

```
## Warning in grepl(db, input): input string 26 is invalid in this locale
## Warning in grepl(db, input): input string 27 is invalid in this locale
## Warning in grepl(db, input): input string 28 is invalid in this locale
## Warning in grepl(db, input): input string 31 is invalid in this locale
## Warning in grep("^\\\\\\bibliography.+", input, value = TRUE): input
string 26 is invalid in this locale
## Warning in grep("^\\\\\\bibliography.+", input, value = TRUE): input
string 27 is invalid in this locale
## Warning in grep("^\\\\\\bibliography.+", input, value = TRUE): input
string 28 is invalid in this locale
## Warning in grep("^\\\\\\bibliography.+", input, value = TRUE): input
string 31 is invalid in this locale
## Warning in grep("^\\\\\\bibliography.+", input, value = TRUE): input
string 36 is invalid in this locale
```

РОЗДІЛ 1

**МОДЕЛЮВАННЯ ТА ПРАКТИЧНЕ ЗАСТОСУВАННЯ  
РОЗРОБЛЕНИХ МЕТОДІВ ТА АРХІТЕКТУР**

Для проведення чисельних експериментів, що наведені у підрозділах 1.1, 1.2 та 1.3 було обрано такі критерії оцінки:

- RMSE (Root Mean Squared Error, середньоквадратична похибка),
- SMAPE (Symmetric Mean Absolute Percentage Error, симетрична абсолютно процентна похибка),

що обчислюються за формулами

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2}, \quad (1.1)$$

та

$$\text{SMAPE} = \frac{1}{N} \sum_{k=1}^N \frac{|y(k) - \hat{y}(k)|}{|y(k) + \hat{y}(k)|} \quad (1.2)$$

відповідно, де  $y$  – шуканий сигнал,  $\hat{y}$  – вихідний сигнал системи.

### **1.1. Моделювання розширенного нейро-фаззі нейрона**

Датасет для першого есперименту (фазовий портрет наведено на рис. 1.1) було сгенеровано за формулою

$$\sin(k + \sin(2k)) \text{ для } k \in [1, 600]. \quad (1.3)$$

Результати експерименту наведено у таблиці нижче, а також проілюстровано залежність точності прогнозу від порядку висновування (рис. 1.2) та кількості функцій належності (рис. 1.3).

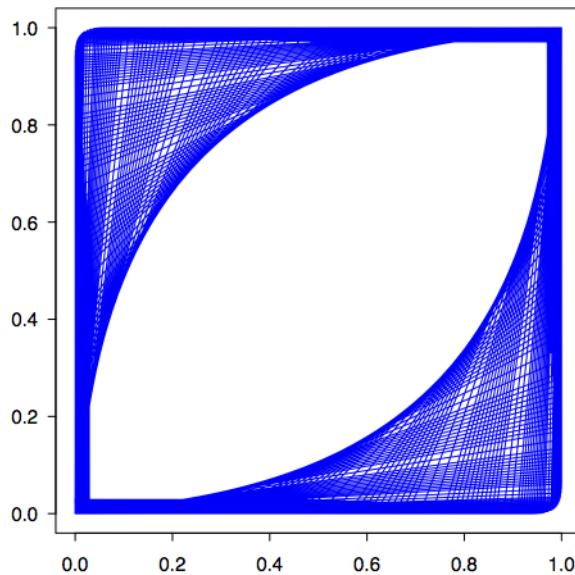


Рис. 1.1. Фазовий портрет штучно сгенерованого датасету

Отже, як видно з таблиці **TODO table** та на рис. 1.4, можна зробити висновок, що точність прогнозу розширеного нео-фаззі нейрону вища від точності звичайного нео-фаззі нейрону (ENFN з нульовим порядком висновування).

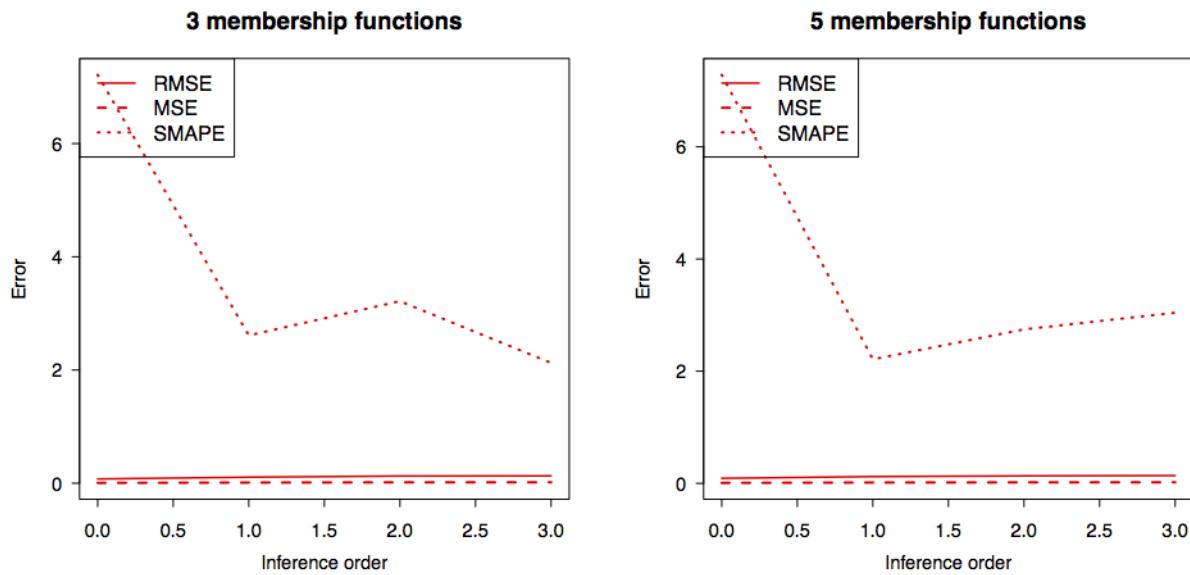


Рис. 1.2. Похибка прогнозу розширеного нео-фаззі нейрону від порядку висновування (для трьох та п'яти функцій належності)

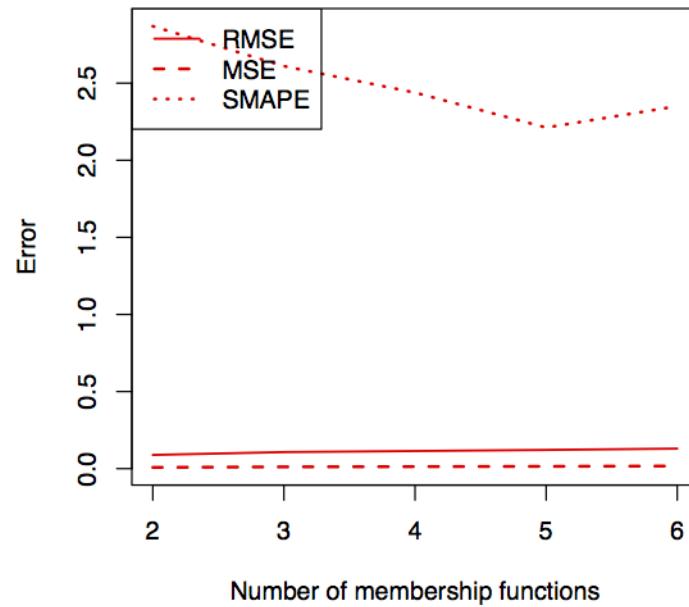


Рис. 1.3. Похибка прогнозу розширенного нео-фаззі нейрону від кількості функцій належності (порядок висновування - 2)

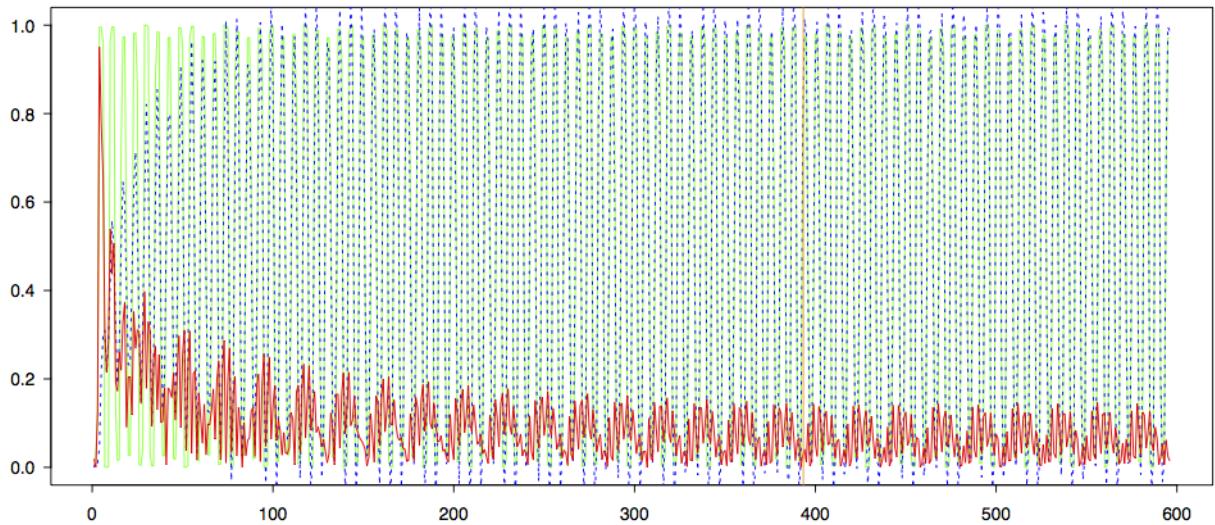


Рис. 1.4. Прогноз розширенного нео-фаззі нейрону з 3 трикутними функціями належності, що реалізує нечітке висновування 3-ого порядку (зелена лінія – шуканий сигнал, синя пунктирна лінія – прогноз розширенного нео-фаззі нейрону, червона лінія – похибка; жовтобагряна вертикальна лінія позначає закінчення тренувальної частини датасету)

Нечітке висновування 0 порядку			
Функції належності		RMSE	SMAPE
	3	0.07593154138757	7.2130842903092
	4	0.0816015155371995	3.74330006151476
	5	0.0899940897089563	7.28423693259539
	6	0.098548606433121	4.97414996889677
Нечітке висновування I порядку			
Функції належності		RMSE	SMAPE
	3	0.107342623622113	2.61031587883906
	4	0.114302161682684	2.43845462677015
	5	0.121556611017793	2.21187746150696
	6	0.129922583600673	2.3504927587044
Нечітке висновування II порядку			
Функції належності		RMSE	SMAPE
	3	0.127570672613453	3.21603084937958
	4	0.121809301731276	2.49191939755954
	5	0.134120568445479	2.74555892364861
	6	0.146724724859333	2.39716069975922
Нечітке висновування III порядку			
Функції належності		RMSE	SMAPE
	3	0.130173960343431	2.12509843555904
	4	0.154580667425542	2.17237774290269
	5	0.139334351692536	3.04236918521622
	6	0.150497843534837	2.43818195954604

Таблиця 1.1

## Точність прогнозу розширеного нео-фаззі нейрону на штучно сгенерованому датасеті

Для подальшої апробації розширеного нео-фаззі нейрону розглянемо задачу прогнозування хаотичного ряду, що описується диференціальним рівнянням Мекі-Гласса [147 vik]:

$$y'(t) = \frac{0.2(t - \tau)}{1 + y^{10}(t - \tau)} - 0.1y(t), \quad (1.4)$$

при цьому значення часового ряду в кожній точці обчислене за допомогою методу Рунге-Кутта четвертого порядку. Часовий крок прийнятий рівним 0.1, початкові умови:  $x(0) = 1.2$ .

Традиційно завдання прогнозування полягає у визначенні  $x(t+6)$ х часового ряду (5.2) з параметром затримки 17 по значенням  $x(t+18)$ ,  $x(t+12)$ ,  $x(t+6)$  і  $x(t)$ . Перед початком обробки отриманий часовий ряд нормувався таким чином, щоб його значення лежали в інтервалі  $[0, 1]$  (область визначення трикутних функцій належності та кубічних сплайнів, що використовуються у синапсах розширеного нео-фаззі нейрону).

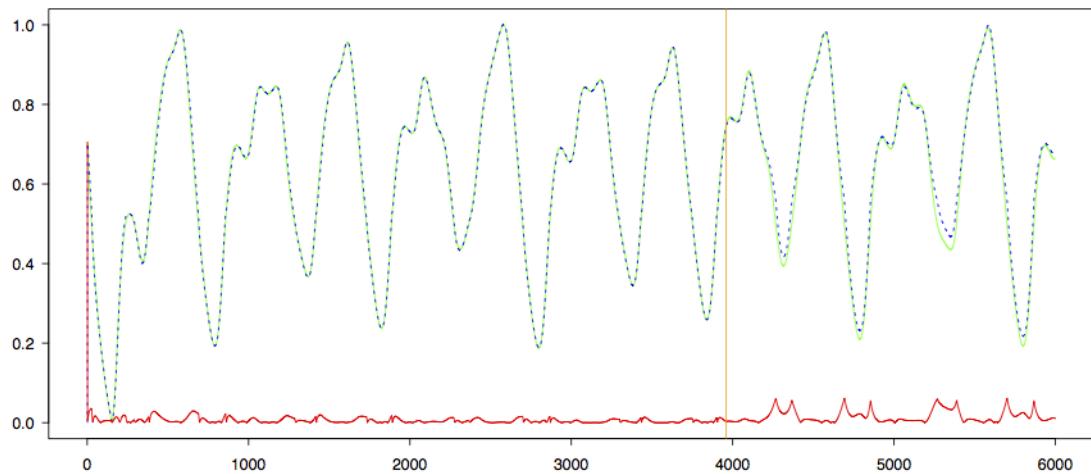


Рис. 1.5. Прогнозування хаотичного часового ряду Макі-Гласса розширенним нео-фаззі нейроном з 3 трикутними функціями належності, що реалізує нечітке висновування 3-ого порядку (зелена лінія – шуканий сигнал, синя пунктирна лінія – сигнал на виході нейрону, червона лінія – похибка; жовто-багряна вертикальна лінія позначає закінчення тренувальної частини датасету)

На рис.1.5 зображений результат прогнозування розширеного нео-фаззі нейрона з трьома функціями належності, що реалізує нечітке висновування 3-ого порядку. Для порівняння на рис. 1.6 наведено прогнозування традиційного нео-фаззі нейрона (що реалізує нечітке висновування нульового порядку) з аналогічними функціями належності (3 трикутні функції належності).

Нечітке висновування 0 порядку			
Функції належності	RMSE	SMAPE	
3	0.143482904587951	0.325353000527264	
4	0.106294989490131	0.264501060116694	
5	0.094578207548207	0.259630642224594	
6	0.094578207548207	0.259630642224594	
Нечітке висновування I порядку			
Функції належності	RMSE	SMAPE	
3	0.0491962584140483	0.136034984261077	
4	0.0300830061654526	0.0978878141219238	
5	0.0312245619190396	0.10776716206662	
6	0.0277981639612314	0.0989376310206592	
Нечітке висновування II порядку			
Функції належності	RMSE	SMAPE	
3	0.0227629037042365	0.0658002874902031	
4	0.0295377560571133	0.108260856276106	
5	0.0283735498990618	0.0946515030316458	
6	0.0251818186025806	0.0847398989365615	
Нечітке висновування III порядку			
Функції належності	RMSE	SMAPE	
3	0.0221516072948077	0.0622040928066835	
4	0.0261638552968437	0.0827913036508308	
5	0.0274818416982828	0.114797690969503	
6	0.0237394982902305	0.0846415828789446	

Таблиця 1.2

**Точність прогнозу ряду Мекі-Глассу розширенним нео-фаззі нейроном від порядку висновування та кількості функцій належності**

На рис. 1.7 показана залежність похибки від порядку висновування розширенних нео-фаззі нейронів з трьома та п'ятьма дзвонуватими функціями належності.

