

# Синтез структури нечітких нейромереж для задач прогнозування курсів акцій на ринках цінних паперів

Доповідач: студент гр. КА-41м Левін Д.К.

Науковий керівник: д.т.н. проф. Зайченко Ю.П.

# Актуальність роботи

---

Ринок цінних паперів є найбільш доступне джерело фінансування економічного зростання. Він полегшує доступ усім суб'єктам економіки до отримання необхідних грошових ресурсів. Для того щоб передбачити поведінку динаміки цін на ринку цінних паперів здійснюють прогнозування показників фондового ринку. За допомогою ефективно діючого ринку цінних паперів можна гнучко перерозподіляти кошти між галузями, концентрувати їх на найбільш перспективних напрямках науково-технічного прогресу.

# Дані

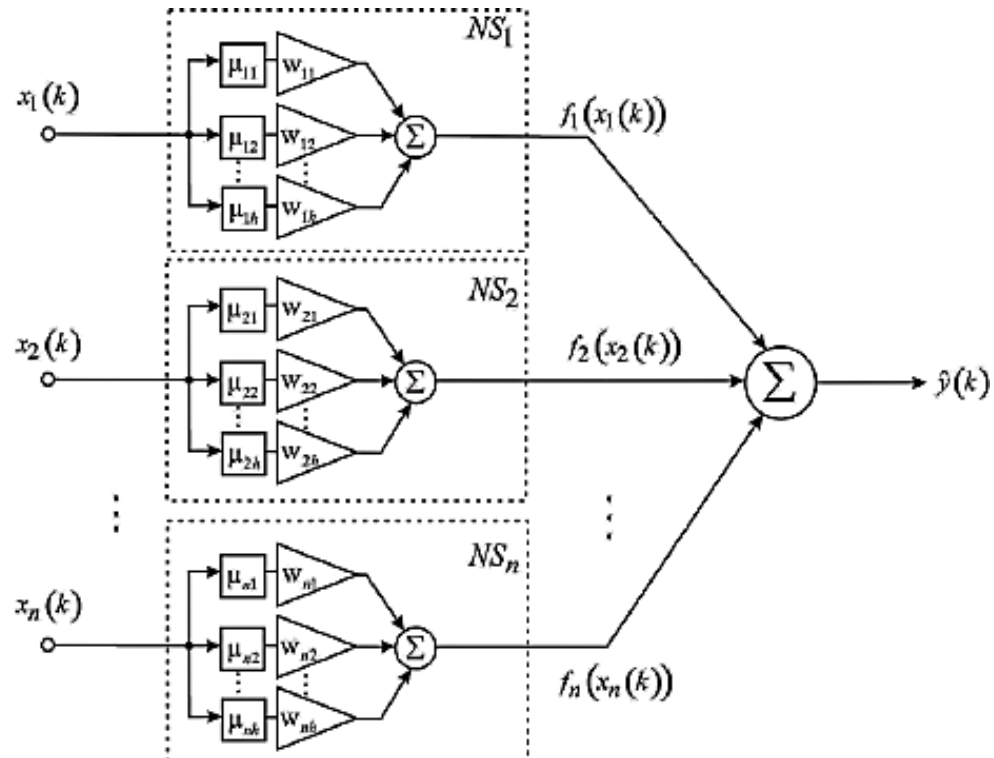
---

Фінансові індекси 17 найбільших компаній США з різних галузей на біржах NASDAQ та NYSE (**ЩОХВИЛИНИ**, з 1.01.2015 по 1.05.2016):

- |                     |                |
|---------------------|----------------|
| 1. Apple            | 10. Coca-Cola  |
| 2. Boeing           | 11. McDonald`s |
| 3. Caterpillar      | 12. Microsoft  |
| 4. Walt Disney      | 13. P&G        |
| 5. General Electric | 14. AT&T       |
| 6. Google           | 15. Verizon    |
| 7. Hewlett-Packard  | 16. Walmart    |
| 8. IBM              | 17. Yahoo      |
| 9. Intel            |                |

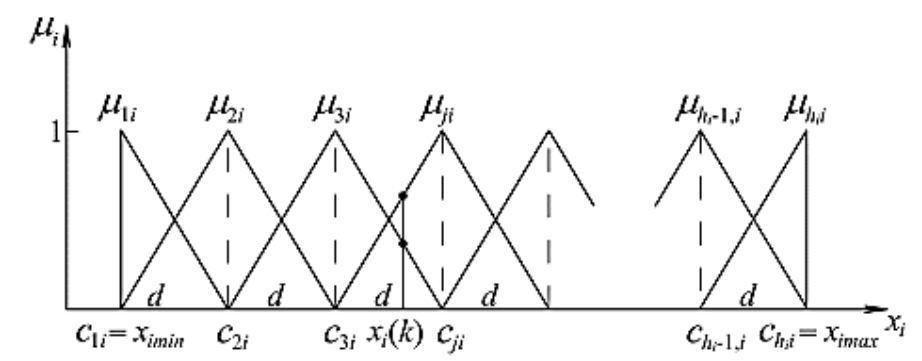
# Використані моделі

Нео-фаззі нейрон (neo-fuzzy neuron):



