

УДК 004.82

*Резнік Р.Ю., студент 2 курсу
спеціальності «Комп'ютерні
технології обробки даних
(DataScience)»*

*Штовба С.Д., д-р техн. наук, професор,
професор кафедри інформаційних
технологій*

МОДИФІКАЦІЯ АЛГОРИТМУ ВАНГА-МЕНДЕЛЯ ШЛЯХОМ ВИКОРИСТАННЯ СУПЕРЕЧЛИВИХ ПРАВИЛ

Донецький національний університет імені Василя Стуса, м. Вінниця

Труднощі ідентифікації багатофакторних залежностей часто обумовлені відсутністю достатнього обсягу точних експериментальних даних або наявністю результатів спостережень в формі нечітких знань та лінгвістичних експертних оцінок. Один із шляхів подолання цих труднощів полягає у застосуванні нечіткої ідентифікації, тобто методів побудови моделей на основі теорії нечітких множин та нечіткої логіки. Найчастіше нечітка ідентифікація проводиться коли модель залежності являє собою нечітку базу знань. Власне для використання нечіткої бази знань її потрібно сформувати, тобто, згенерувати нечіткі правила.

Одним із найбільш відомих алгоритмів синтезу правил нечіткої бази знань є алгоритм Ванга-Менделя [1]. Ідея алгоритму полягає у генерації одного нечіткого правила з кожного рядка вибірки експериментальних даних. Потім із цим правил-кандидатів обирають ті, що мають максимальний ступінь активації. При цьому із множини суперечливих правил залишають лише одне.

В доповіді пропонується модифікацію алгоритму Ванга-Менделя, в якій дозволяється використання суперечливих правил [2] за принципом «головного конкурента». До бази знань вноситься не лише правило переможець з максимальним ступенем активації, але і друге за рангом правило з тим самим антецедентом. Таким чином, формується база знань із суперечливими правилами. Під суперечливими розуміються правила з однаковими антецедентами та різними консеквентами, наприклад:

Якщо $(x_1 = \text{низький} \text{ та } x_2 = \text{середній})$, тоді $y = \text{середній}$,

Якщо $(x_1 = \text{низький} \text{ та } x_2 = \text{середній})$, тоді $y = \text{високий}$.

Якщо в базі знань вказана вище пара суперечливих правил, тоді результат логічного виведення за ними буде між *середній* та *високим*, тобто відповідатиме терму *вище середнього*, якого немає в терм-множині вихідної змінної y . Таким чином, застосування суперечливих правил підвищує «роздільну здатність» опису залежності, яку моделює база знань. Відповідно, з'являються шанси підвищити точність нечіткої бази знань за рахунок використання суперечливих правил.

В класичному алгоритмі Ванга-Менделя спершу визначають ступені належності всіх змінних $x_1^{(i)}$, $x_2^{(i)}$ та $y^{(i)}$ до нечітких множин.

Для прикладу на рис. 1 $x_1^{(1)}$ належить до терму В1 зі ступенем належності 0,8, 0,2 до В2 і нульову ступінь для інших термів. Так само $x_2^{(2)}$ має одиничний ступінь належності до множини СЕ, а до всіх інших нульову.

Далі, для кожної змінної потрібно відібрати множини з максимальним ступенем належності. Для прикладу див. рис. 1, $x_1^{(1)}$ має максимальну ступінь належності до множини В1, $x_2^{(1)}$ має максимальну ступінь належності до множини S1, та $y^{(1)}$ має максимальну ступінь належності до множини СЕ. Власне з цих максимальних ступенів належності до множин і формуються відповідні правила, наведені кроки потрібно повторити для кожної пари вхід-вихід.

$(x_1^{(1)}, x_2^{(1)}, y^{(1)}) \rightarrow [x_1^{(1)}(0,8 \text{ В1, max}), x_2^{(1)}(0,7 \text{ S1, max}), y^{(1)}(0,9 \text{ СЕ, max})] \rightarrow$

Правило 1 : Якщо x_1 це В1 та x_2 це S1, то $y \in \text{СЕ}$.