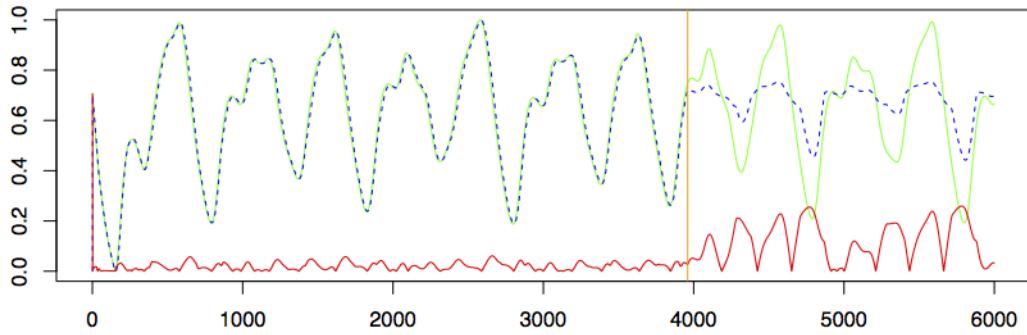


Рис. 1.6. Прогнозування хаотичного часового ряду Макі-Гласса традиційним нео-фаззі нейроном з 3 трикутними функціями належності (зелена лінія – шуканий сигнал, синя пунктирна лінія – сигнал на виході нейрону, червона лінія – похибка; жовтобагряна вертикальна лінія позначає закінчення тренувальної частини датасету)

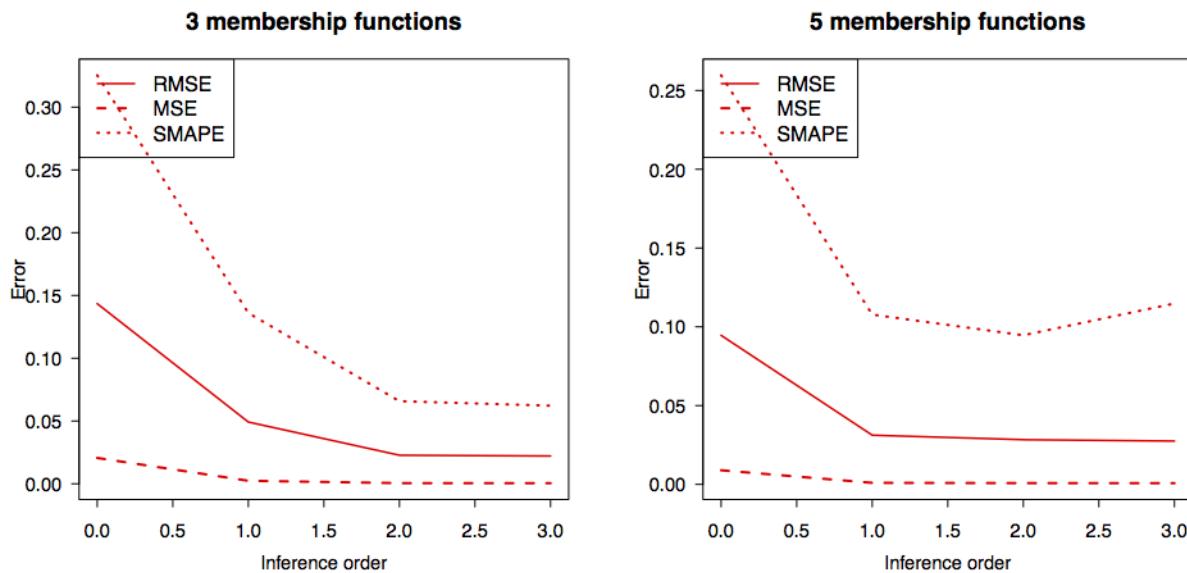


Рис. 1.7. Похибка прогнозування хаотичного часового ряду Макі-Гласса розширенними нео-фаззі нейронами з 3 та 5 дзвонуватими функціями належності від порядку нечіткого висновування

Як видно з таблиці **tableref** та рис. 1.7, розширенний нео-фаззі нейрон, що реалізує нечітке висновування вищого від 0 порядку, прогнозує хаотичний часовий ряд за рівнянням Мекі-Гласса з суттєво вищою точністю ніж

традиційний нео-фаззі нейрон.

Також пропонований розширений нео-фаззі нейрон було апробовано на реальному часовому ряді «Споживання електоренергії у м. Сімферополь за 2007 рік» (фазовий портрет наведено на рис. 1.8).

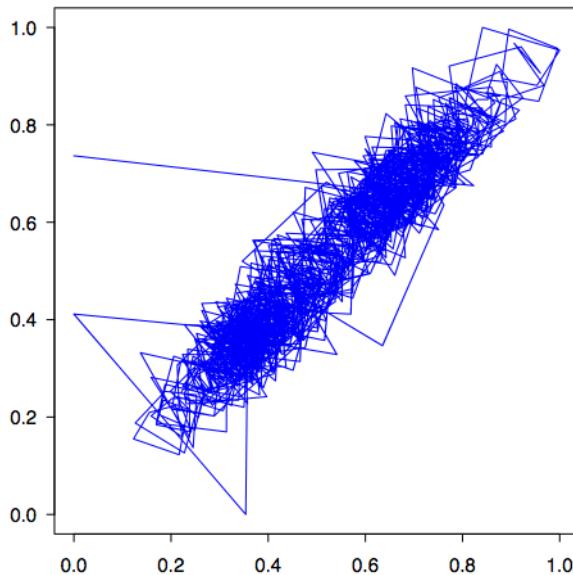


Рис. 1.8. Фазовий портрет часовогого ряду «Споживання електоренергії у м. Сімферополь за 2007 рік»

Найліпший прогноз ( $\text{RMSE} \approx 0.14$ ,  $\text{SMAPE} \approx 0.21$ ) надав нейрон, що реалізує нечітке висновування 1-ого порядку з 6-ма функціями належності (рис. 1.9)

Експерименти, що описані у цьому підрозділі, підтверджують, що розширені нео-фаззі нейрони, які реалізують нечітке висновування довільного порядку, мають підвищенну точність прогнозування хаотичних рядів (як штучно сгенерованих, так і реальних) порівняно з традиційними нео-фаззі нейронами (які реалізують нечітке висновування нульового порядку).

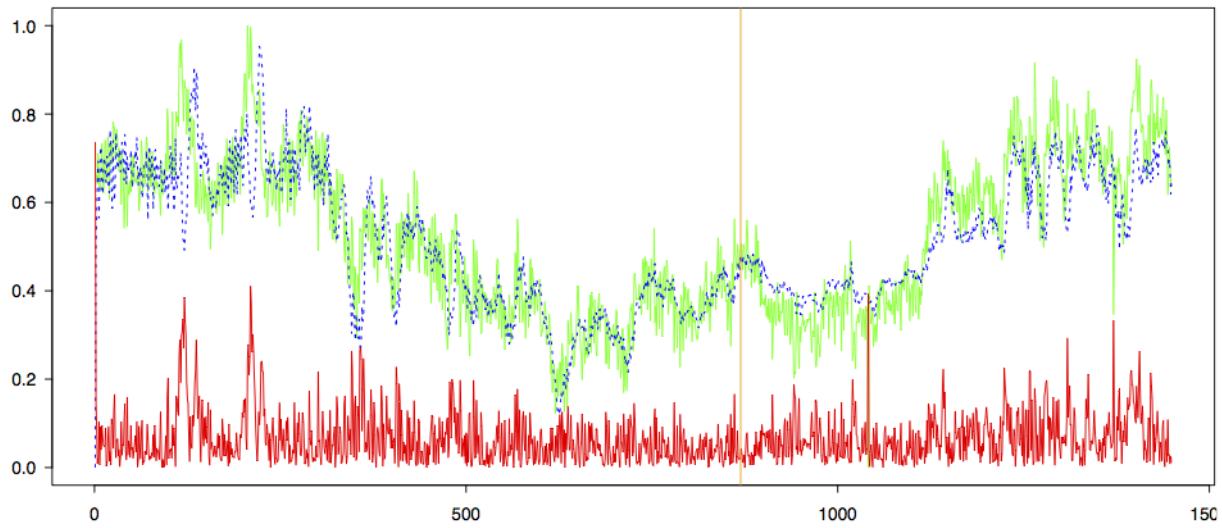


Рис. 1.9. Прогнозування часового ряду «Споживання електоренергії у м. Сімферополь за 2007 рік» розширеним нео-фаззі нейроном з 6-ма дзвонувати функціями належності, що реалізує нечітке висновування 1-ого порядку (зелена лінія – шуканий сигнал, синя пунктирна лінія – сигнал на виході нейрона, червона лінія – похибка; жовтобагряна вертикальна лінія позначає закінчення тренувальної частини датасету)

## 1.2. Моделювання гібридної каскадної нейро-фаззі мережі з розширенними нео-фаззі нейронами та оптимізацією пулу нейронів

Низку експериментів для апробації гібридної каскадної нейро-фаззі мережі на розширенних нео-фаззі нейронах з оптимізацією пулу нейронів було проведено на датасетах, що їх надала дослідницька група «The Applications of Machine Learning (AML)» з Університету Аалто, що у Фінляндії (Aalto University School of Science, Espoo, Finland). Одним з таких датасетів є часовий ряд «Споживання електроенергії у Польщі за період з 1990-х років» (фазовий портрет наведено на рис 1.10).

Вихідний сигнал гібридної каскадної нейро-мережі наведено на рис. 1.11, а похибки розширенних нео-фаззі нейронів та нейронів-узагальнювачів для кожного каскаду системи наведено у таблиці **TODO**.

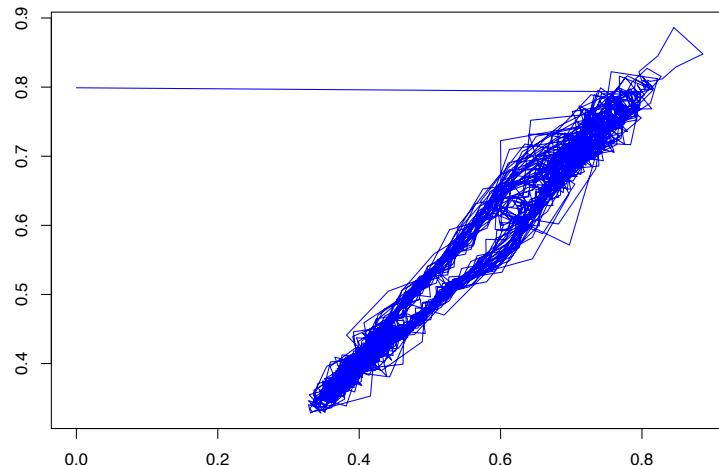


Рис. 1.10. Фазовий портрет часового ряду Прогнозування часового ряду «Споживання електроенергії у Польщі за період з 1990-х років»

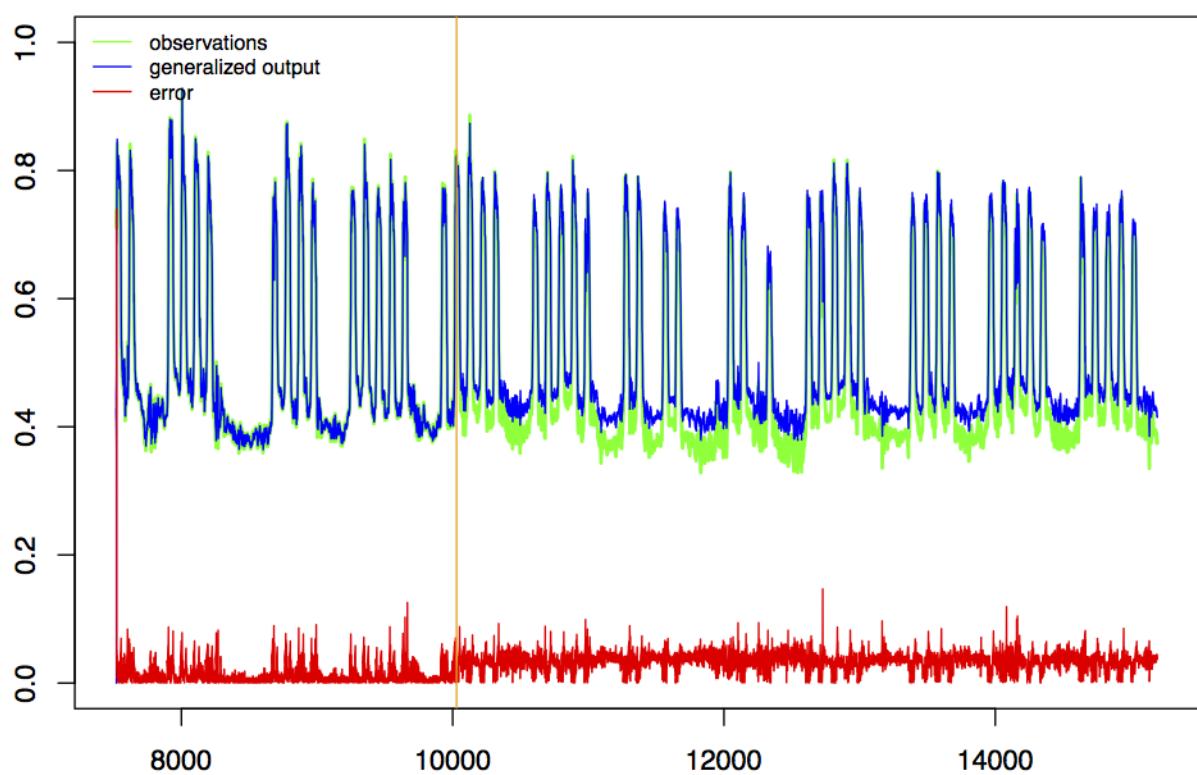


Рис. 1.11. Прогнозування часового ряду «Споживання електоренергії у м. Сімферополь за 2007 рік» гібридної каскадної нейро-фаззі мережі з оптимізацією пулу нейронів (жовтобагряна вертикальна лінія позначає закінчення тренувальної частини датасету)

	SMAPE	RMSE
Каскад I		
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.080830	0.39288
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.075014	0.036680
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.074955	0.038170
Нейрон-узагальнювач I каскаду	0.059835	0.030302
Каскад II	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.080835	0.039290
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.059837	0.036683
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.074966	0.038195
Нейрон-узагальнювач II каскаду	0.059821	0.036683
Каскад III	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.080934	0.039333
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.059869	0.036711
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.75009	0.038213
Нейрон-узагальнювач III каскаду	0.059869	0.030320
Каскад IV	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.080892	0.039316
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.059869	0.030303
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.075034	0.038213
Нейрон-узагальнювач IV каскаду	0.059849	0.030303
Нейрон-узагальнювач системи	0.059821	0.030302

Таблиця 1.3

**Результати прогнозування часового ряду «Споживання електоренергії у м. Сімферополь за 2007 рік» нейронів (зокрема нейронів-узагальнювачів) каскадної нейро-фаззі системи**

Фазовий портрет другого датасету «Коливання рівню приловоотливної зони» (Subtidal coastal level of fluctuations) наведено на рис. 1.11, похибки вузлів системи – у таблиці TODO.