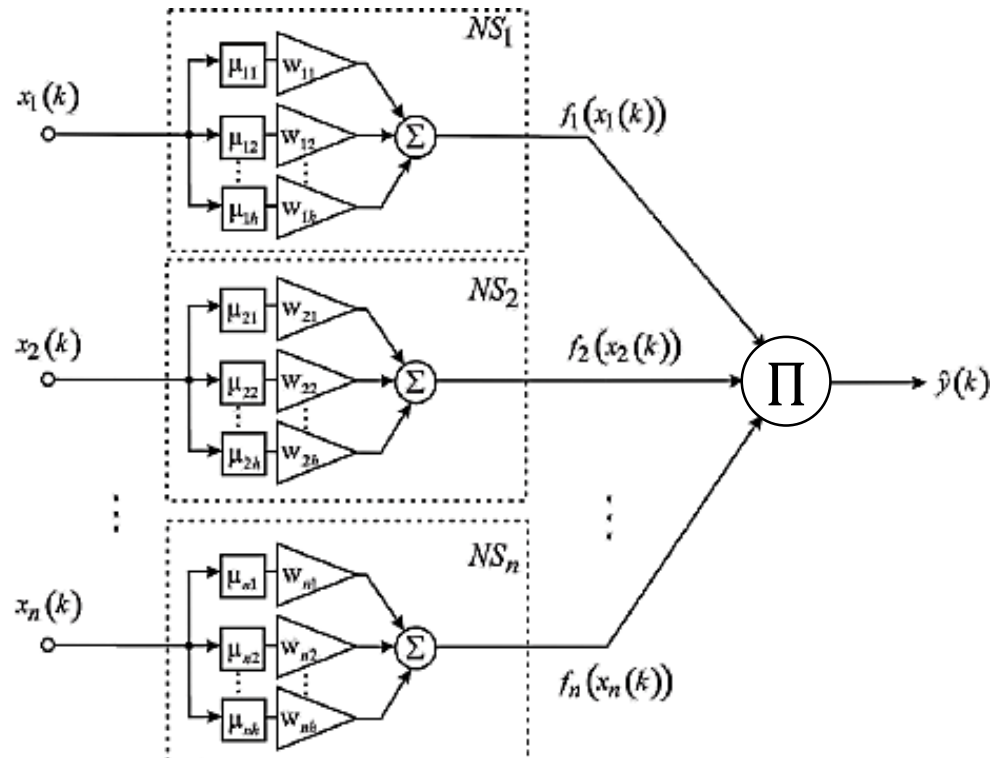
$$y = \sum_{i=1}^n \sum_{j=1}^m w_{ij} \mu_{ij}(x_i)$$

Функції належності:



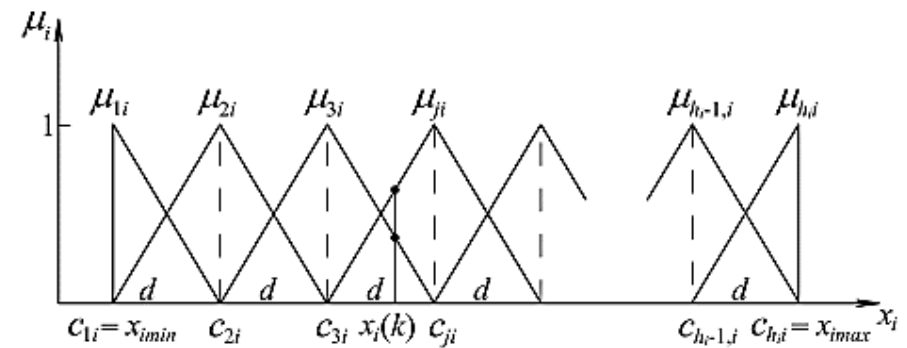
# Використані моделі

Мультиплікативний нео-фаззі нейрон:



