- Validation

0

0

none

5 3

SANN - Results: Spreadsheet1

Net. ID	Net. name	Algorithm
3	SOFM 8-2	Kohonen 1000

Active neural networks

Predictions (Kohonen) | Graphs | Kohonen graph | Custom predictions |

Summary

Save

Cancel

Options

Sample

- Train
- Test
- Validation

0

0

none

SANN - Results: Spreadsheet1

Net. ID	Net. name	Algorithm
4	SOFM 8-6	Kohonen 1000

Active neural networks

Predictions (Kohonen) | Graphs | Kohonen graph | Custom predictions |

Summary

Save

Cancel

Options

Sample

- Train
- Test
- Validation

0

0

none

SANN - Results: Spreadsheet1

Net. ID	Net. name	Algorithm
5	SOFM 8-15	Kohonen 1000

Active neural networks

Predictions (Kohonen) | Graphs | Kohonen graph | Custom predictions |

Summary

Save

Cancel

Options

Sample

- Train
- Test
- Validation

0

0

none

Kohonen graph –Select All

SANN - Results: Архив1 in neuro

Active neural networks

Net. ID	Net. name	Algorithm
1	SOFM 8-6	Kohonen 1000

Select\Deselect active networks Delete networks

Predictions (Kohonen) | Graphs | Kohonen graph | Custom predictions

Summary
Save
Cancel
Options
Sample
 Train
 Test
 Validation

6
22
multiple

Kohonen Kohonen **Select All** Clear

Sheet for selected cells (Архив1 in neuro)

Kohonen spreadsheet for selected cells (Архив1 in neuro)
Number of selected cases: 22, Number of selected cells: 6
Samples: Train

Case name	Neuron location	Neuron ID	Activation	X1	X2	X3	X4
1	(2, 1)	4	0,525324	55,00000	86,00000	93,00000	89,00000
2	(1, 1)	1	0,543065	74,00000	55,00000	86,00000	93,00000
4	(1, 3)	3	0,846809	47,00000	98,00000	74,00000	55,00000
5	(2, 3)	6	0,982716	90,00000	47,00000	98,00000	74,00000
7	(2, 2)	5	0,433032	63,00000	52,00000	90,00000	47,00000
8	(2, 3)	6	0,679896	91,00000	63,00000	52,00000	90,00000
9	(1, 3)	3	0,624299	71,00000	91,00000	63,00000	52,00000
11	(1, 3)	3	0,799830	78,00000	78,00000	71,00000	91,00000
13	(1, 2)	2	0,007484	69,00000	89,00000	78,00000	78,00000
14	(1, 3)	3	0,532084	70,00000	69,00000	89,00000	78,00000
15	(2, 1)	4	0,651505	52,00000	70,00000	69,00000	89,00000
16	(2, 2)	5	0,423332	67,00000	52,00000	70,00000	69,00000
17	(2, 3)	6	0,416690	87,00000	67,00000	52,00000	70,00000
19	(2, 1)	4	0,470843	48,00000	86,00000	87,00000	67,00000
21	(1, 1)	1	0,496103	52,00000	60,00000	48,00000	86,00000
22	(2, 3)	6	0,757757	85,00000	52,00000	60,00000	48,00000
23	(2, 3)	6	0,641415	96,00000	85,00000	52,00000	60,00000
24	(1, 3)	3	0,734147	69,00000	96,00000	85,00000	52,00000
25	(2, 1)	4	0,432251	61,00000	69,00000	96,00000	85,00000
26	(1, 1)	1	0,452933	67,00000	61,00000	69,00000	96,00000
27	(1, 1)	1	0,576822	53,00000	67,00000	61,00000	69,00000
28	(2, 3)	6	0,420322	86,00000	53,00000	67,00000	61,00000



КЛАСИФІКАЦІЯ ОБ'ЄКТІВ

Класифікація (лат., розряд і ...фікація): система розподілу предметів, явищ або понять на класи, групи тощо за спільними ознаками, властивостями.

Класифікація — це багаторівневий, послідовний поділ обсягу поняття з метою систематизації, поглиблення та отримання нових знань стосовно членів поділу. Результатом класифікації є система підпорядкованих понять: подільне поняття є родом, а нові поняття є видами цього роду, підвидами видів тощо.

Типові постановки задач класифікації:

Задача ідентифікації, яка полягає у тому, щоб вирізнити певний конкретний об'єкт серед йому подібних (наприклад, впізнати серед інших людей свою дружину).

1. Віднесення об'єкта до того чи іншого класу. Це може бути, наприклад, задача розпізнавання літер або прийняття рішення про наявність дефекту у деякій технічній деталі.

2. Віднесення об'єкта до певного класу відображає найтипівішу проблему класифікації, і, коли говорять про розпізнавання образів, найчастіше мають на увазі саме цю проблему. Саме вона розглядається тут у першу чергу.