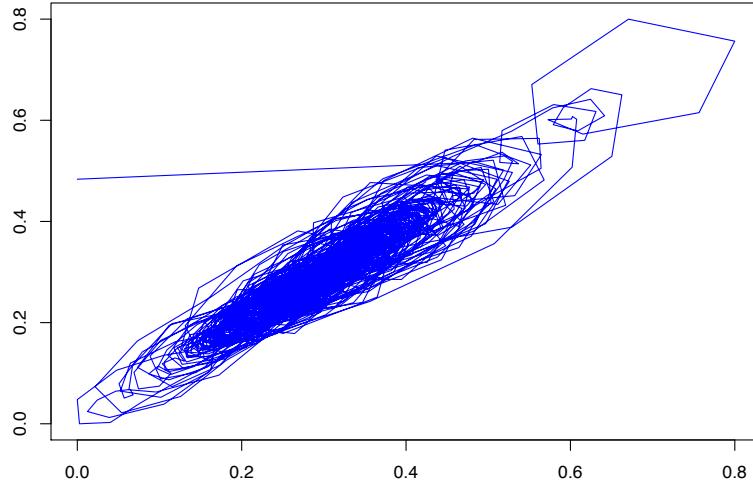
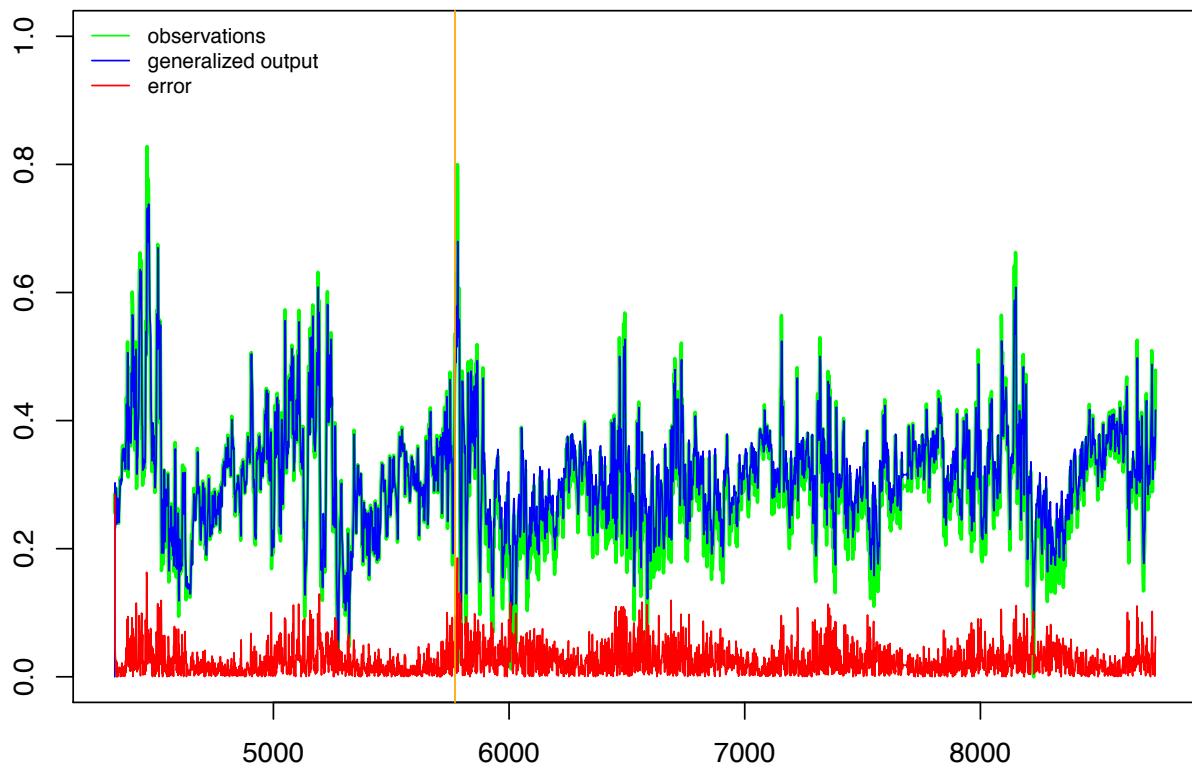


Рис. 1.12. Фазовий портрет часового ряду Прогнозування часового ряду «Коливання рівню приловоотливної зони»



Прогнозування часового ряду «Коливання рівню приловоотливної зони» гібридної каскадної нейро-фаззі мережі з оптимізацією пулу нейронів (жовтобагряна вертикальна лінія позначає закінчення тренувальної частини датасету)

	SMAPE	RMSE
Каскад I		
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.110067	0.036612
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.105192	0.035474
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.103129	0.034814
Нейрон-узагальнювач I каскаду	0.105598	0.035301
Каскад II	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.110023	0.036593
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.105118	0.035457
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.103148	0.035818
Нейрон-узагальнювач II каскаду	0.102377	0.034698
Каскад III	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.110013	0.036591
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.105126	0.035458
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.103155	0.034820
Нейрон-узагальнювач III каскаду	0.103155	0.034820
Каскад IV	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.110026	0.036594
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.105153	0.035464
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.103168	0.034822
Нейрон-узагальнювач IV каскаду	0.102545	0.034083
Нейрон-узагальнювач системи	0.102377	0.030302

Таблиця 1.4

**Результати прогнозування часового ряду «Коливання рівню приловоотливної зони» нейронів (зокрема нейронів-узагальнювачів) каскадної нейро-фаззі системи**

Останній датасет було взято зі змагань у прогнозуванні часових рядів «European Symposium on Time Series Prediction 2008», фазовий портрет наведено на рис. 1.13, результат роботи пропонованої системи – на рис. 1.2.

Результати прогнозування нейронів окремих каскадів, нейронів-узагальнювачів, а також системи в цілому можна побачити у таблиці table.

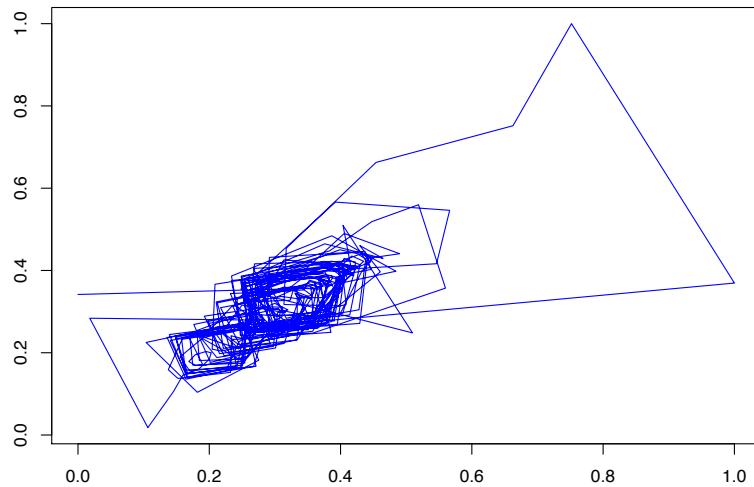
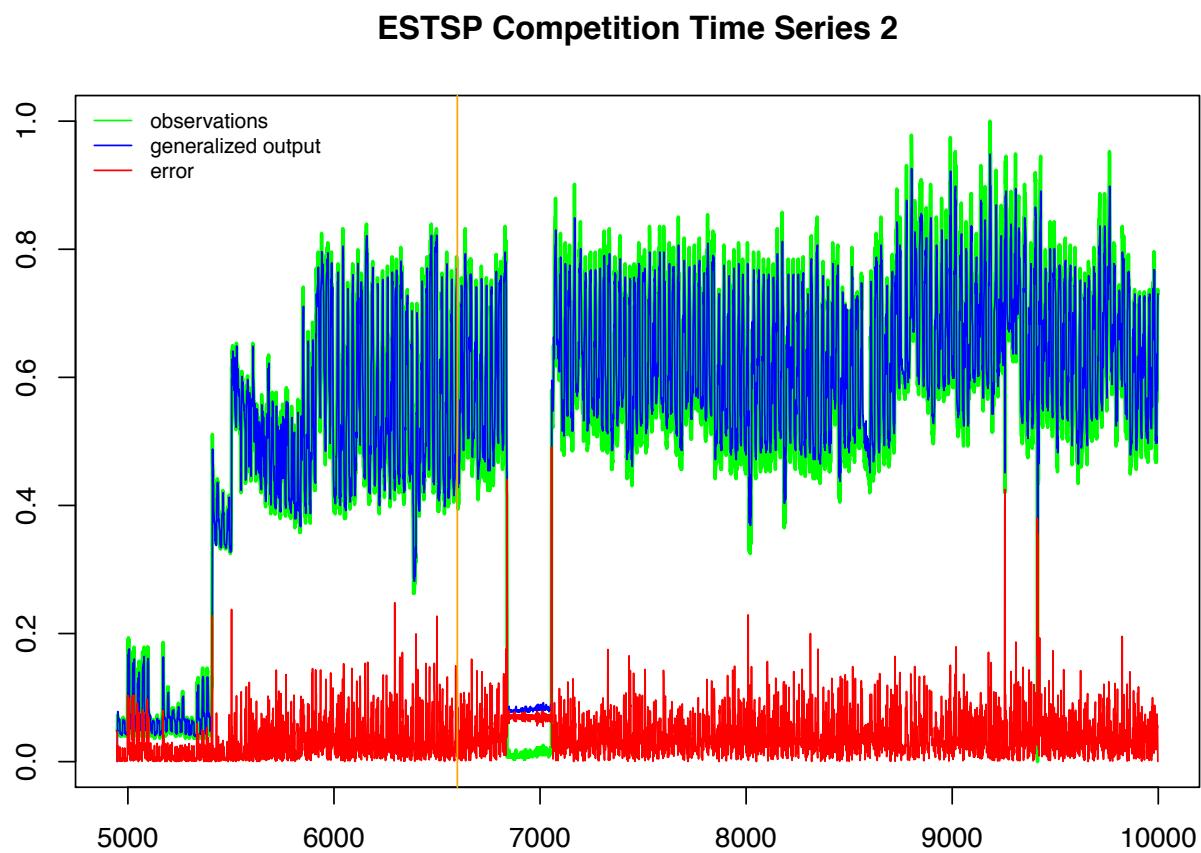


Рис. 1.13. Фазовий портрет часовогого ряду Прогнозування часовогого ряду «ESTSP Competition Time Series»



Прогнозування часовогого ряду «ESTSP Competition Time Series» гібридної каскадної нейро-фаззі мережі з оптимізацією пулу нейронів (жовтобагряна вертикальна лінія позначає закінчення тренувальної частини датасету)

	SMAPE	RMSE
Каскад I		
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.175623	0.062985
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.149139	0.056551
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.146186	0.054414
Нейрон-узагальнювач I каскаду	0.159002	0.055274
Каскад II	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.175602	0.062942
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.149092	0.056487
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.146278	0.054415
Нейрон-узагальнювач II каскаду	0.158996	0.055264
Каскад III	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.175592	0.062933
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.149090	0.056487
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.146270	0.054415
Нейрон-узагальнювач III каскаду	0.158986	0.055262
Каскад IV	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.175605	0.062947
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.149127	0.056487
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.146323	0.054419
Нейрон-узагальнювач IV каскаду	0.159015	0.055268
Нейрон-узагальнювач системи	0.102377	0.030302

Таблиця 1.5

**Результати прогнозування часового ряду «ESTP Competition Time Series» окремих нейронів (зокрема нейронів-узагальнювачів) каскадної нейро-фаззі системи**

### 1.3. Моделювання багатовимірної гібридної каскадної нейро-фаззі мережі, що еволюціонує, з оптимізацією пулу нейронів

В якості тестового датасету для моделювання багатовимірної гібридної каскадної нейро-фаззі мережі з оптимізацією пулу нейронів (що ґрунтуються на багатовимірних нео-фаззі нейронах, як описано у [linkchapter](#)) застосовувався багатовимірний ряд, сгенерованих за домогою диференціальних рівнянь

моделі Лоренца:

$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = -xz + rx - y, \\ \dot{z} = xy - bz. \end{cases} \quad (1.5)$$

У моделі Лоренца присутні три невідомих функції, а також кілька невідомих параметрів [21].

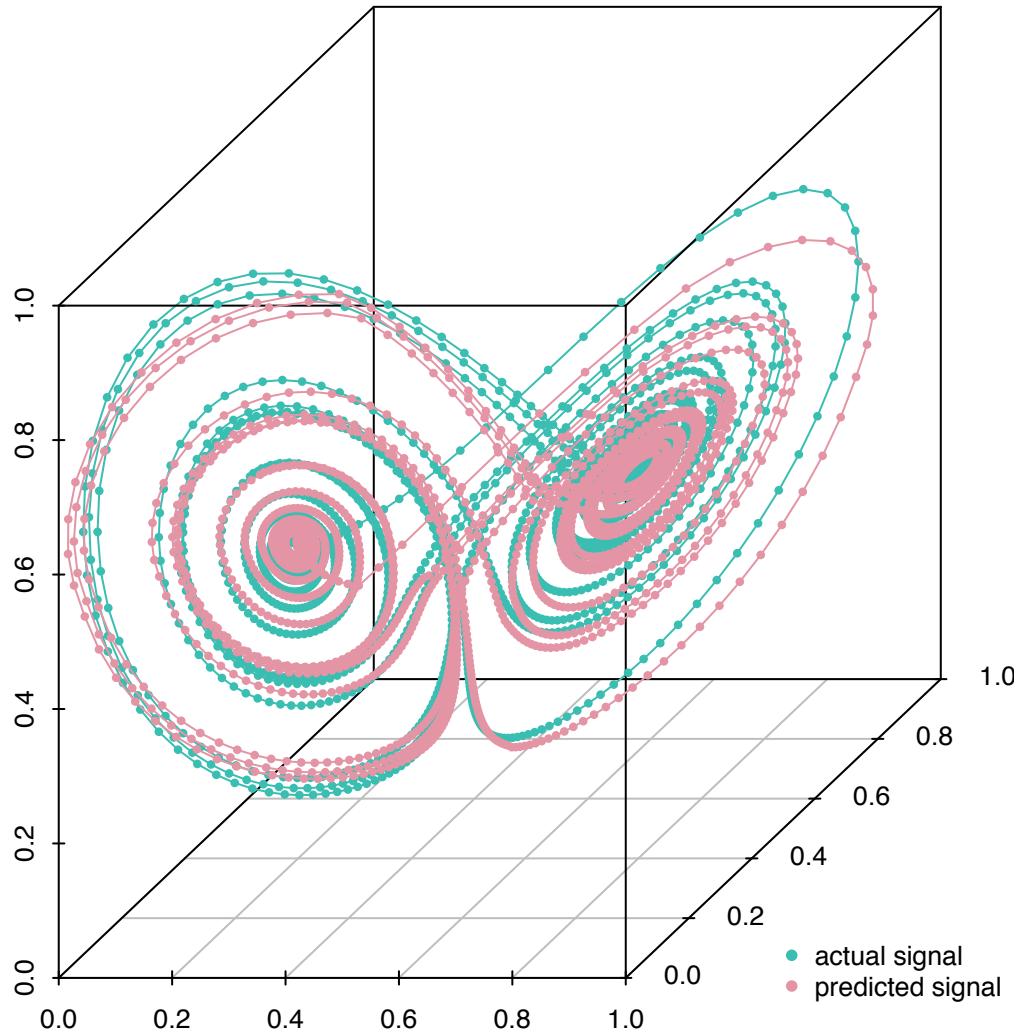


Рис. 1.14. Прогнозування багатовимірного часового ряду гібридною каскадною нейро-фаззі мережею з оптимізацією пулу нейронів

При плавній зміні параметра динамічна система змінює тип свого аттра-

ктора. Рішення системи рівнянь Лоренца 1.5 при значенні параметра  $r$ , що перевищує біфуркаційних, виглядає майже ідентично випадковому процесу. У певному сенсі, атTRACTОР Лоренца є стохастичними автоколиваннями, зо підтримуються у динамічній системі за рахунок зовнішнього джерела.

Каскад I	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.132351	0.64333
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.106885	0.055176
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.118058	0.059517
Нейрон-узагальнювач IV каскаду	0.106823	0.055111
Каскад II	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.132370	0.064339
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.106840	0.55165
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.118059	0.059517
Нейрон-узагальнювач IV каскаду	0.106840	0.55165
Каскад III	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.132325	0.064324
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.106858	0.055171
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.118059	0.059513
Нейрон-узагальнювач IV каскаду	0.106858	0.055171
Каскад IV	SMAPE	RMSE
Нейрон I (3 функції належності, нечітке висновування III порядку)	0.132258	0.0642449
Нейрон II (5 функцій належності, нечітке висновування IV порядку)	0.106805	0.055144
Нейрон III (4 функції належності, нечітке висновування V порядку)	0.118015	0.059424
Нейрон-узагальнювач IV каскаду	0.106805	0.055144
Нейрон-узагальнювач системи	0.106789	0.55105

Таблиця 1.6

### Результати прогнозування багатовимірного часового ряду нейронами (зокрема нейронами-узагальнювачами) МІМО гібридної каскадної нейро-фаззі мережі з оптимізацією пулу нейронів

У фазовому просторі дивний атTRACTОР має топологію деякого клубка траєкторій, в межах якого можна виділити дві області. У кожен момент часу рішення знаходиться в одній з цих областей, причому зміна станів системи в

одну або іншу область є абсолютно непередбачуваною.

АтTRACTор Лоренца демонструє ще одну особливість, притаманну дивним атTRACTорам - чутливість до початкових умов. АтTRACTори, тобто нерухомі точки і граничні цикли, характеризуються тим, що для різних початкових умов сімейства рішень сходилися до одного асимптотичному рішення, тобто різні категорії вийшли з різних точок, які відповідають різним початковим умовам, сходилися при  $t \rightarrow \infty$  в одну точку або близькі криві. Тому поведінку звичайних систем, що мають атTRACTори поблизу нерухомих точок і граничних циклів, на великих часах добре передбачувано. З дивними атTRACTорами все зовсім інакше. Які б близькі початкові умови не вибиралися, при  $t \rightarrow \infty$  рішення будуть розходитися, віддаляючись одне від одного в фазовому просторі. Оскільки в реальних задачах початкові умови відомі з деякою погрешністю, абсолютно неможливо вказати поведінку такого атTRACTора при досить великому  $t$ , тому поведінка систем, що описуються дивними атTRACTорами, є абсолютно непередбачуваною.