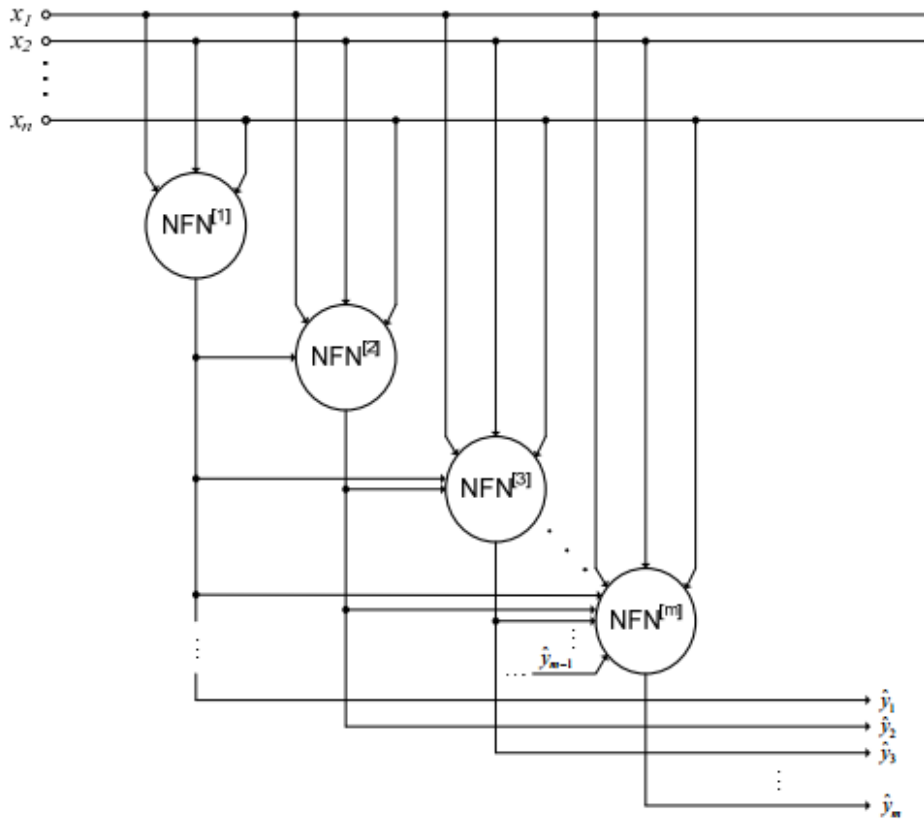
$$y = \prod_{i=1}^n \sum_{j=1}^m w_{ij} \mu_{ij}(x_i)$$

Функції належності:

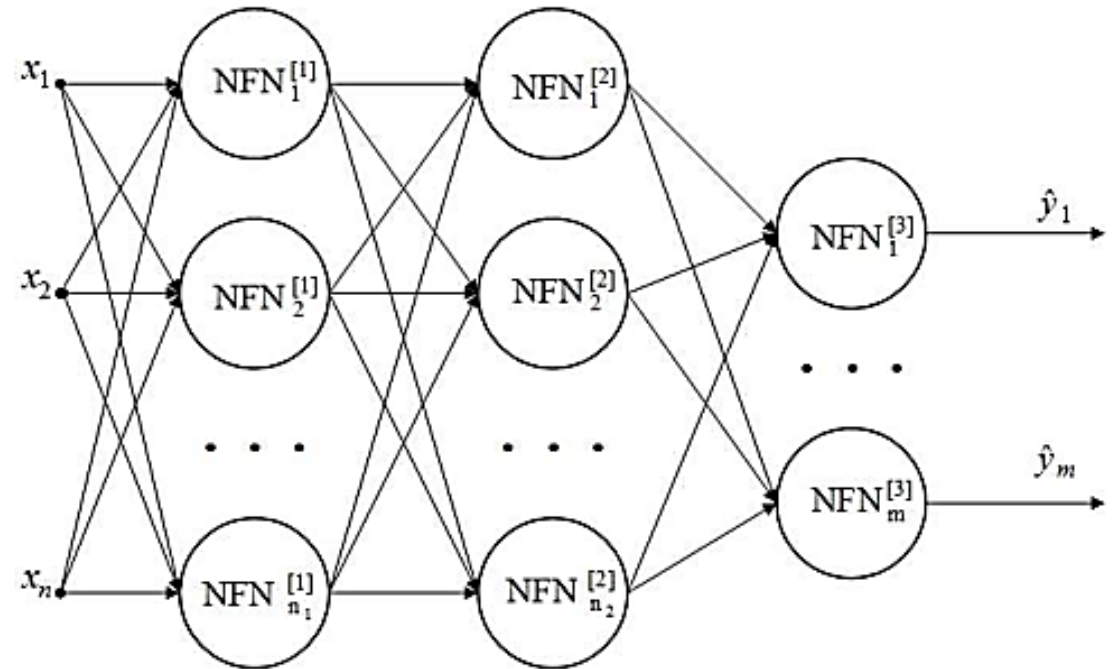


# Нео-фаззі структури

Нео-фаззі каскад



Нео-фаззі мережа



# Ініціалізація моделі

---

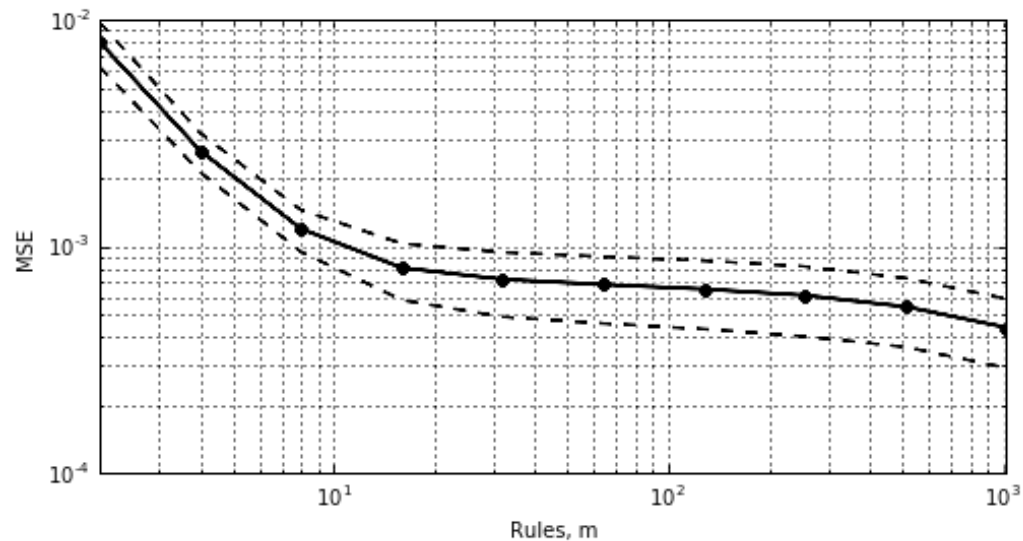
1. Випадкові значення
2. Константне значення
3. Запропонований метод:

$$\left\{ \begin{array}{l} d_i = \frac{c_{i,j+1} - c_{ij}}{2} = \frac{x_i^{max} - x_i^{min}}{2m}, j = 2..m \\ w_{i1} := \frac{1}{n} M\{y^k | x_i^k \in [c_{i,1}, c_{i,1} + d_i)\} \\ w_{im} := \frac{1}{n} M\{y^k | x_i^k \in [c_{i,m} - d_i, c_{i,m}]\} \\ w_{ij} := \frac{1}{n} M\{y^k | x_i^k \in [c_{i,j} - d_i, c_{i,j} + d_i)\} \end{array} \right.$$

# Порівняння методів ініціалізації

Дані:  $\Delta t = 1$  год.

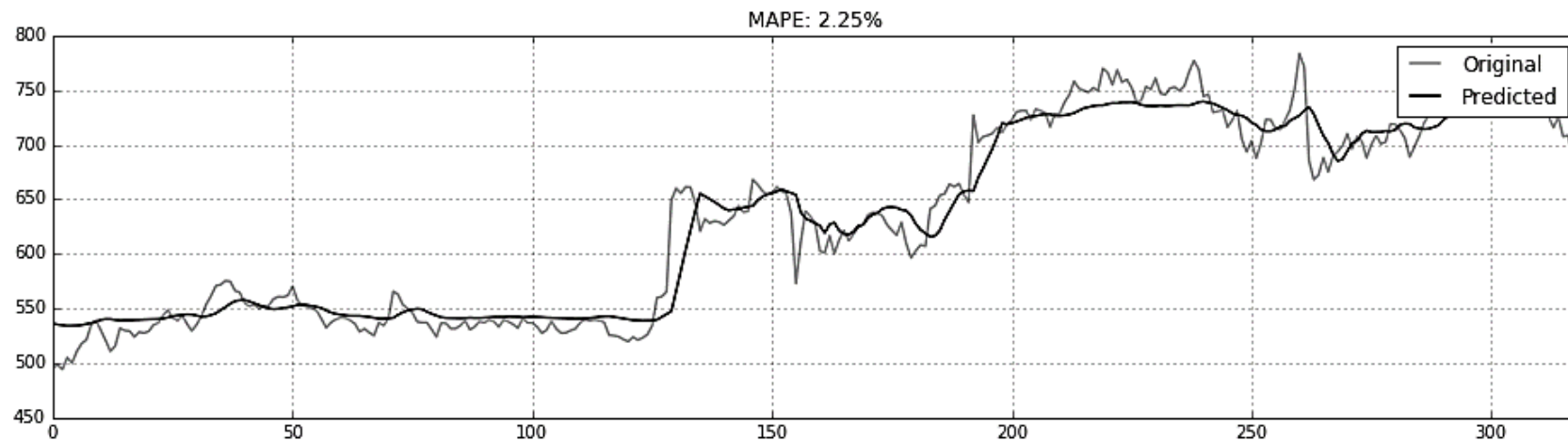
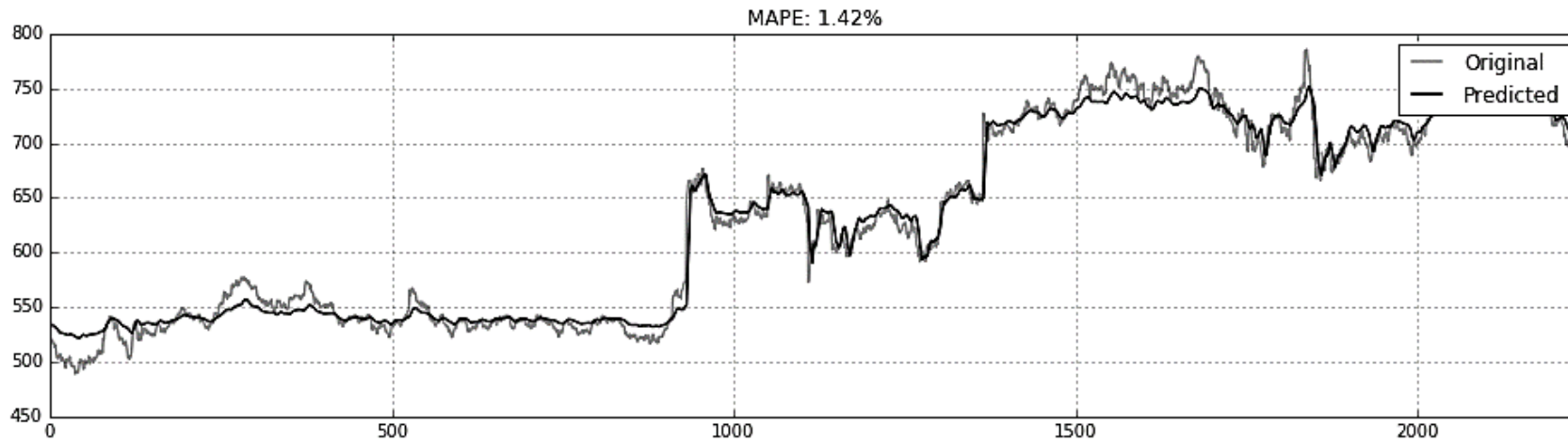
Нейрон:  $n = 8$ ,  $m = 2..1024$ ,