Об'єкт відноситься до кластера, **якщо відстань від об'єкта до центра кластера менша за радіус кластера**. Якщо ця умова не виконується для двох та більше кластерів, об'єкт є спірним. Неоднозначність даного завдання може бути усунена експертом або аналітиком.



X1	X2	X3	X4	X5	X6	X7	Y	Номер кластеру
55	86	93	89	59	52	53	68	4
74	55	86	93	89	59	52	53	1
98	74	55	86	93	89	59	52	3
47	98	74	55	86	93	89	59	6
90	47	98	74	55	86	93	89	5
52	90	47	98	74	55	86	93	6
63	52	90	47	98	74	55	86	3
91	63	52	90	47	98	74	55	3
71	91	63	52	90	47	98	74	2
78	71	91	63	52	90	47	98	3
78	78	71	91	63	52	90	47	4
89	78	78	71	91	63	52	90	5
69	89	78	78	71	91	63	52	6
70	69	89	78	78	71	91	63	4
52	70	69	89	78	78	71	91	1
67	52	70	69	89	78	78	71	6
87	67	52	70	69	89	78	78	6
86	87	67	52	70	69	89	78	3
48	86	87	67	52	70	69	89	4
60	48	86	87	67	52	70	69	1
52	60	48	86	87	67	52	70	1
85	52	60	48	86	87	67	52	6

SANN - New Analysis/Deployment: Арыш1 in neuro

New analysis/Deployment

Deployment

Deploy models from previous analyses

New analysis

Load network files

File name	Net. ID	Net. name	Hidden act.	Output act.

PMML file list

New analysis

- Regression
- Classification
- Time series (regression)
- Time series (classification)
- Cluster analysis

Select an analysis type from the list above to start a new analysis. To deploy models from previous analyses, use the deployment option.

OK Cancel Options Open Data

Select variables for analysis

1 - X1	1 - X1	1 - X1
2 - X2	2 - X2	2 - X2
3 - X3	3 - X3	3 - X3
4 - X4	4 - X4	4 - X4
5 - X5	5 - X5	5 - X5
6 - X6	6 - X6	6 - X6
7 - X7	7 - X7	7 - X7
8 - Y	8 - Y	8 - Y
9 - Номер кластеру	9 - Номер кластеру	9 - Номер кластеру

Spread Zoom Spread Zoom Spread Zoom

Categorical target: 9 Continuous inputs: 1-8 Categorical inputs:

Show appropriate variables only

OK Cancel [Bundles]...

Use the "Show appropriate variables only" option to pre-screen variable lists and show categorical and continuous variables. Press F1 for more information.

SANN - Automated Network Search (ANS): Арыш1 in neuro

Active neural networks

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct.
1	MLP 8-11-6	75,000000	33,333333	66,666667	BFGS 0	CE
2	MLP 8-9-6	62,500000	33,333333	33,333333	BFGS 0	CE
3	MLP 8-7-6	31,250000	33,333333	0,000000	BFGS 1	SOS
4	MLP 8-4-6	100,000000	0,000000	0,000000	BFGS 0	CE
5	MLP 8-10-6	21,250000	33,333333	0,000000	BFGS 1	CE

Quick | MLP activation functions | Weight decay | Initialization

Network types

MLP:

Min. hidden units: 4 Max. hidden units: 12

RBF:

Min. hidden units: 5 Max. hidden units: 6

Train/Retain networks

Networks to train: 20 Networks to retain: 5

Error function

Sum of squares Cross entropy

Train Go to results Save networks Data statistics Summary Cancel Options

SANN - Results: Арыш1 in neuro

Active neural networks

Net. ID	Net. name	Training perf.	Test perf.	Validation perf.	Algorithm	Error funct.
1	MLP 8-11-6	75,000000	33,333333	66,666667	BFGS 0	CE
2	MLP 8-9-6	62,500000	33,333333	33,333333	BFGS 0	CE
3	MLP 8-7-6	31,250000	33,333333	0,000000	BFGS 1	SOS
4	MLP 8-4-6	100,000000	0,000000	0,000000	BFGS 0	CE
5	MLP 8-10-6	21,250000	33,333333	0,000000	BFGS 1	CE

Select/Deselect active networks Delete networks

Build models with CNN Build models with ANS Build models with Subsampling

Predictions | Graphs | Details | Liftcharts | Custom predictions

Number of cases to predict: 7 Clear previous predictions

Custom inputs Custom predictions

#	1.Home...	2.Home...	3.Home...	4.Home...	5.Home...	X1
1	6	3	6	2	6	96
2	3	4	6	4	6	69
3	1	4	6	5	6	61
4	1	1	6	1	6	67
5	6	6	6	1	6	53

Summary Save networks Cancel Options

Samples Train Test Validation Missing