Для генерування тестового датасету використовувались такі параметри:

$$r = 28,$$

$$dt = 0.001;$$

По завершенні експерименту маємо систему з чотирьох (рідше – трьох) каскадів з трьома багатовимірними нейронами MNFN та одним нейроном-узагальнювачем у кожному каскаді. Результати роботи системи зображені на рис. 1.14 та більш детально описані у таблиці tableref.

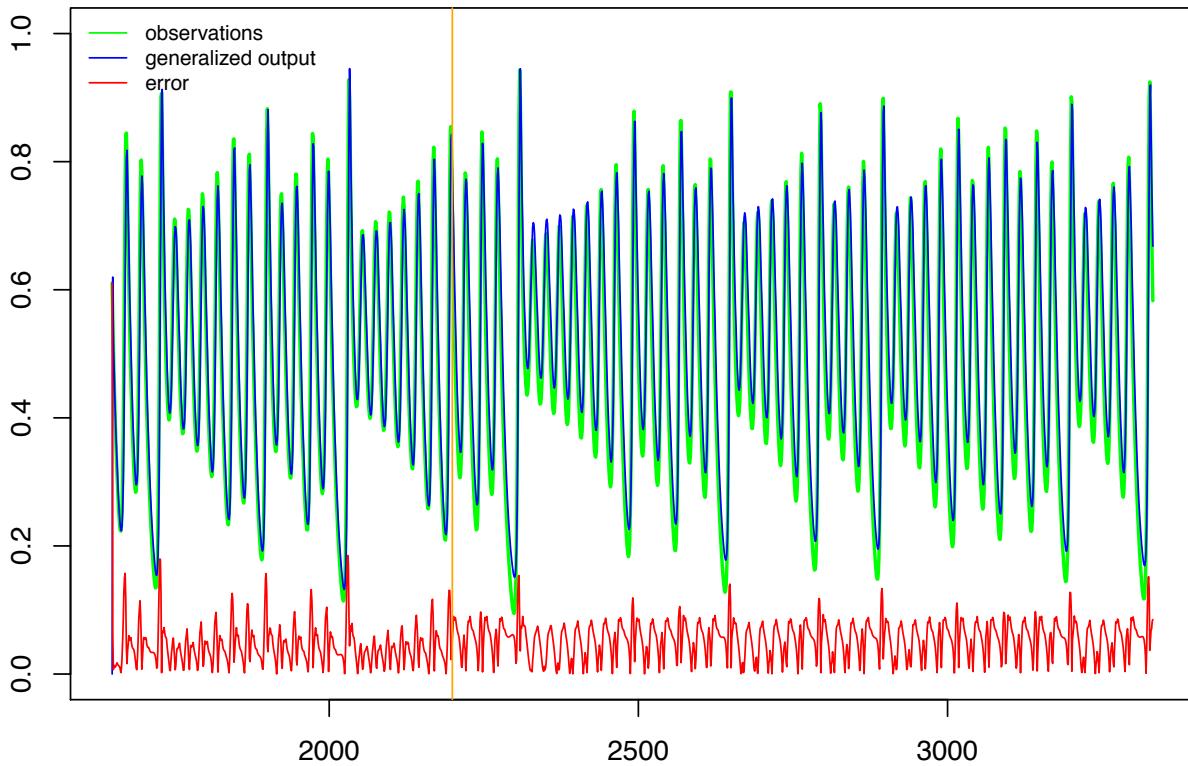


Рис. 1.15. Прогнозування багатовимірного часового ряду гібридною каскадною нейро-фаззі мережею з оптимізацією пулу нейронів

#### 1.4. Моделювання самонавчанної нейро-фаззі системи, що еволюціонує

Одна з основних переваг, притаманних пропонованій самонавчанній нейро-фаззі системі, що еволюціонує, полягає в автоматичному визначенні оптимальної кількості кластерів та значення фаззифікатору на кожному етапі обробляння даних. Першу серію експериментів було проведено на штучно зсинтезованих наборах даних з різним ступенем розмитості та перекриття класів аби дослідити вплив значення параметру фаззіфікації на якість кластерування в режимі реального часу відповідно до обраного критерію дійсності.

Кожен з наборів даних, що їх наведено на рис. 1.16, містить 80 спостережень з 2 ознаками (для очності) у кожному спостереженні. Тестові дані були сгенеровані таким чином, аби у першому наборі класи були чітко розподі-

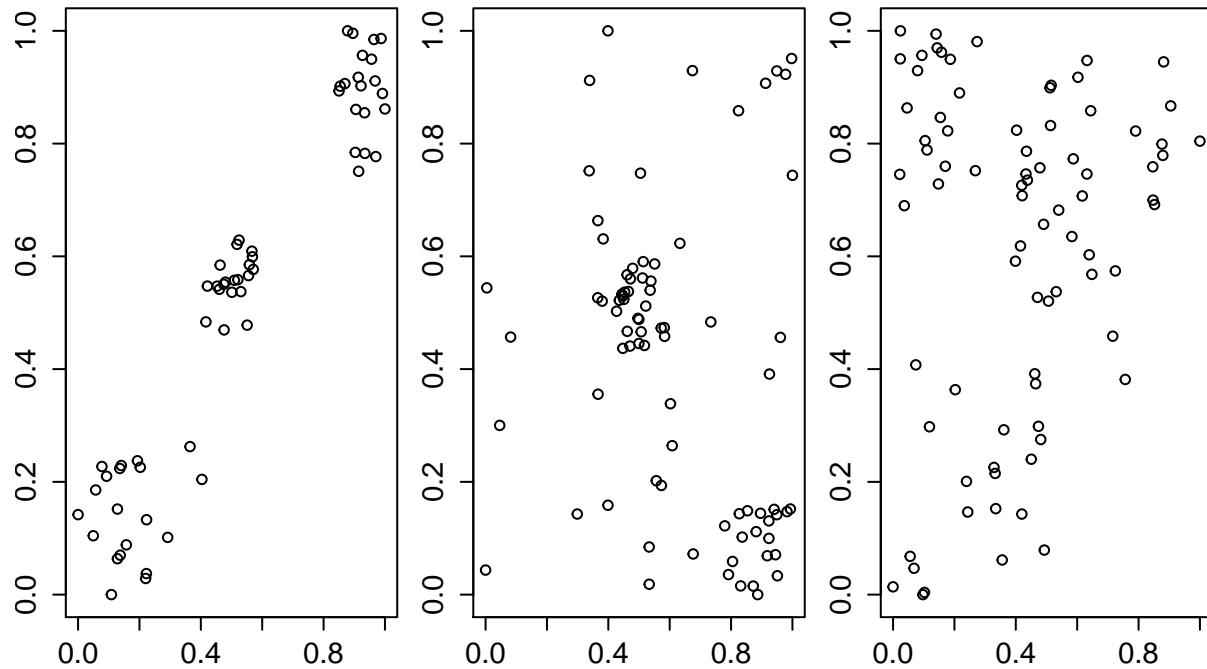


Рис. 1.16. Штучно сгенеровані набори даних

лені (crisp dataset), у другому наборі кластерні границі були дещо розмиті (fuzzy dataset), у третьому випадку класи сильно перетиналися (extra fuzzy dataset). Логічно припустити, що система, яка тестиється, обере менше значення параметру фаззіфікації для первого датасету та більше для останнього, де границі класів спостережень є більш розмитими.

Спостереження надходили до нейро-фаззі мережі у послідовному режимі, вагові коефіцієнти нейронів були проініціалізовані, використовуючи пакетну модифікацію обраного алгоритму кластерування на датасеті з довільних двадцятьох спостережень відповідного набору даних (адже система, як і класичний fuzzy c-means, [чутлива до ініціалзації](#)). Локально оптимальні кількість кластерів та значення параметру фаззіфікації обумовлювалися максимальним середнім значенням рекурентних коефіцієнту розбиття РС (??) та Ксі-Бені індексу (??):  $\max \frac{PC_j^{[m]} + 1 - XB_j^{[m]}}{2}$  (у данному випадку використовувалося від'ємне значення Ксі-Бені індексу  $1 - XB(k)$ , оскільки щоменше  $XB_j^{[m]}$ , то ліпшим є розбиття даних на кластери).

Для первого набора даних (crisp dataset), як і передбачалося, оптималь-

Каскад 1 ( $m = 2$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.91758	0.7446	0.64787	0.59236
Індекс Ксі-Бені	0.052129	0.061034	0.092235	0.1294
Каскад 2 ( $m = 3$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.92643	0.6609	0.50214	0.43305
Індекс Ксі-Бені	0.027232	0.06872	0.17281	0.26914
Каскад 3 ( $m = 4$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.87218	0.5256	0.37605	0.31993
Індекс Ксі-Бені	0.15687	0.4153	0.84699	1.1765
Каскад 4 ( $m = 5$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.73909	0.45445	0.32428	0.27063
Індекс Ксі-Бені	0.12985	0.30637	0.68584	1.0551

Таблиця 1.7

### Індекси валідності (датасет 1)

ним виявився другий каскад ( $m = 3$ ) з трьома кластерами і нейроном-переможцем із найменшим значенням параметру фаззіфікації  $\beta = 2$  (рис. 1.18). Така конфігурація є оптимальною відповідно до обох використовуваних індексів валідності – найменше значення Ксі-Бені індексу  $XB_j^{[m]}$  та найбільший коефіцієнт розбиття  $PC_j^{[m]}$ :

$$PC_1^{[2]} = 0.9009951,$$

$$XB_1^{[2]} = 0.03349166.$$

Лише одне спостереження у цьому датасеті (його позначено багряним квадратом) не належить жодному кластерові з ступінем більшим від 0.6. Індекси валідності нейронів системи наведені у таблиці 5.1.

Для набору даних з середньою вираженістю класів найліпшим виявився нейрон другого каскаду ( $m = 3$ ) і фаззіфікатором  $\beta = 3$  (таблиця 5.2).

Як показано на рис. 1.17, декілька спостережень у центрі (позначені багряними квадратами) можна віднести до 2 кластерів з відносно високим ступінем належності, проте більшість спостережень можна чітко розкластеризувати,

що ілюструється високим значенням коефіцієнту розбиття, та дуже низьким Ксі-Бені індексом:

$$PC_2^{[2]} = 0.9727868,$$

$$XB_2^{[2]} = 0.087474.$$

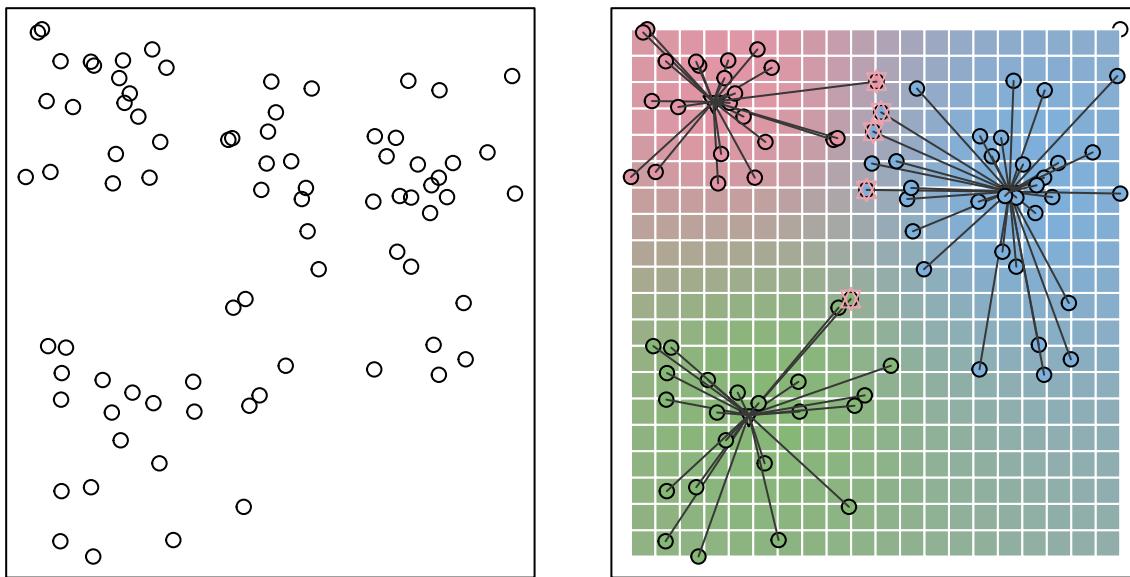


Рис. 1.17. Набір даних з нечіткими межами класів (fuzzy dataset)

Для набору з найменш чіткими межами класів (таблиця 5.3), система обрала нейроном-переможцем вузол третього каскаду ( $m = 4$ ) з високим параметром фаззіфікації  $\beta = 4$ :

$$PC_3^{[3]} = 0.335525,$$

$$XB_3^{[3]} = 0.2128333.$$

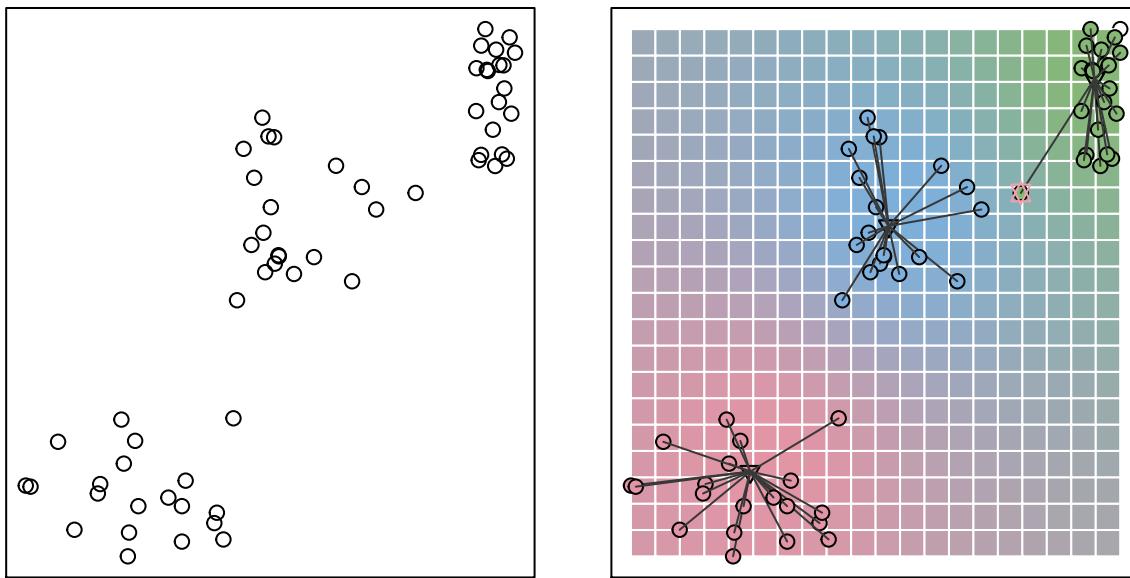


Рис. 1.18. Набір даних з чітко вираженими класами (Crisp dataset)

Каскад 1 ( $m = 2$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.78414	0.58928	0.53853	0.52239
Індекс Kci-Бені	0.16668	0.30834	0.3745	0.38723
Каскад 2 ( $m = 3$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	50084	0.71164	0.97275	0.4191
Індекс Kci-Бені	0.009751	0.031235	0.087474	0.1323
Каскад 3 ( $m = 4$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.91888	0.47532	0.32777	0.28912
Індекс Kci-Бені	0.052563	0.1757	0.27516	0.33766
Каскад 4 ( $m = 5$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.85618	0.34327	0.24778	0.22445
Індекс Kci-Бені	0.048316	0.19887	0.34307	0.41228
Каскад 5 ( $m = 6$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.81295	0.30709	0.21636	0.19214
Індекс Kci-Бені	0.060896	0.19702	0.31393	0.38668

Таблиця 1.8

### Індекси валідності (датасет 2)

На рис. 1.19 спостереження, для яких ступінь належності до будь-якого кластеру не перевищує 0.6, позначені багряними квадратами. Як і очікува-

лося, для цього набору даних кількість таких спостережень значно вища від попередніх датасетів з більш компактними та «чіткими» класами.