Індекс	Новий метод	Випадкова	Константа
AAPL	0.0006	0.2898	0.3735
BA	0.0005	0.2574	0.4589
CAT	0.0004	0.2501	0.3678
DIS	0.0005	0.2563	0.3291
GE	0.0005	0.3088	0.3312
GOOG	0.0004	0.2982	0.3185
HPQ	0.0003	0.2867	0.3451
IBM	0.0003	0.2680	0.3881
INTC	0.0005	0.2535	0.3339
KO	0.0006	0.2567	0.2357
MCD	0.0002	0.3011	0.2379
MSFT	0.0007	0.2893	0.3119



# Порівняння методів ініціалізації



# Алгоритми навчання

---

1. Алгоритм зворотного поширення помилки (SGD)
2. Rprop
3. AdaGrad, RMSProp
4. Пропонується: модифікація RProp для ефективного навчання НФН:

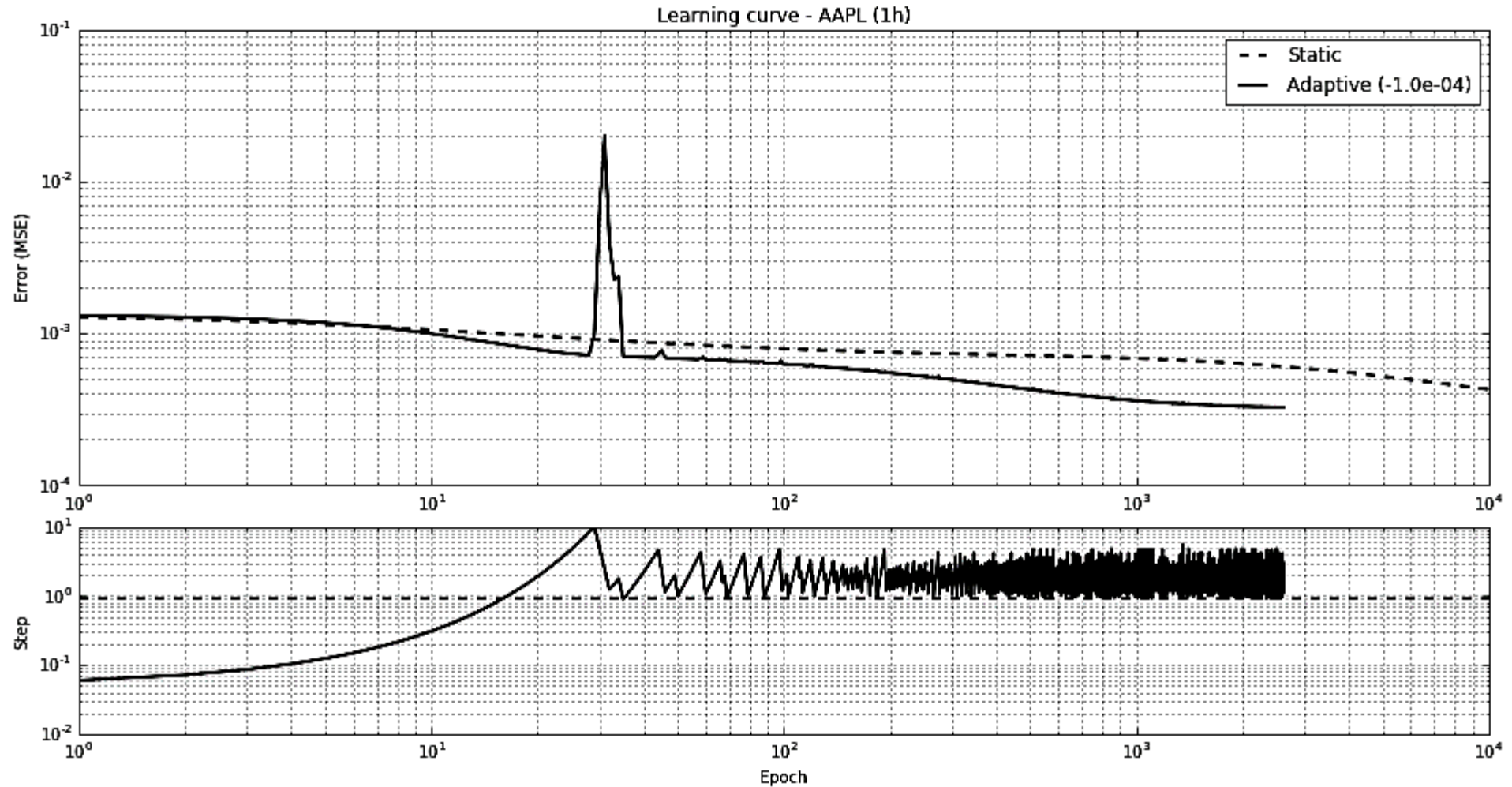
$$\Delta := \begin{cases} \eta^+ \Delta, & E(t) \leq E(t-1) \\ \eta^- \Delta, & E(t) > E(t-1) \end{cases}$$
$$w_{ij} = \begin{cases} w_{ij} - \Delta \cdot g_{ij} \\ w_{ij}, & E(t) \geq E(t-1) \end{cases}$$