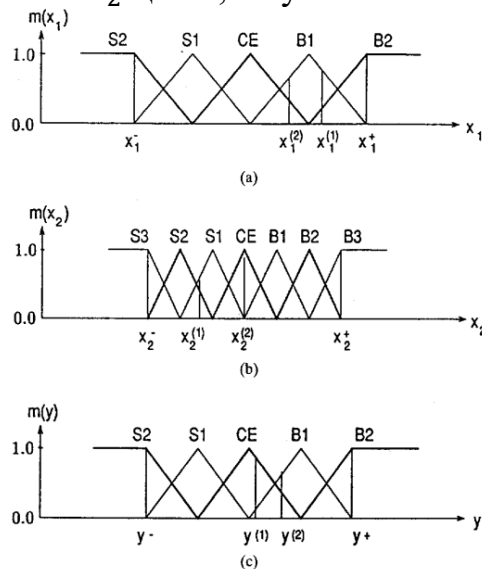


Рисунок 6 – Поділ вхідного та вихідного просторів на нечіткі області та відповідні функції належності [1]

На кінець навмисно накладають обмеження, щоб в базу знань не потрапило конфліктного правила. Шляхом призначення ступеня кожному правилу, створеному з пар даних, і відбором того правила з конфліктної групи, яке має максимальну ступінь. Ступінь правила визначається як добуток ступенів його складових, що власне створює це правило. Таким чином отримана база знань виходить без конфліктних правил.

Ми ж допускаємо відбір суперечливих правил. Для кожного відібраного на попередньому кроці правила долучається і конкурент – правило з ідентичним антецедента та з іншим консеквентом, яке має другий ранг.

Для прикладу з рис. 1 y^2 має максимальну ступінь належності до множини В1 0,7 та найближчу конкуруючу множину СЕ з ступенем 0,45. На кінець потрібно сформулювати правило з вихідним термом з максимальним ступенем та конкурентним консеквентом. Наведені кроки потрібно повторити для кожної пари вхід-вихід.

$(x_1^{(1)}, x_2^{(1)}, y^{(1)}) \rightarrow [x_1^{(1)}(0,8 \text{ B1, max}), x_2^{(1)}(0,7 \text{ S1, max}), y^{(1)}(0,9 \text{ CE, max})] \rightarrow$
 Правило 1 : Якщо x_1 це B1 та x_2 це S1, то $y \in \text{CE}$.

$(x_1^{(1)}, x_2^{(1)}, y^{(1)}) \rightarrow [x_1^{(1)}(0,8 \text{ B1, max}), x_2^{(1)}(0,7 \text{ S1, max}), y^{(1)}(0,1 \text{ B1, max})] \rightarrow$

Правило 2 (Правило конкурент) : Якщо x_1 це B1 та x_2 це S1, то $y \in \text{B1}$.

Проведемо обчислювальні експерименти з ідентифікації залежностей за допомогою нечітких баз Мамдані з суперечливими правилами. Експерименти проведемо за методикою з роботи [3] для таких двофакторних залежностей рис. 2:

$$y = x_1^2 - x_2^3 * \tan(0.1 * x_1) \in [-20,25], x_2 \in [3,16]; \quad (\text{a})$$

$$y = \frac{x_1^2 + x_2}{0.1e^{x_1 + x_2}}, x_1 \in [-5,5], x_2 \in [3,7]; \quad (\text{б})$$

$$y = x_1^2 - x_2^3 * \tan(0.1 * x_1) \in [-2,3], x_2 \in [1,5]. \quad (\text{в})$$

Для кожної з залежностей синтезовано множину нечітких правил Мамдані. Для залежності (а) застосуємо лінгвістичне розбиття вхідних 4x4, для залежності (б) – 4x3, для залежності (в) – 4x5. Відповідно, для цих залежностей синтезуємо 16, 12 та 20 несуперечливих правил-кандидатів класичним алгоритмом Ванга-Менделя та 32, 24 і 40 правил модифікованим алгоритмом. Далі, побудуємо криві навчання (рис. 3) у формі залежності *RMSE* від кількості відібраних до бази знань правил. *RMSE* розраховано на тестовій вибірці із 100 точок. Криві навчання побудовано для найкращого набору правил, який визначався адитивним та субтрактивним жадібними алгоритмами.