Каскад 1 ( $m = 2$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.85094	0.71415	0.61734	0.57085
Індекс Ксі-Бені	0.10584	0.11462	0.13797	0.16101
Каскад 2 ( $m = 3$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.61668	0.42848	0.37779	0.35884
Індекс Ксі-Бені	0.1754	0.20364	0.22364	0.23995
Каскад 3 ( $m = 4$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.33458	0.44082	0.79405	0.29615
Індекс Ксі-Бені	0.20989	0.129	0.051039	0.26282
Каскад 4 ( $m = 5$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.50244	0.33067	0.26029	0.23318
Індекс Ксі-Бені	0.37268	0.61417	0.79695	0.93626
Каскад 5 ( $m = 6$ )	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
Коефіцієнт розбиття	0.53279	0.29731	0.22648	0.19858
Індекс Ксі-Бені	0.27407	0.47298	0.60569	0.70716

Таблиця 1.9

### Індекси валідності (датасет 3)

Для очності у всіх наведених рисунках кольором позначені не тільки розкластеровані спостереження і центри кластерів, а й задній план (фон) малюнків, що дозволяє візуально визначити, до якого кластеру система віднесла б нові спостереження. Не дивно, що, тоді як для перших двох датасетів важко визначити домінуючий колір, оскільки кластери їх спостережень більш менш компактні та явно виражені, для останнього набору даних домінуючий колір – сірий, сформований кольорами усіх кластерів, що ілюструє великий ступінь перекриття класів і, відповідно, високе значення оптимального параметру фаззіфікації  $\beta$ , що обрала система.

Ця низка експериментів проілюструвала як важливо вірно визначати параметр фаззіфікації, оптимальне значення якого у випадку обробляння даних у послідовному режимі з високою вирогідністю змінюється у часі, а саме здатність визначати оптимальне значення цього параметру в онлайн режимі є

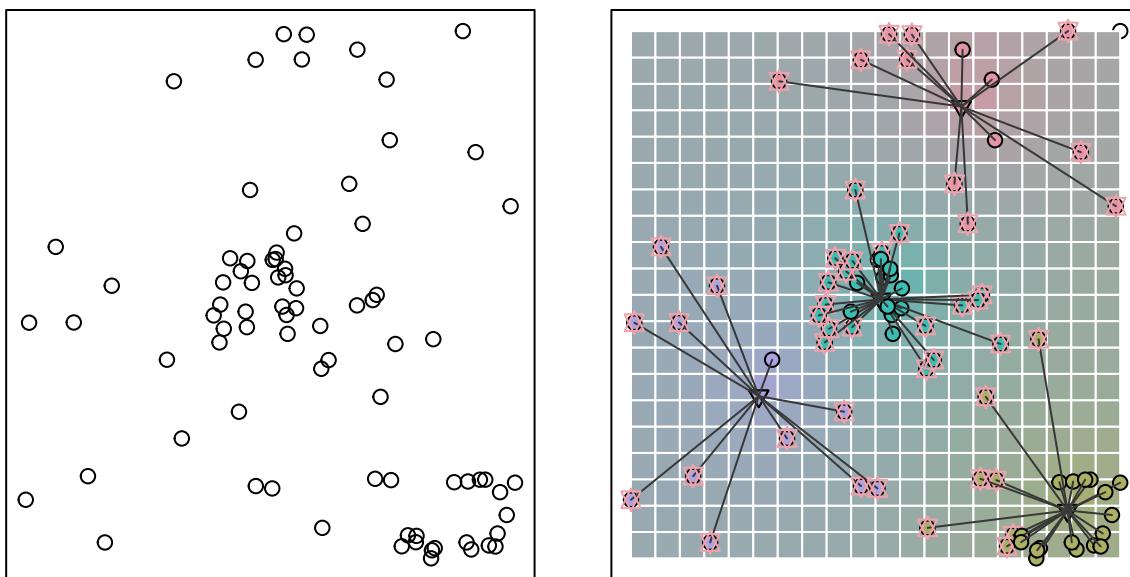


Рис. 1.19. Набір даних з класами, що перетинаються (extra fuzzy dataset)

відмінною особливістю запропонованої самонавчанної нейро-системи.

Наступна низка експериментів була проведена на наборі даних «Іриси Фішера» (Fisher's Iris data set).

Це багатовимірний датасет для задачі класифікації, на прикладі якого англійський статистик та біолог Рональд Фішер в 1936 році продемонстрував роботу розробленого ним методу дискримінантного аналізу. Іноді його також називають «Ірисами Андерсона» (через те, що дані були зібрані американським ботаніком Едгаром Андерсоном). Цей набір даних став класичним і часто використовується в літературі для ілюстрації роботи різних статистичних алгоритмів.

Проте цей датасет рідко використовується у кластерному аналізі, адже межі класів «*Virginica*» та «*Versicolor*» не можна чітко визначити, ґрунтуючись на даних, що їх використовував Фішер (що легко продемонструвати за допомогою ієрархічного кластерування, рис. 1.20). Саме цим і цікавий для нас цей набір даних: коли класичні методи чіткого кластерного аналізу не справляються з задачею, може стати у нагоді система, що реалізує нечітке кластерування зі змінним параметром фаззифікації та кількістю кластерів. Для більшості методів кластерного аналізу, зокрема для методу нечітких се-

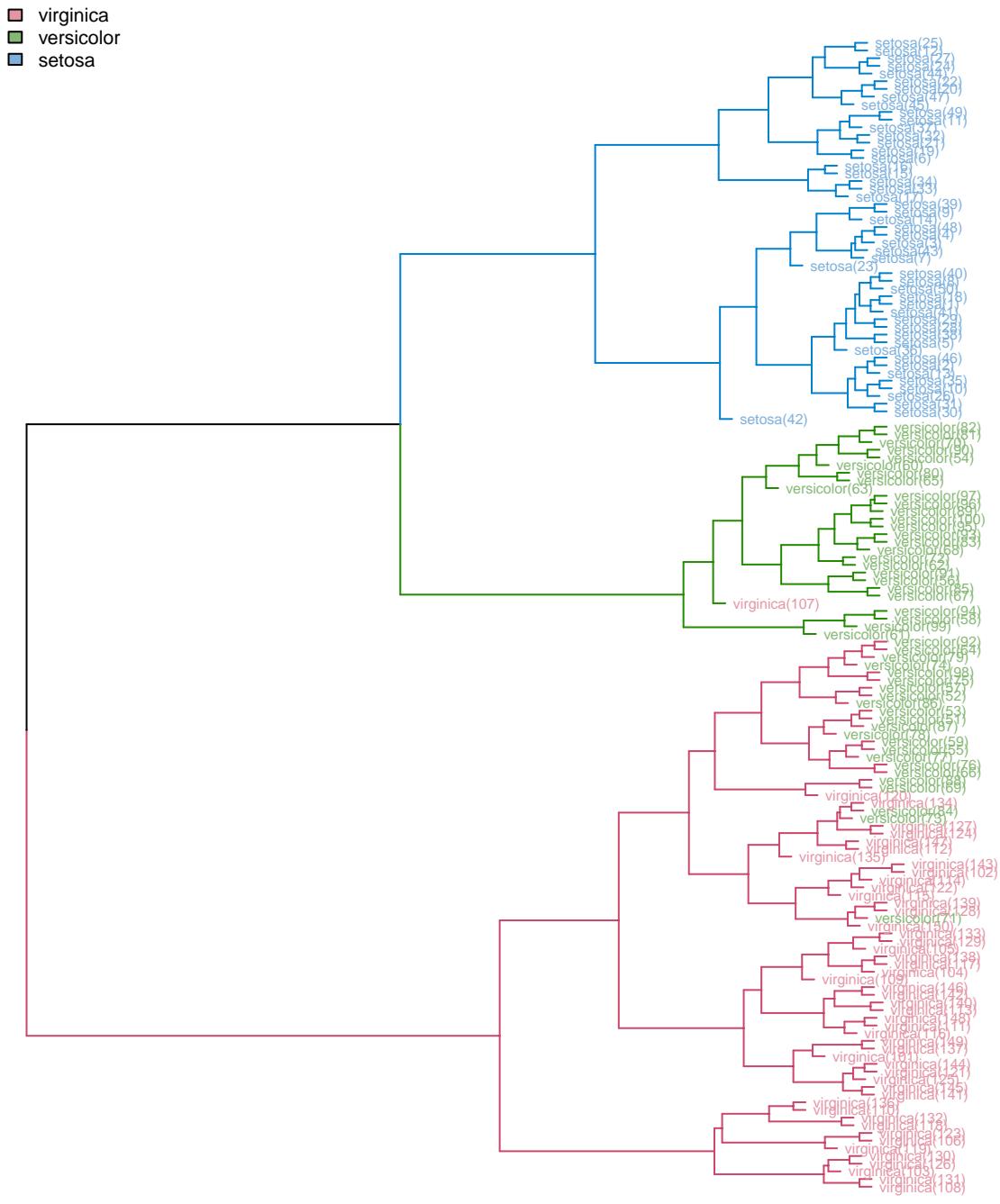


Рис. 1.20. Ієрархічне класерування датасету «Іриси Фішера»

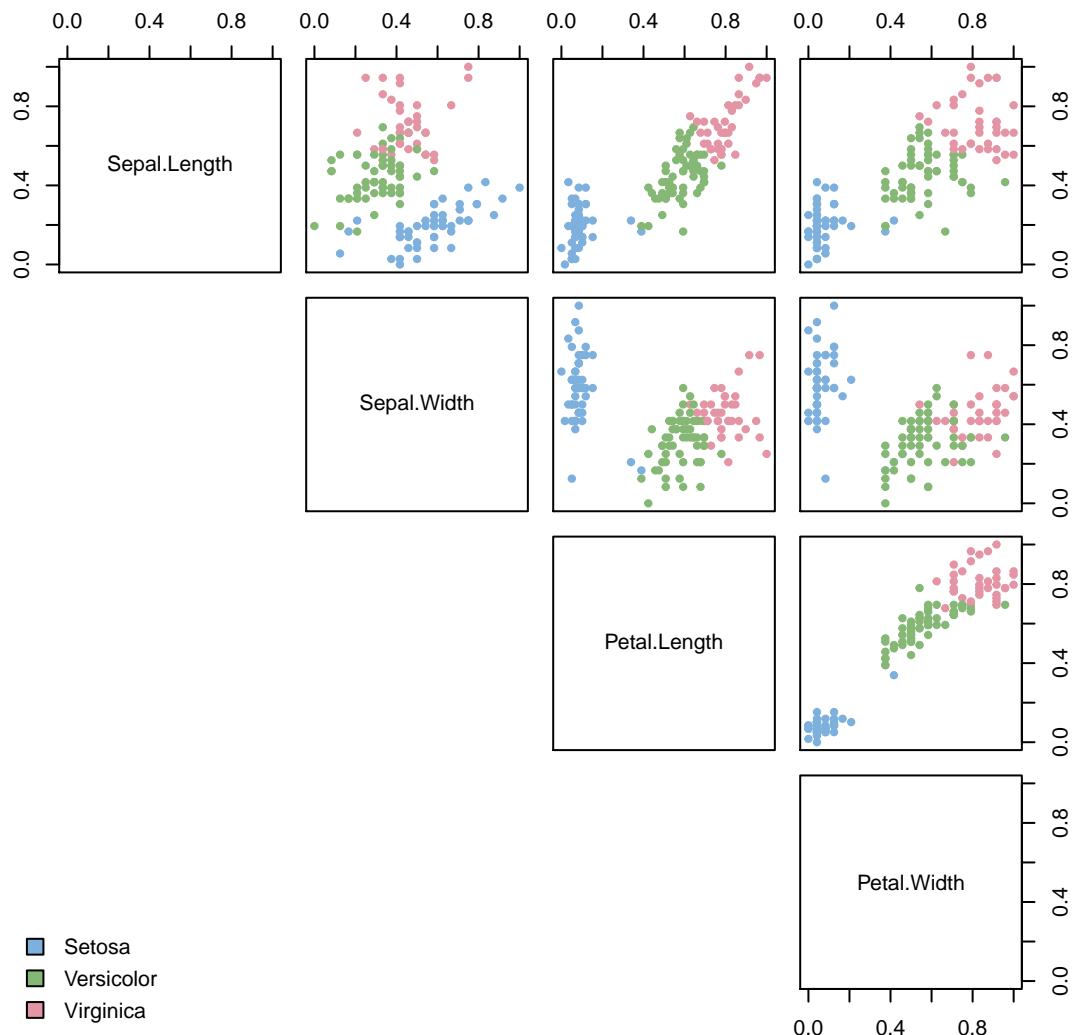


Рис. 1.21. Розкластерований датасет «Іриси Фішера» при  $m = 3$ ,  $\beta = 2$  (Точність кластерування – 96%)

редніх (fuzzy c-means), необхідно заздалегідь задати кількість кластерів, і очевидним рішенням є прийняття  $m = 3$ , адже маємо три класи: Iris Virginica, Iris Versicolor та Iris Setosa (рис. 1.21).

	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.8313073	0.8741245	0.8709475	0.8888124
min	0.7859722	0.7533766	0.6615745	0.7656498
max	0.8534013	0.9166667	0.935051	0.9604701

Таблиця 1.10

### Точність кластерування при $m = 3$

Точність кластерування за допомогою методу нечітких середніх за таких умов ( $m = 3$ ,  $\beta = 2$ ) рідко перевищує 83% (таблиця 5.4). (Оскільки для обраного датасету існують мітки з вірною класифікацією, ефективність кластеризації вимірювалася у відсотках точності щодо еталонного розбиття після дефаззіфікації.) Проте, якщо не обмежувати пропоновану систему у кількості кластерів (система ініціалізується інтервалом допустимих значень  $m$  (кількість кластерів) та параметру фаззифікації  $\beta$ ), велими цікавими є результати кластерування нейронів кожного з каскадів.

У таблиці 5.5 наведена точність розбиття даних, коли  $m \gg 3$  кластерів відповідно. Варто зазначити, що нейрони у пулі кожного каскаду реалізують метод нечітких середніх зі змінним значення фазифікатору, а отже є чутливими до довільно ініціалізованих цетрах кластерів, тому у таблицях наведені середня, мінімальна та максимальна точності кластерування (після дефаззіфікації).

$m = 7$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.8972948	0.9150268	0.9242503	0.9178207
min	0.8536056	0.8461905	0.8723182	0.8600289
max	0.9621849	0.9736172	0.9810146	0.9663462
$m = 8$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9065560	0.9296311	0.9243606	0.9248976
min	0.8217056	0.8562179	0.8577202	0.8590278
max	0.9474588	0.9789402	0.9848214	0.9747899
$m = 9$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9258154	0.9282887	0.9308971	0.9229753
min	0.8689921	0.8270525	0.8684641	0.8556390
max	0.9849170	0.9806397	0.9664112	0.9748284
$m = 10$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9213191	0.9285106	0.9332528	0.9282907
min	0.8663370	0.8722271	0.8652272	0.8766667
max	0.9663420	0.9838095	0.9723656	0.9756335
$m = 11$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9295315	0.9408977	0.9317242	0.9295800
min	0.8520268	0.8964924	0.8890781	0.8788656
max	0.9716166	0.9848485	0.9704892	0.9798627
$m = 12$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9349407	0.9433244	0.9337934	0.9306632
min	0.8815133	0.8949802	0.8798160	0.8486111
max	0.9795274	0.9783497	0.9630952	0.9772727
$m = 13$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9420998	0.9398127	0.9375204	0.9357708
min	0.8823175	0.8614025	0.8882479	0.8828348
max	0.9807518	0.9788034	0.9753452	0.9748873

Таблиця 1.11

Точність кластерування для  $m \in [7, 13]$ ,  $\beta \in [2, 5]$

На рис 1.22 зображено залежність точності кластерування від кількості кластерів. Цікаво, що при, здавалося б, очевидному рішенні обрати кількість кластерів рівною трьом, отримуємо чиненайгіршу точність кластерування (при  $\beta = 2$ ) після дефаззіфікації щодо еталонного розбиття (Для порівняння

на рис. 1.23 та рис. 1.24 наведені розбиття, що їх запропонували нейрони-переможці деяких каскадів, де  $m \gg 3$ ).