# Порівняння алгоритмів навчання

Нейрон:  $n = 8$ ,  $m = 64$

Індекс	$\Delta t = 1$ хв.		$\Delta t = 1$ год.		$\Delta t = 1$ день.	
	NFN-Rprop	SGD	NFN-Rprop	SGD	NFN-Rprop	SGD
AAPL	2140	7194	1173	10000	170	485
BA	3442	10000	1116	7923	94	201
CAT	2046	8074	856	4406	131	302
DIS	2569	9879	944	8665	165	372
GE	2708	10000	885	5535	113	235
GOOG	3548	10000	1185	9939	119	253
HPQ	2881	10000	816	4931	171	347
IBM	3295	10000	850	4625	114	248
INTC	2843	10000	870	5162	184	395
KO	2395	10000	1049	8542	100	287
MCD	3096	10000	881	4664	183	545
MSFT	2069	8172	1066	7167	152	365

# Алгоритми навчання



# Гіпотеза ефективного ринку

---

Вся суттєва інформація негайно і повною мірою відбивається на ринковій курсовій вартості цінних паперів: ринок є ефективним щодо деякої інформації, якщо вона відразу і повністю відбивається в ціні активу, що робить цю інформацію непотрібною для отримання надприбутків.

**Наслідок:** ціни акцій є процесами випадкових блукань, для яких найкращий прогноз:  $y_{t+1}^* = y_t$ .

# Перевірка гіпотези ефективного ринку

---

Тест на наявність нульових кореляцій Льюнга-Бокса:

$$Q = n(n + 1) \sum_{j=1}^h \frac{\rho_k^2}{n - j}$$

Тест відношення дисперсій Ло-МакКінлі:

$$VR(k) = \frac{1}{k} \frac{\sigma^2(k)}{\sigma(1)} - 1$$

# Тест Льюнга-Бокса

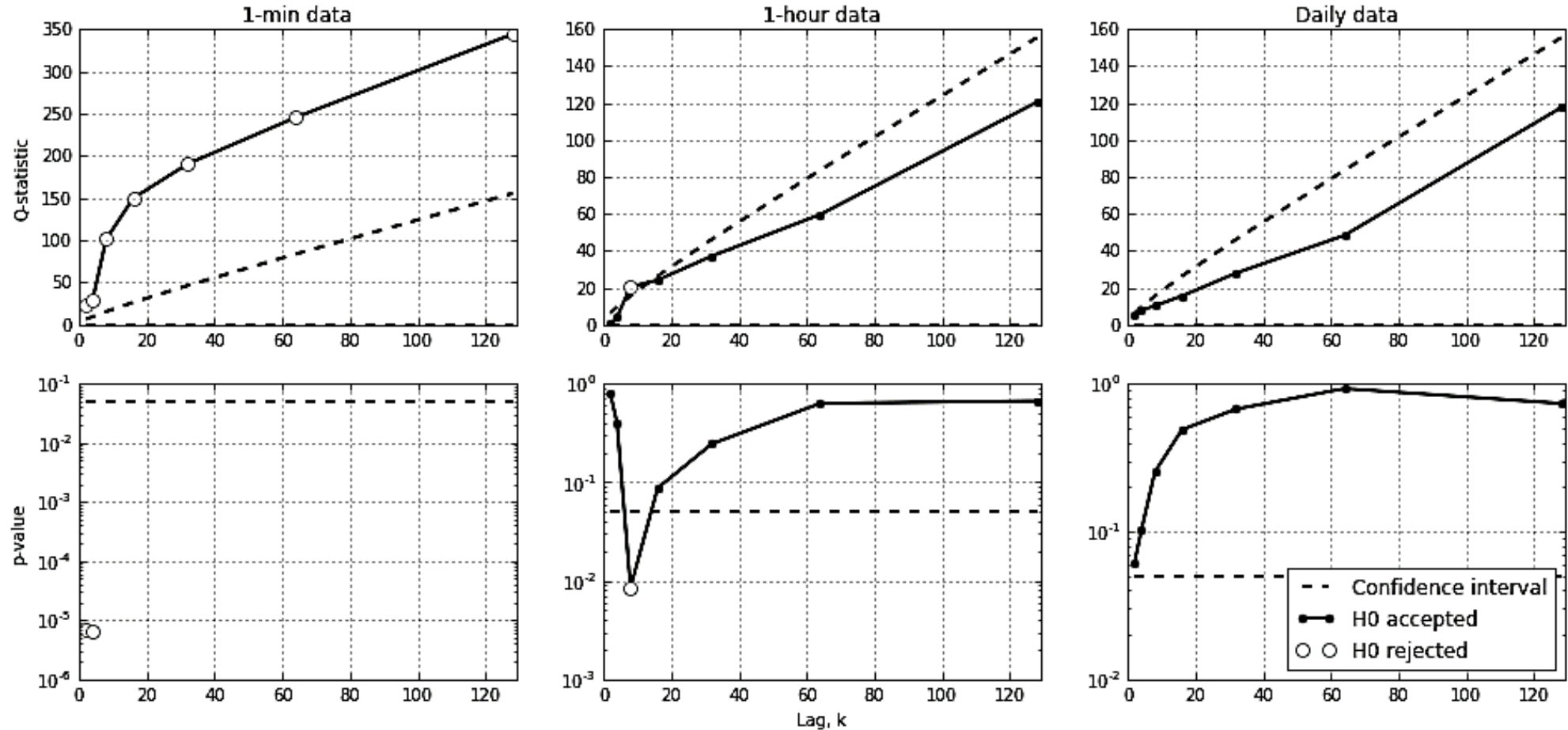
---

Індекс	$\Delta = 1$ год.		$\Delta = 1$ день	
	$h = 8$	$h = 128$	$h = 8$	$h = 128$
AAPL	0.00	0.00	0.00	0.00
DIS	0.02	0.00	0.02	0.00
INTC	0.13	0.02	0.17	0.79
MCD	0.00	0.07	0.14	0.04
PG	0.02	0.19	0.02	0.29
VZ	0.02	0.78	0.67	0.47
YAHOO	0.01	0.67	0.25	0.74

Решта індексів показали наявність кореляцій тільки на масштабі в 1 хв.