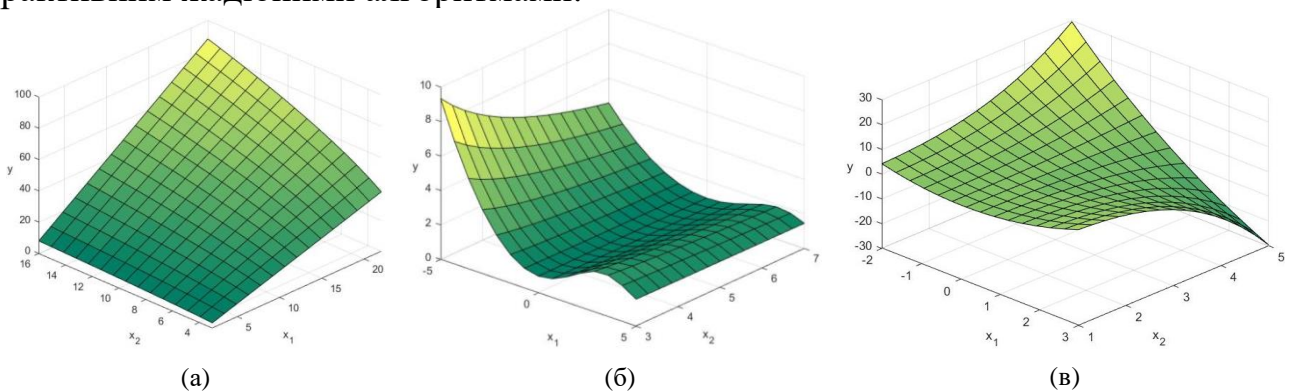


Рисунок 7 – Поверхні еталонних залежностей

Порівнюючи криві навчання на рис.3 та дані зведені в табл. 1, бачимо доцільність використання суперечливих правил-конкурентів алгоритму Ванга-Менделя для синтезу нечітких баз знань Мамдані. Майже в усіх експериментах нечіткі бази знань отримані модифікованим алгоритмом Ванга-Менделя кращі за точністю ніж бази знань отримані класичним алгоритмом. Для близької до лінійної залежності (1) покращення суттєве. Для суттєво нелінійних залежностей (2) та (3) покращення точності за рахунок використання суперечливих правил незначне.