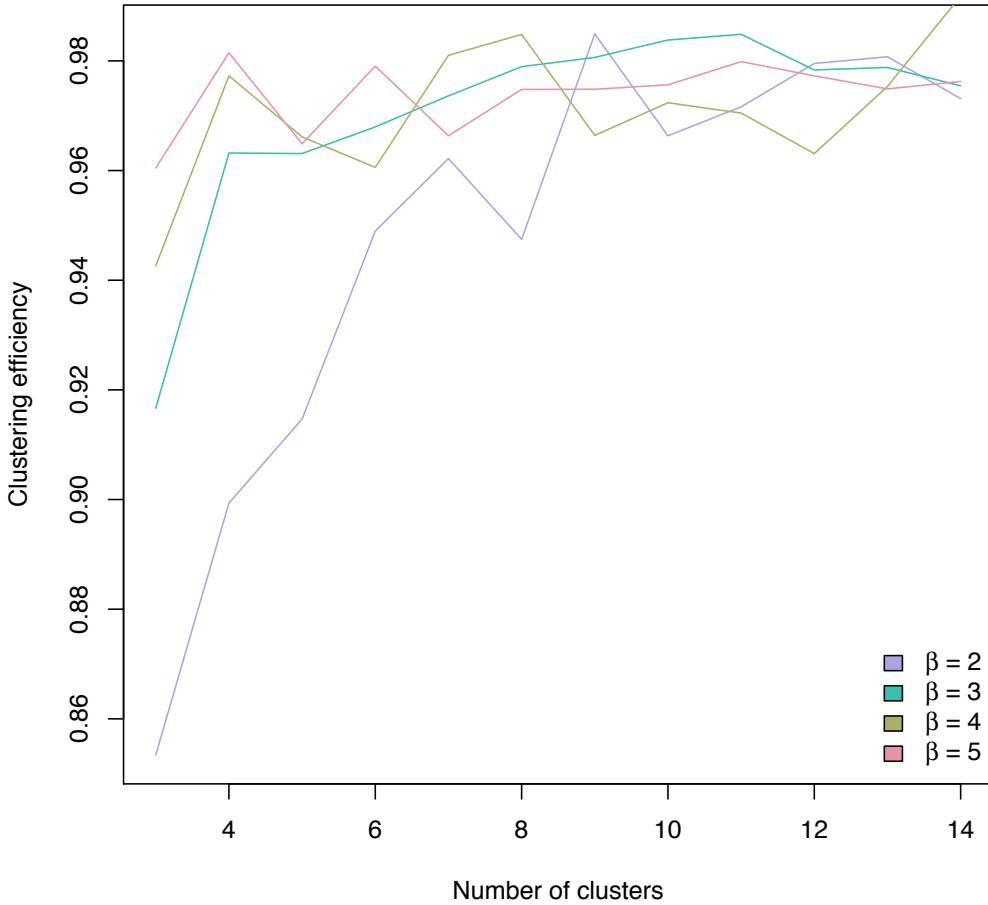


Рис. 1.22. Точність кластерування від кількості кластерів та параметру фаз-зіфікації

Цьому легко знайти пояснення, адже метод нечітких  $k$ -середніх (а саме цей метод у цьому експерименті реалізовують вузли пулів кожного каскаду) добре розпізнає кластери лише гіперсферичної форми. Проте кластер довільної (негіперсферичної) форми, можна розбити на декілька гіперсферичних підкластерів, що й відбувається у каскадах, де  $m > 3$ , що пропонують розбиття на дрібні кластери. На рисунках 1.25 та 1.26 наведені розбиття деяких каскадів, де кількість кластерів більша від кількості класів еталонної вибірки; тут можна побачити, що декілька кластерів, що після дефазифікації будуть віднесені до одного класу, наприклад, Iris Virginica розташовані поруч один з одним, тобто є складовими більшого кластеру негіперсферичної форми.

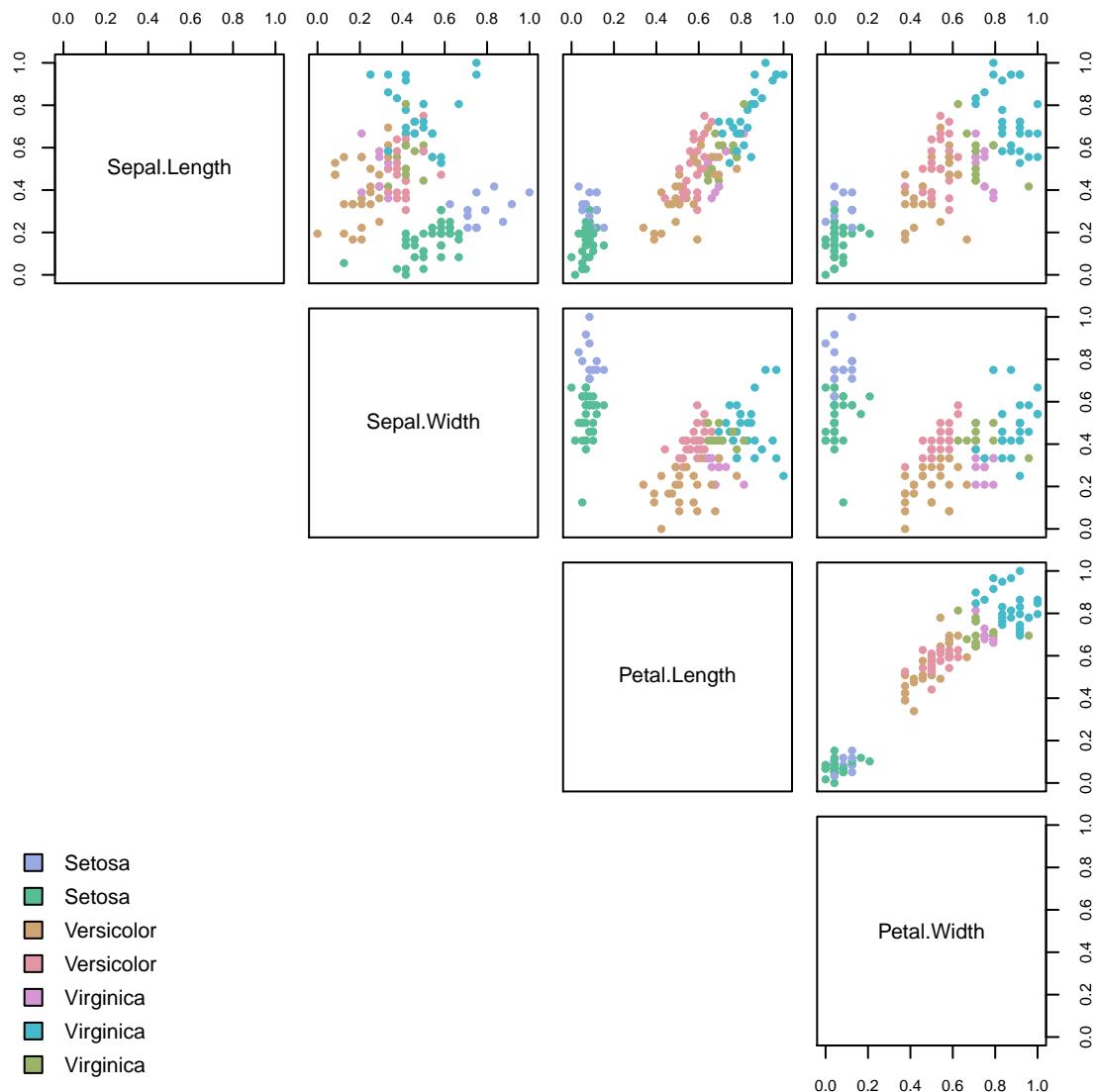


Рис. 1.23. Розкластерований датасет «Іриси Фішера» при  $m = 7$ ,  $\beta = 5$  (Точність кластерування  $\approx 93\%$ )

	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
avg	0.9369168	0.9448829	0.9383179	0.9403416
min	0.8847819	0.8953380	0.8787879	0.9069805
max	0.9731262	0.9754579	0.9918301	0.9762515

Таблиця 1.12

Точність кластерування при  $m = 14$

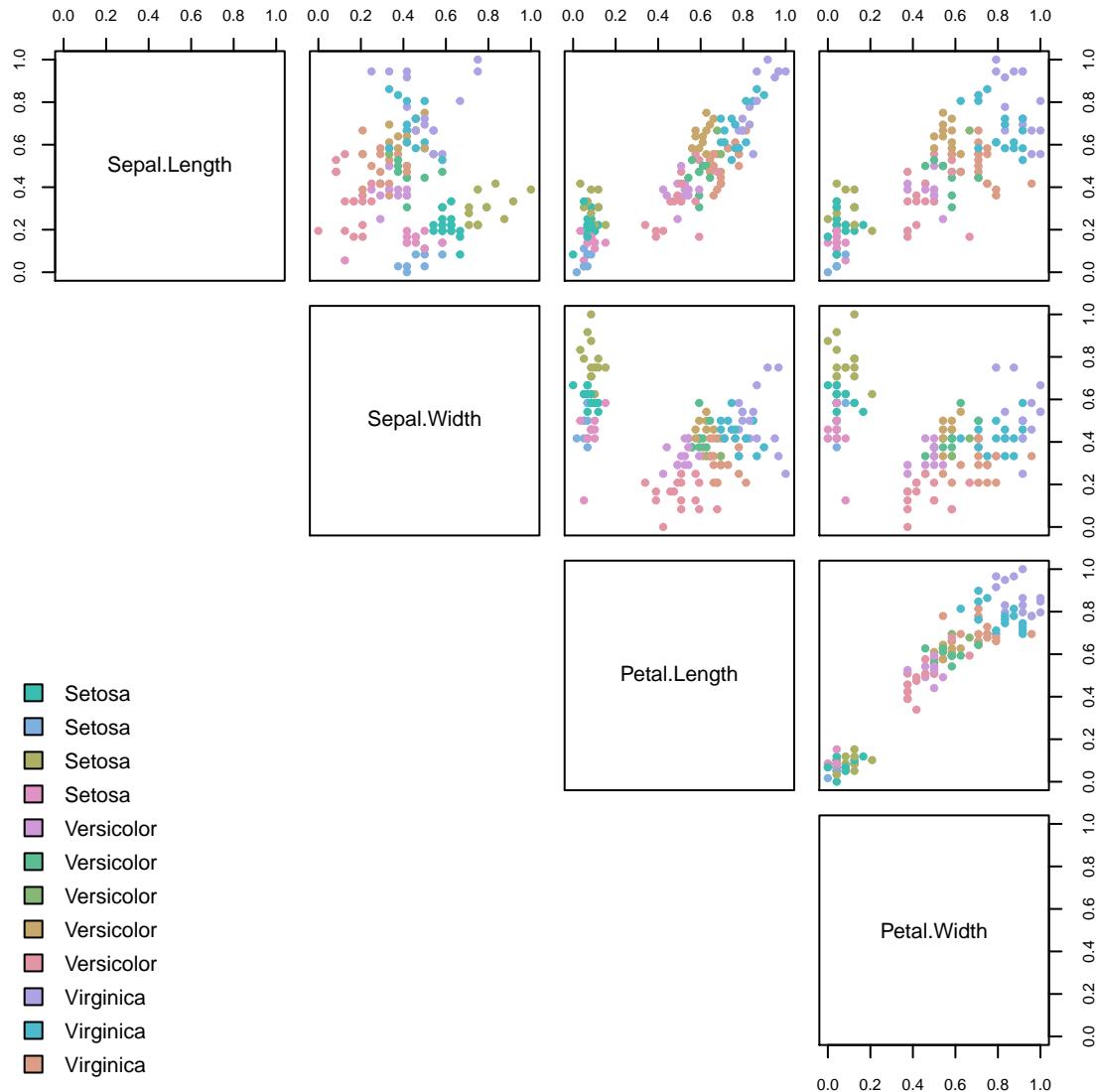


Рис. 1.24. Розкластерований датасет «Іриси Фішера» при  $m = 12$ ,  $\beta = 4$  (Точність кластерування  $\approx 96\%$ )

Таким чином, видається доречним, навіть у випадку, коли відоме еталонне розбиття датасету, дозволити системі обрати кінцеву кількість кластерів самостійно, особливо у випадку, коли вузли системи реалізують одинаковий метод кластерування.

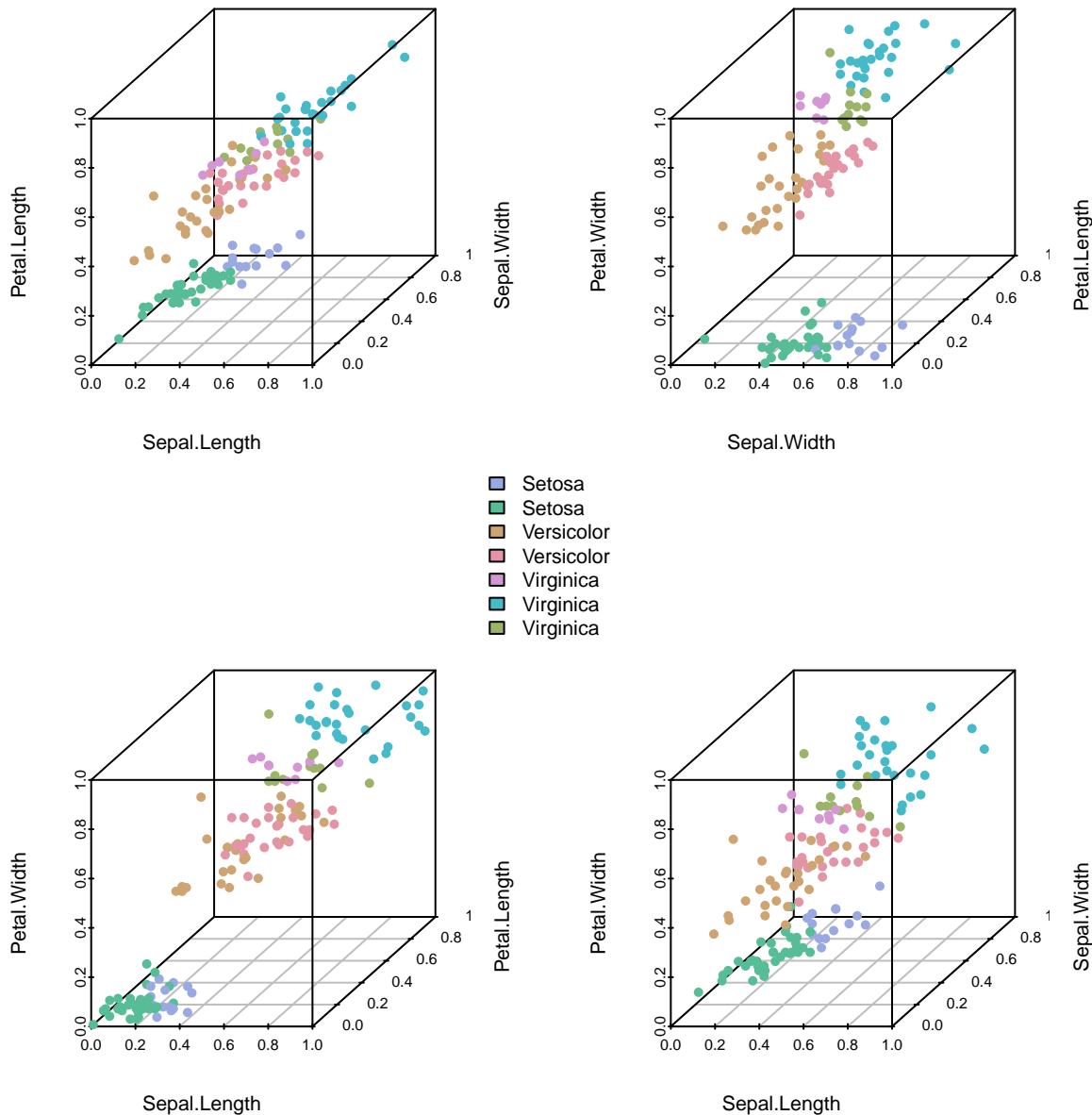


Рис. 1.25. Розкластерований датасет «Іриси Фішера» при  $m = 7$ ,  $\beta = 5$  (Точність кластерування  $\approx 93\%$ )

Варто зауважити, що у цьому випадку для визначення локально опти-

мального розбиття доцільно використовувати модифіковані індекси валідності, чи такі, що не залежать від відстані центрів кластерів, наприклад ті, що ґрунтуються на щільності (density-based).