# Тест Льюнга-Бокса

Variance-ratio test - YAHOO



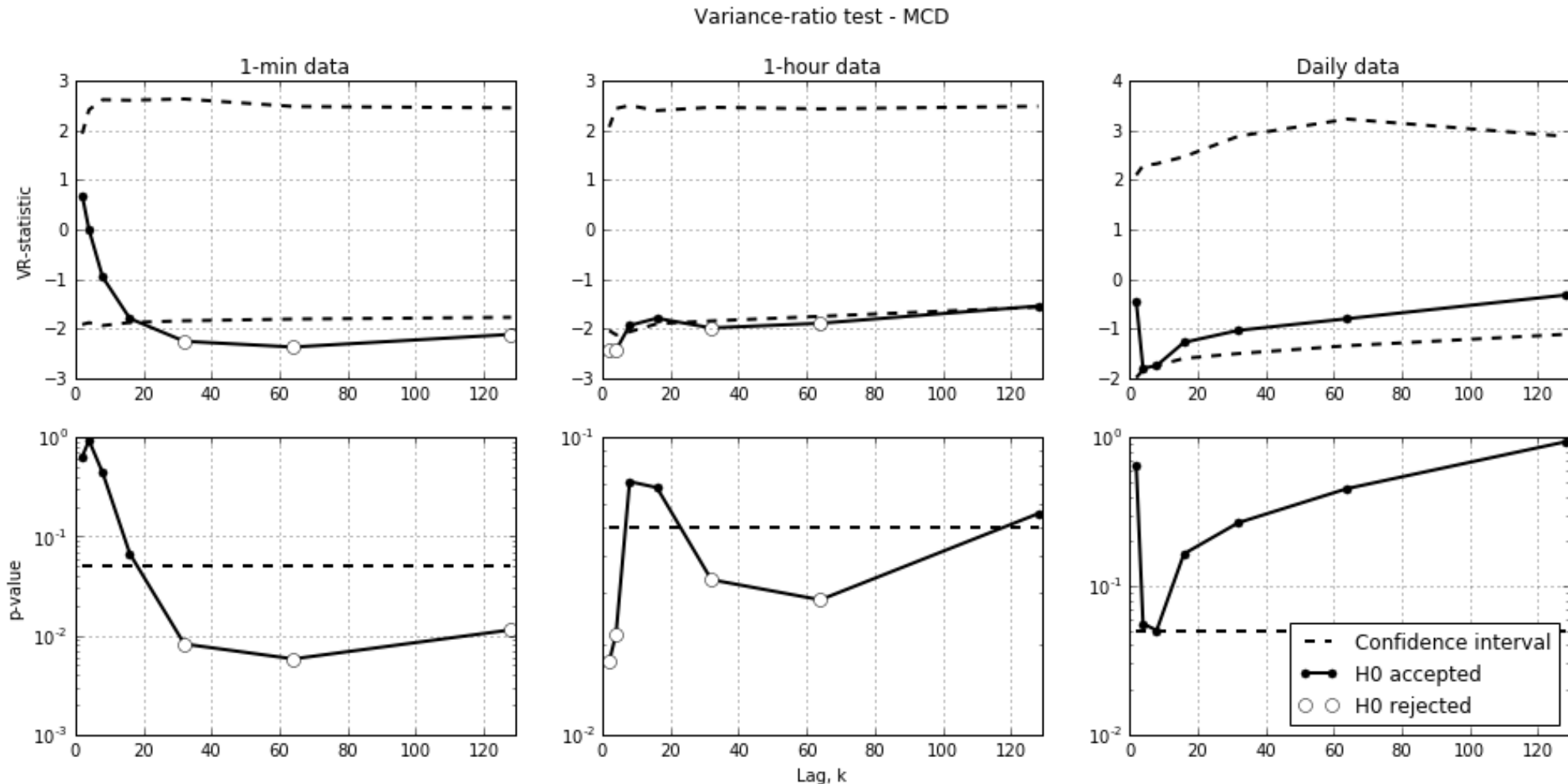


# Тест Ло-МакКінлі

Індекс	$\Delta = 1$ хв.				$\Delta = 1$ год.				$\Delta = 1$ день
	16	32	64	128	4	16	32	64	4
<i>h</i>	16	32	64	128	4	16	32	64	4
AAPL	0.00	0.00	0.00	0.01	0.12	0.25	0.18	0.29	0.13
CAT	0.77	0.55	0.34	0.20	0.04	0.21	0.28	0.27	0.33
DIS	0.14	0.18	0.07	0.08	0.04	0.64	0.51	0.57	0.64
MCD	0.07	0.01	0.01	0.01	0.02	0.07	0.03	0.03	0.06
T	0.00	0.00	0.00	0.00	0.12	0.02	0.02	0.05	0.05
VZ	0.47	0.13	0.01	0.01	0.10	0.11	0.27	0.66	0.63
YAHOO	0.43	0.69	0.60	0.64	0.70	0.35	0.24	0.18	0.01

Решта індексів і значень  $h$  не показали відхилень у поведінці дисперсій.

# Тест Ло-МакКінлі



# Результати прогнозування

Середнє значення MSE обох моделей нео-фаззі нейрону, обчислене за допомогою кросс-валідації на 16 розбиттях.

Індекс	RWH	NFN		mNFN	
AAPL	0.00036	0.00036	+0.5%	0.00036	+0.1%
BA	0.00022	0.00022	-2.0%	0.00022	-1.4%
CAT	0.00022	0.00023	-2.3%	0.00022	-1.5%
DIS	0.00033	0.00034	+2.5%	0.00033	-0.8%
GE	0.00027	0.00026	-2.3%	0.00027	-0.6%
GOOG	0.00026	0.00026	-1.7%	0.00026	-1.3%
HPQ	0.00014	0.00014	-0.0%	0.00014	-0.6%
IBM	0.00018	0.00019	-1.7%	0.00018	+1.0%
INTC	0.00028	0.00028	-1.9%	0.00028	-1.7%
KO	0.00035	0.00036	-2.4%	0.00036	-1.9%
MCD	0.00017	0.00017	+1.1%	0.00017	-0.5%
MSFT	0.00038	0.00038	+1.0%	0.00039	-1.7%
PG	0.00018	0.00019	-2.6%	0.00018	+1.0%
T	0.00031	0.00030	+2.5%	0.00030	+1.4%
VZ	0.00027	0.00027	+0.6%	0.00027	-0.4%
WMT	0.00010	0.00010	+1.5%	0.00010	+1.1%
YAHOO	0.00017	0.00017	-2.4%	0.00017	-1.6%
AAPL	0.00036	0.00036	-0.5%	0.00036	-0.1%

# Висновок

---

Існують підстави вважати, що ціни акцій на фондових ринках узгоджуються з гіпотезою ефективного ринку, і за результатами проведених тестів більшість досліджених індексів підтвердили цю гіпотезу. Проте не всі фондові активи є процесами випадкового блукання і містять деяку інформацію про майбутній рух ціни, що дозволило отримати для цих активів прогнози, що кращі за базову модель.

# Новизна робота

---

1. Запропоновано модифікацію нео-фаззі нейрону – мультиплікативний НФН, що продемонстрував не гірші результати за оригінальну модель, проте теоретично застосовний до інших класів задач.
2. Модифіковано алгоритм Rprop для навчання нео-фаззі нейрону, що дозволило суттєво скоротити час навчання порівняно з традиційним алгоритмом
3. Розроблено новий метод ініціалізації параметрів нео-фаззі нейрону, який дозволяє суттєво покращити початкове рішення і тим самим покращити характеристики навчання
4. Розроблені моделі досліджено в рамках задачі прогнозування цін акцій компаній на ефективних ринках

# Подальші дослідження

---

1. Аналіз більшої кількості даних – 500 індексів найбільших світових компаній за останні 36 років на різних масштабах часу для дослідження довгострокових трендів
2. Аналіз високочастотних даних торгів, адже існують підстави вважати, що високочастотний трейдинг містить більш виражені тренди.
3. Дослідження сторонньої інформації, яка впливає на ціни фінансових активів – новин, публікацій в соціальних мережах
4. Дослідження впливу торгівельних стратегій та ціну активу.