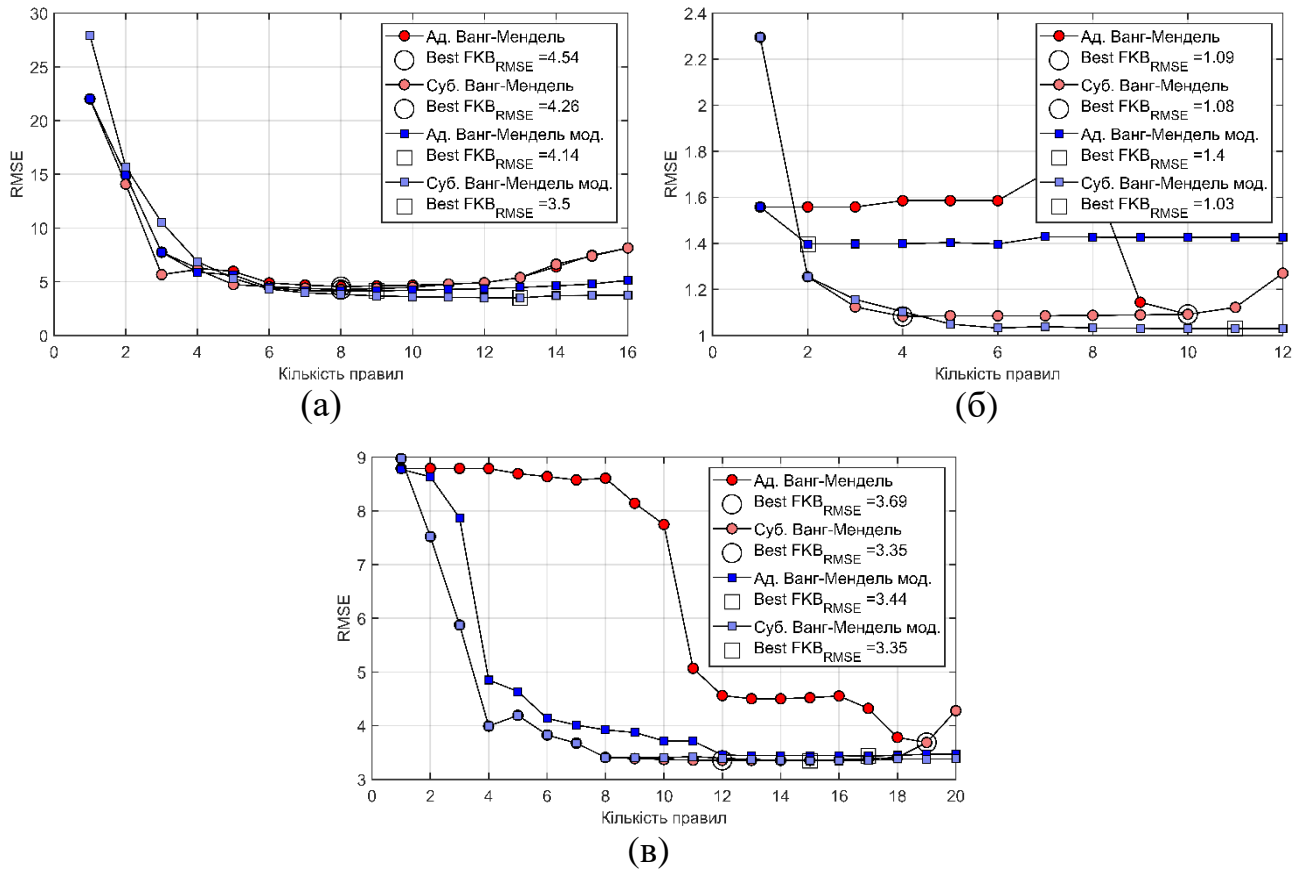


Рисунок 8 – Криві навчання залежностей (а-в)

Таблиця 1 – Показник точності алгоритмів залежностей (а-в)

		Залежність (а)	Залежність (б)	Залежність (в)
		RMSE		
Класичний Ванг-Мендель	Адитивний	4,54	1,09	3,69
	Субтрактивний	4,26	1,08	3,35
Модифікований Ванг-Мендель з головним конкурентом	Адитивний	4,14	1,4	3,44
	Субтрактивний	3,5	1,03	3,35

Запропонована модифікація алгоритму Ванга-Менделя, а саме: використання правил конкурентів при синтезі баз знань Мамдані, забезпечує кращу точність, за рахунок підвищення роздільної здатності опису.

Список літературних джерел:

1. Wang L. X., Mendel J. M. Generating fuzzy rules by learning from examples //IEEE Transactions on systems, man, and cybernetics. – 1992. – Т. 22. – №. 6. – С. 1414-1427.
2. Резнік Р.Ю., Штовба С.Д. Синтез нечіткої бази знань Мамдані з використанням суперечливих правил. Матеріали III Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених "Прикладні інформаційні технології" (22 квітня 2022 року) - Вінниця: ДонНУ імені Василя Стуса., с. 34-37.
3. Штовба С.Д., Галуцак А.В. Ідентифікація багатofакторних залежностей за допомогою баз знань. Лабораторний практикум: електронний навчальний посібник. – Вінниця: Вінницький національний технічний університет, 2016. – 96 с.