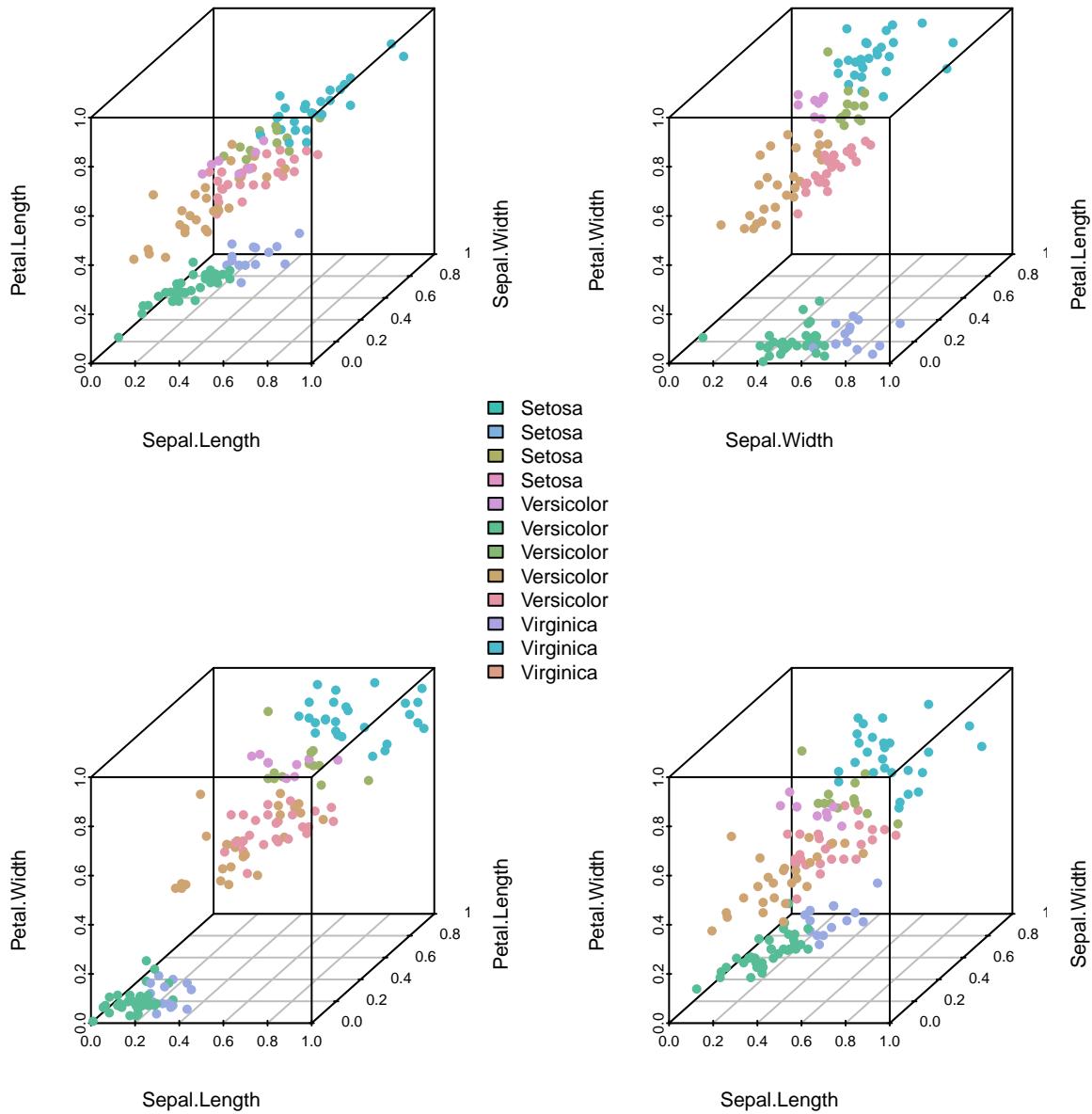


Рис. 1.26. Розкластерований датасет «Іриси Фішера» при  $m = 12$ ,  $\beta = 4$  (Точність кластерування  $\approx 96\%$ )

Наступну серію експериментів було проведено на датасеті «Знання студентів про електричні машини постійного струму».

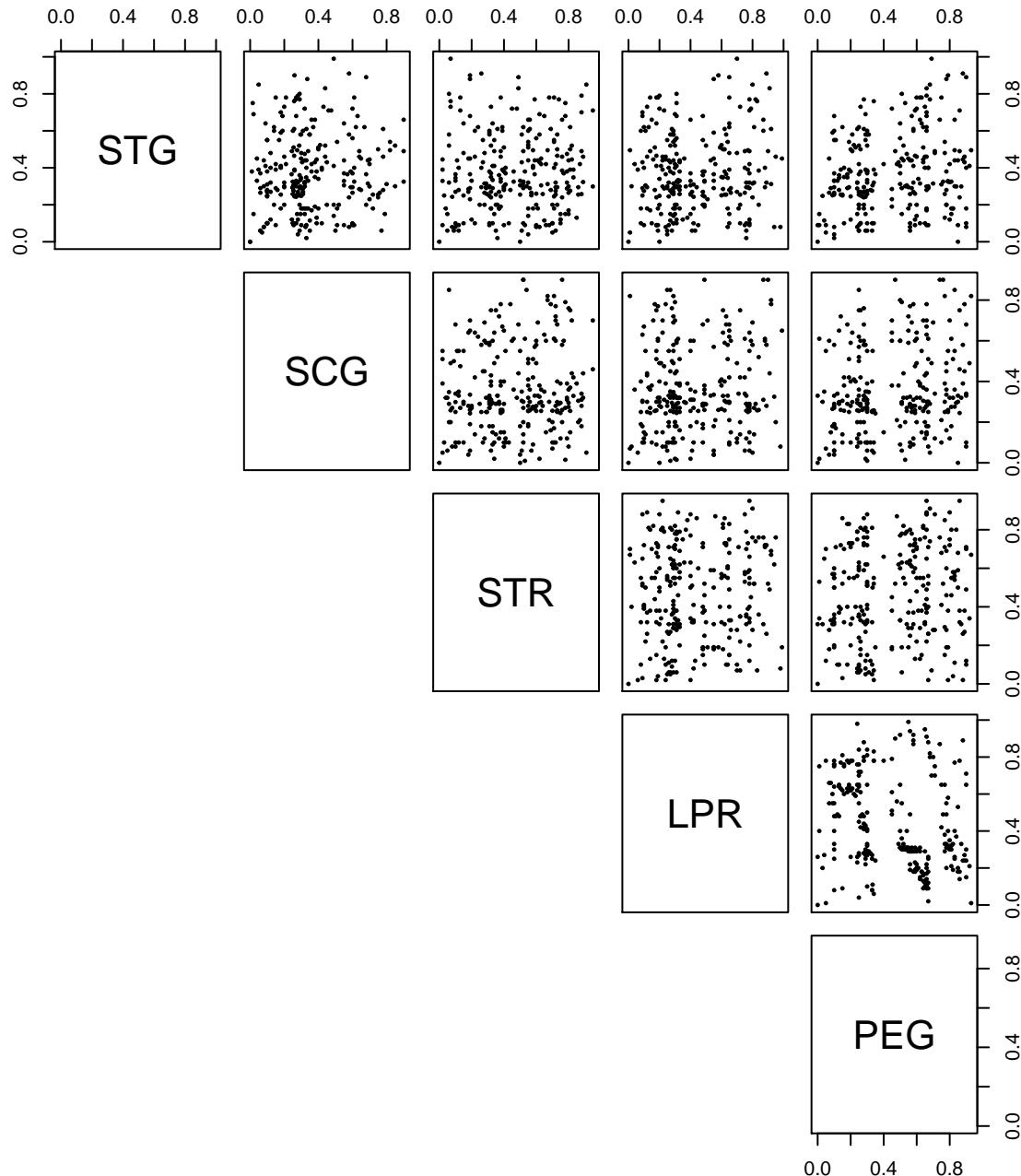


Рис. 1.27. Датасет «Знання студентів про електричні машини постійного струму»

Цей датасет було додано до UCI репозиторію у 2013 році, він містить 403 паттерни, кожен з п'ятьма атрибутами:

1. STG: кількість часу, що його витратив(витратила) студент(ка) на навчання цільового матеріалу,
2. SCG: Кількість повторювань навчання цільового матеріалу студен-

том(студенткою),

3. STR: Кількість часу, що його використав(використала) студент(ка) на навчання матеріалу, пов'язаного з цільовим матеріалом,
4. LPR: Оцінка, що її отримав(отримала) студент(ка) на іспиті з предмету, пов'язаного з цільовим предметом,
5. PEG: Оцінка, що її отримав(отримала) студент(ка) на іспиті з цільового предмету,

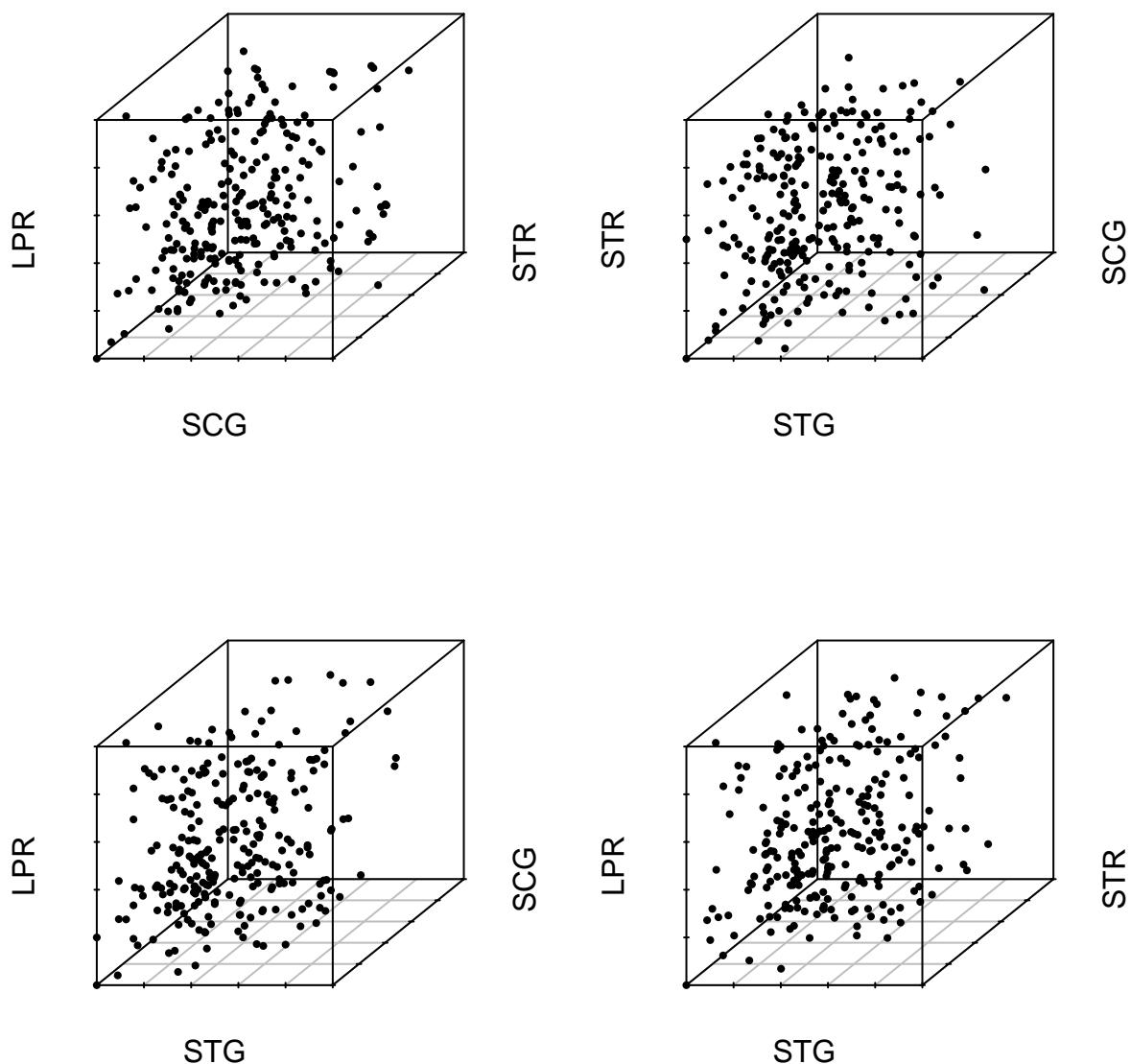


Рис. 1.28. Датасет «Знання студентів про електричні машини постійного струму»

Попарні графіки атрибутів наведено на рис. 1.27 та у тримірному просторі на рис. 1.28.

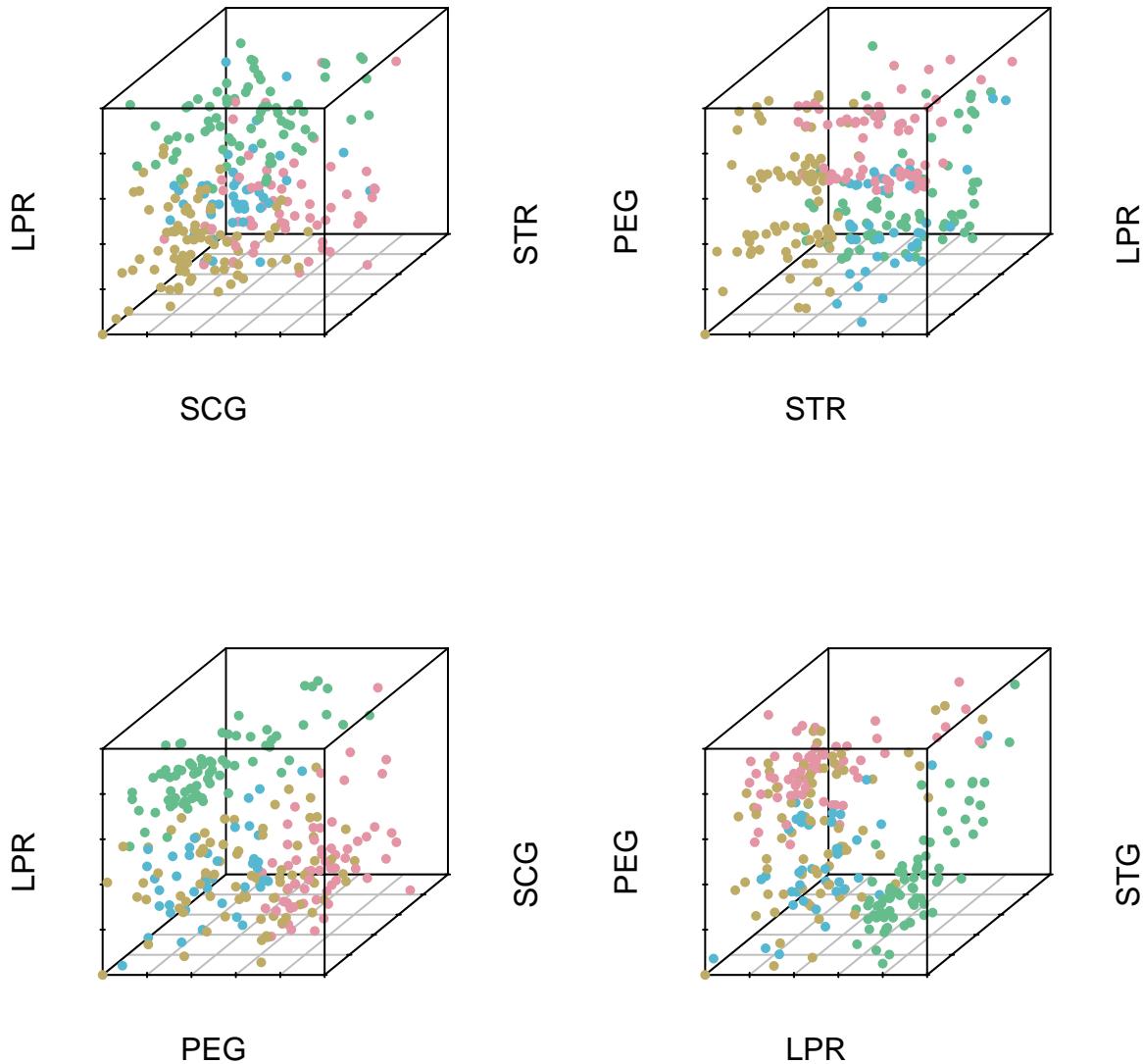


Рис. 1.29. Датасет «Знання студентів про електричні машини постійного струму»

Для цього експерименту нейрони-узагальнювачі керувалися рекурентним Ксі-Бені Індексом при визначенні локально-оптимального нейрона (з найлішшим параметром фаззіфікації) та каскаду (з оптимальною кількістю кластерів).

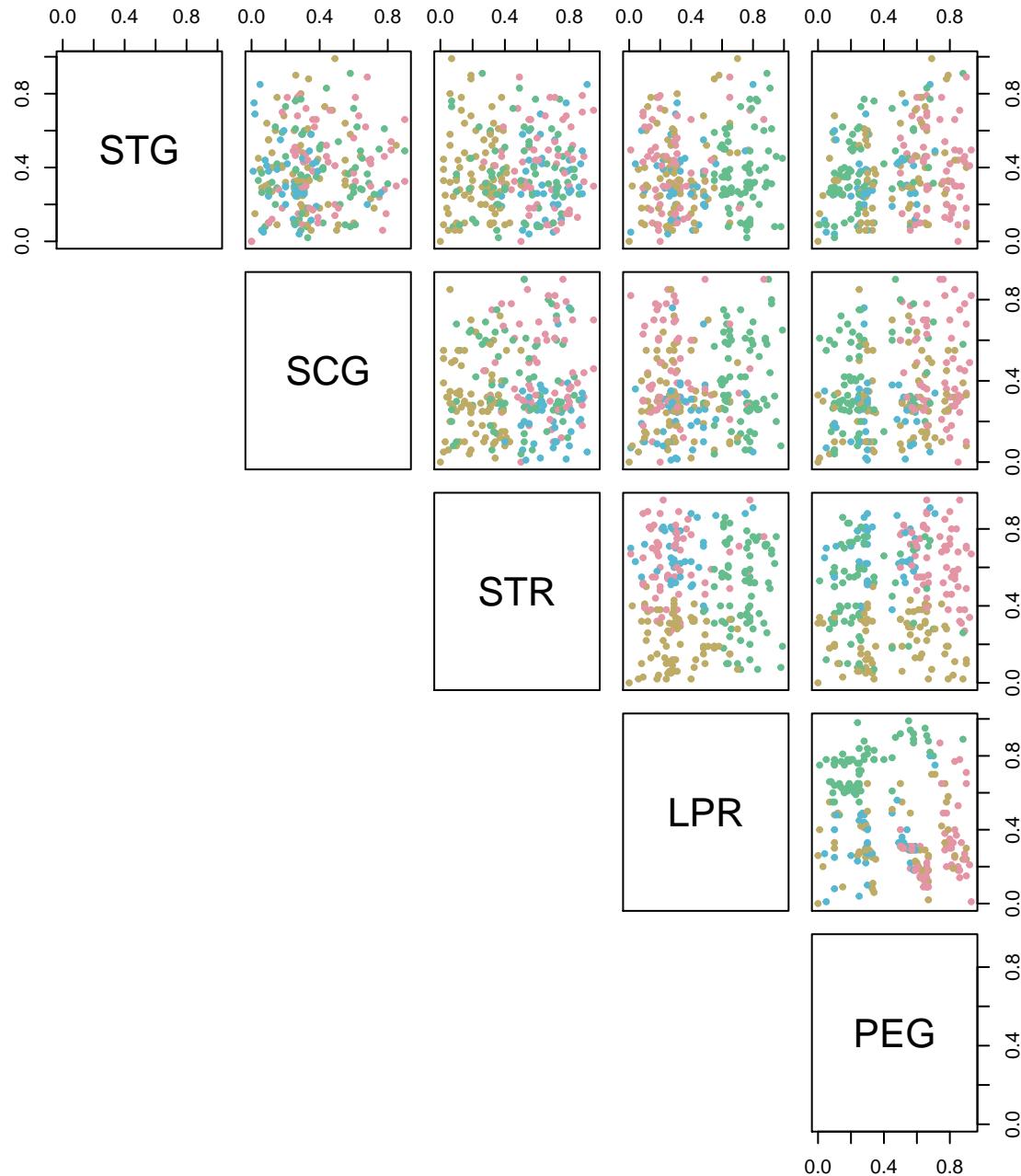


Рис. 1.30. Датасет «Знання студентів про електричні машини постійного струму»

Оптимальне розбиття, що його наведено на рис.1.29 та на рис.1.30, надав другий нейрон третього каскаду з  $m = 4$ ,  $\beta = 2$  з коефіцієнтом Ксі-Бені 0.38155.

## 1.5. Розв'язування практичних задач за допомогою розробленої самонавчанної гібридної каскадної системи, що еволюціонує

Проблема здорового харчування - одна з найактуальніших у наші дні. Повноцінне харчування передбачає споживання достатньої кількості білків, жирів, вуглеводів, вітамінів, макро- і мікроелементів для нормального функціонування організму в цілому. Багато хвороб шлунково-кишкового тракту «молодіють» - це гастрити, виразкова хвороба шлунка і різні порушення обміну речовин. Фізичне здоров'я, стан імунітету, довголіття, психічна гармонія - все це безпосередньо пов'язано з проблемою здорового харчування людини. Для студентів проблема харчування стоїть особливо гостро. У зв'язку з браком часу у студентів немає можливості дотримуватися правильного режиму прийомів їжі в кількості 3-4 разів. Також характерний в основному сидячий спосіб життя - гіподинамія. У поєднанні з незбалансованим раціоном харчування це згубно впливає на організм і його стан. Звісно, вирішення проблеми здорового харчування потребує комплексного підходу, проте інформованість - невід'ємна складова правильного підбору раціону здорового харчування. Насьогодні нескладно знайти інформацію щодо рекомендованої денної кількості калорій, білків, жирів та вуглеводів, проте важко дати оцінку конкретному прийому їжі, наприклад, придбаному у їdalyni, де немає етикеток з такою інформацією. Мобільний додаток «Spoon app» може стати у нагоді, коли користувач прагне бути проінформованим щодо поживності конкретної трапези, роблячи аналіз світлинни тарілки з їжею. Вхідними даними мобільного додатку є світлина, що її користувач має зробити таким чином, аби тарілка знаходилася у центрі, а також тип тарілки (звичайна, глибока, дуже глибока) аби на виході додаток мав обґрунтовану кількість калорій та поживність порції, що була зображена на світлині.

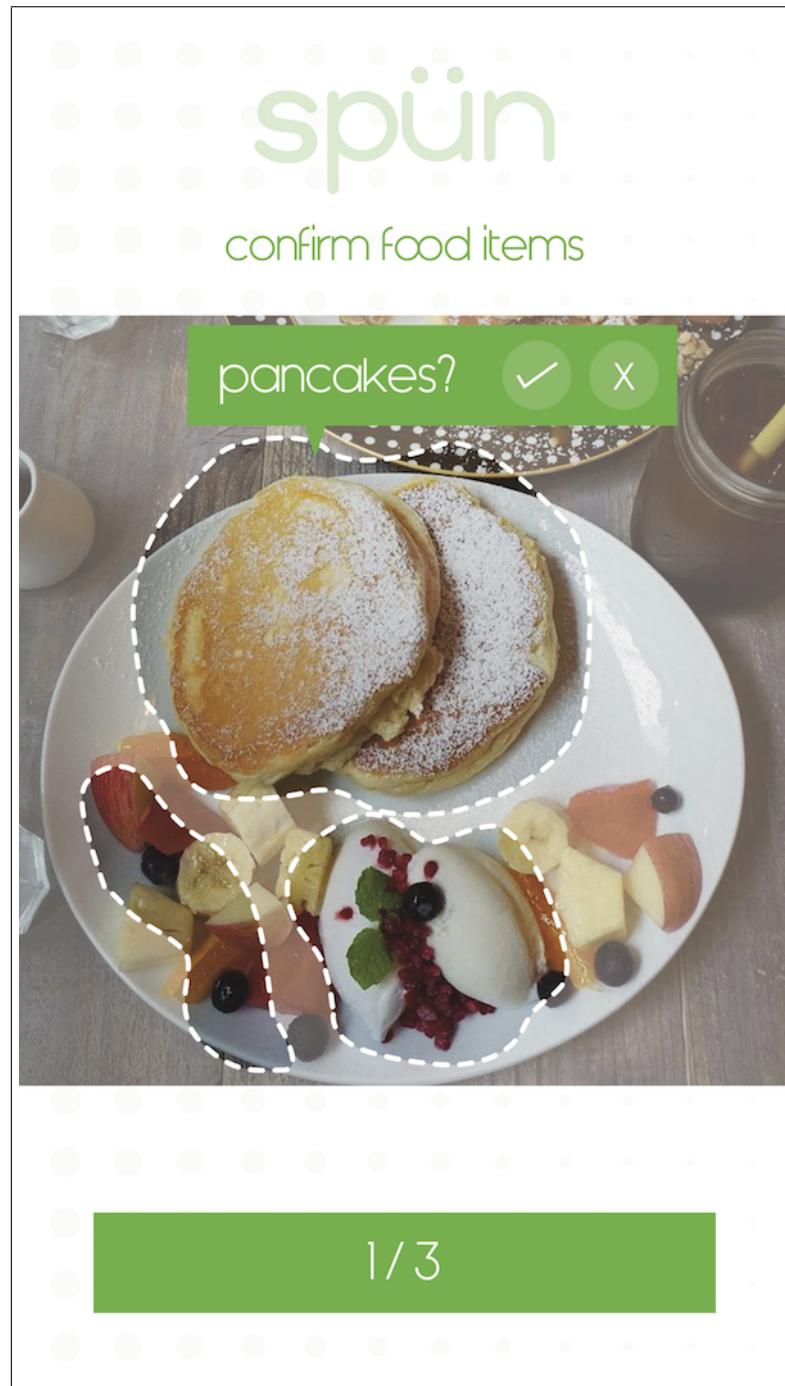


Рис. 1.31. Другий етап аналізу світлини з прийомом їжі мобільним додатком Spoon App (навчання з підкріпленням)

Аналіз світлини можна розбити на декілька етапів:

1. кластерування даних зображених на світлині (відбувається на стороні клієнту)
2. ідентифікація окремих складових прийому їжі: класифікація кожного зображення, після розбиття світлини на кластери на першому етапі

(відбувається на серверній стороні), визначення типу продукту за допомогою бази даних, що знаходиться на сервері, та подальше визначенні кількості калорій, співвідношення білкі, жирів та вуглеводів.

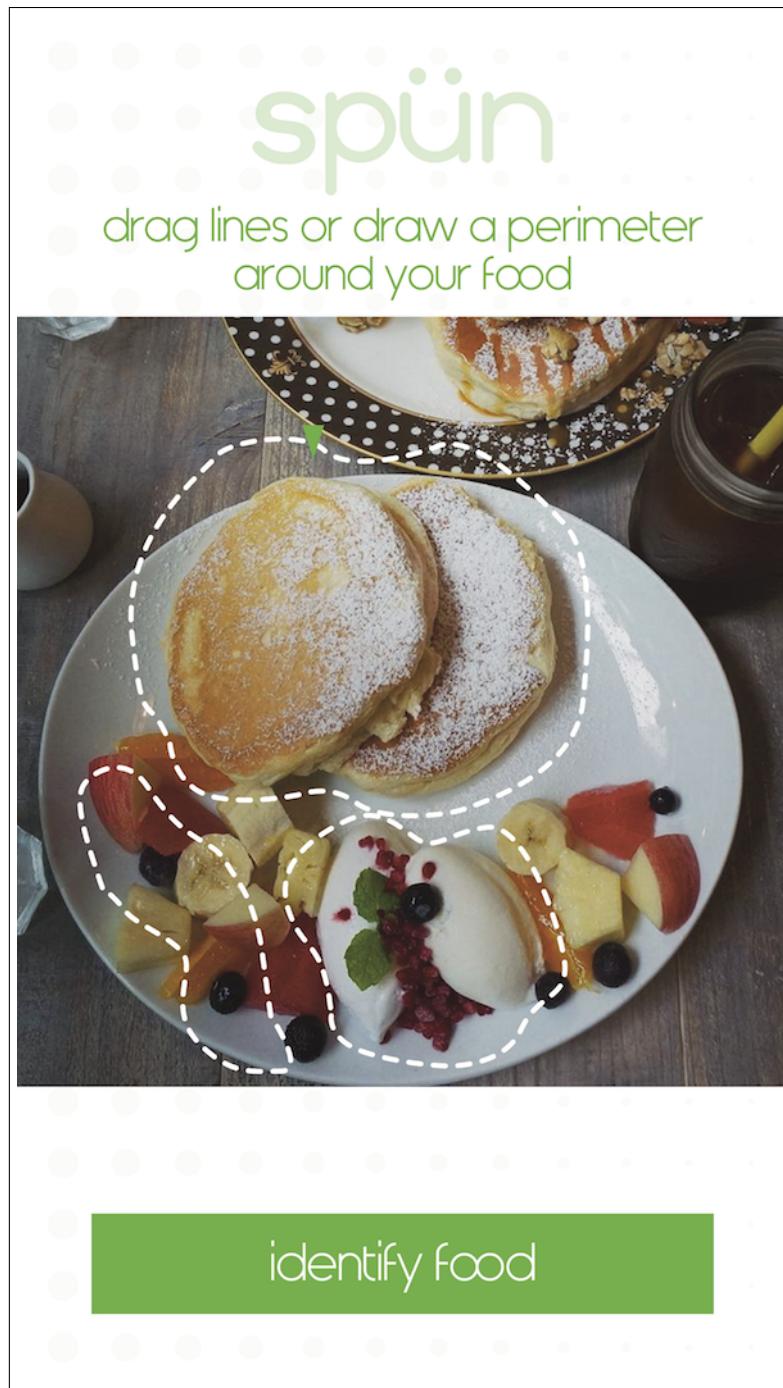


Рис. 1.32. Перший етап аналізу світлини з прийомом їжі мобільним додатком Spoon App (кластерування за умови невизначеності щодо кількості кластерів)

Другий етап у певному сенсі є навчанням з підкріпленням, адже користу-

вачеві пропонується підтвердити чи скорегувати кінцевий результат класифікації, як зображене на рис. 1.31. Проте більш цікавим нам видається саме перший етап аналізу світлини, і пропонована гібридна самонавчана система використовується саме на цьому етапі, адже вона краще від інших існуючих систем задовольняє умовам, що їх було висунуто на етапі формування технічних вимог до програмного забезпечення (Software Requirements Specifications):

- кластерування має проходити за умови невизначеності щодо кількості кластерів,
- оскільки кластерування відбувається на стороні клієнту, важливо мінімізувати обслуговальну складність алгоритму, а отже перевага надається методам послідовного кластерування.

По завершенні аналізу на першому етапі мобільний додаток попропонує розбиття світлини на  $t$  кластерів, як показано на рис. 1.32. Варто зауважити, що межі кластерів, що їх визначила самонавчана система, дещо відрізняються від тих, що зображені на рис. 1.32, для того, щоб користувачеві було зручніше візуально сприймати розбиття світлини на кластери, пунктирні лінії, що зображують межі кластерів, на декілька міліметрів віддалені від меж дійсних кластерів, проте на сервер для подальшої класифікації відправляється світлина з розбиттям, що запропонувала система. На цьому етапі користувач може скорегувати розбиття, перетягнувши пунктирну лінію меж кластерів, чи зовсім видалити пропонований кластер. Хоча навчання з підкріпленим у прямому сенсі, відбувається лише на етапі класифікації (база даних на сервері оновлюється, коли користувач корегує результат класифікації), а на цьому етапі маємо саме навчання без учителя, це все ж таки дає змогу у певному сенсі дати оцінку кластеруванню системи: вважаємо кластерування успішним, якщо користувач не робив жодних змін до пропонованого розбиття, та неуспішним, коли розбиття було скореговане. Після бета тестування мобільного додатку маємо наступні результати:

Цікавим видається перебіг подій, коли користувач видалив один кластер,

кількість та межі кластерів залишилися незмінними	608
межі кластеров було дещо змінено користувачем, проте кількість кластерів залишилась незмінною	61
користувач змінив кількість кластерів та межі пропонованих кластерів	52
користувач видалив кластер(и), межі інших пропонованих кластерів лишилися незмінні	29

Таблиця 1.13

**Результати бета тестування первого етапу (кластерування за умови невизначеності щодо кількості кластерів) аналізу світлини мобільним додатком «Spoon App»**

проте залишим межі інших кластерів незмінними, у такому випадку, як зазначалося вище, вважається що кластерування не було успішним. Проте подальший аналіз показав, що у 90% таких випадків користувач залишив на тарілці нейстівний предмет (виделку, ложку тощо), і хоча система вірно відвела йому окремий кластер, не має сенсу класифікувати його та визначати калорійність цього предмету, тому логічно, що користувач видалив його на цьому етапі. У 10% випадків, як показав аналіз світлин з початковим кластеруванням, що запропонувала самонавчана система, та світлин після корегування користувачем, користувач свідомо видаляв складову прийому їжі, зазвичай найменш корисну (тітєчко, шоколад тощо). Тому видається доцільним ці 4% світлин також віднести до таких, що були вірно розкластеровані системою (яка, проте, не бере до уваги людський фактор).