

# БАЄСОВСЬКІ НЕЙРОННІ МЕРЕЖІ: ПОРІВНЯЛЬНИЙ АНАЛІЗ ТОЧНОСТІ КЛАСИФІКАЦІЇ

О. М. Ахметшин, Д. ф. - м. н., О. О. Киргизов, Магістр

Дніпропетровський національний університет, Україна, 49050,  
м. Дніпропетровськ, пр. Науковий 13.  
Електронна пошта: akhm@mail.dsu.dp.ua

В статті приведені теоретичні відомості про Баєсовські та Ймовірнісні нейронні мережі – класифікаційні мережі, засновані на баєсовським вирішальним правилі. Представлено структури та метод побудови мереж, а також результати класифікації на прикладі медичної бази даних онкологічних захворювань.

## 1. ВСТУП

Більшість задач – включаючи діагностику захворювань, розпізнавання образів та пророкування можуть бути представлені як задачі класифікації, так як кожна вимагає ідентифікації стану та має множину параметрів, що їх описують. Створення точних класифікаторів з предкласифікованих даних являється важливим та активно розвиваючимся розділом в машинному навчанні та дослідженню даних.

При побудові автоматичних систем класифікації, виникають три основні задачі: представлення даних, виділення інформативних ознак та побудова системи класифікації.

Одними з методів, на основі яких можна прийняти рішення про клас нового образу, являються обчислення центроїда кожного класу, або використання критерію "найближчого сусіда". Більш "витончені" методи на додаток до розгляду відстані враховують щільність розподілу сусідніх образів.

В основі побудови оптимальних автоматичних систем класифікації, в загальному випадку, використовується баєсовське класифікаційне правило або його різновиди. Цей підхід, заснований на припущенні, що задача вибору рішення сформульована в термінах теорії ймовірностей і відомі всі ймовірні величини, що представляють інтерес. Останнім часом бурхливо розвиваючимся методом є Баєсовські нейронні класифікаційні мережі, що враховують окрім окремого вмісту кожного параметру, що описують образи класифікації, ще й взаємозв'язки між ними.

Представимо дві нейронні мережі, що виконують класифікацію даних за баєсовським правилом [1]:

• Ймовірнісні нейронні мережі (ІНМ) –

використовують лише умовні спільні щільності розподілу параметрів кожного класу даних.

- Баєсовські нейронні мережі (БНМ) – використовують умовні спільні щільності розподілу параметрів як добуток умовних щільностей кожного з параметрів, враховуючи залежності між ними для кожного класу даних [2,3].

## 2 ЙМОВІРІСНІ НЕЙРОННІ МЕРЕЖІ

Один з видів нейронних мереж, що враховують щільність розподілу образів, які можна використовувати для класифікації називаються ймовірнісні нейронні мережі (ІНМ). В основу класифікації у мережі ІНМ покладене використання правила Баєса. Це вимагає оцінки функції щільності ймовірностей. Оцінка проводиться за допомогою методу Парзена, у якому використовується вагова функція, що має центр у точці, що представляє навчальний образ. Така вагова функція називається функцією потенціалу або ядром:

$$f(r) = \exp(-r^2) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right) \quad (1)$$

Приклад архітектури мережі для рішення задачі класифікації показаний на рис. 1.

Задачею вхідного шару є розподіл даних вхідного образу для шару нейронів. У даному

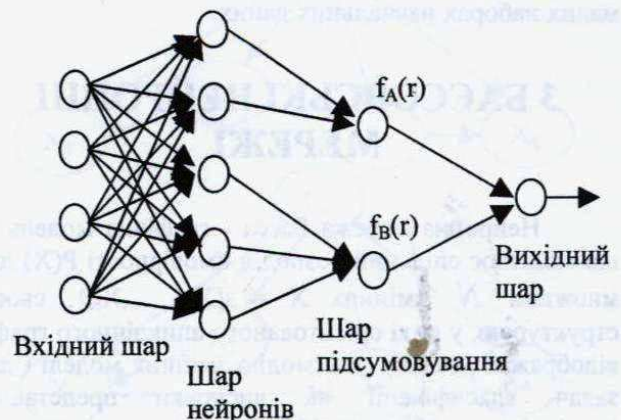


Рис. 1. Приклад архітектури ІНМ



випадку кожен вхідний набір даних має чотири параметри. Шар нейронів має по одному елементу для кожного образу з набору навчальних даних. Вхідний шар і шар нейронів утворять повнозв'язну

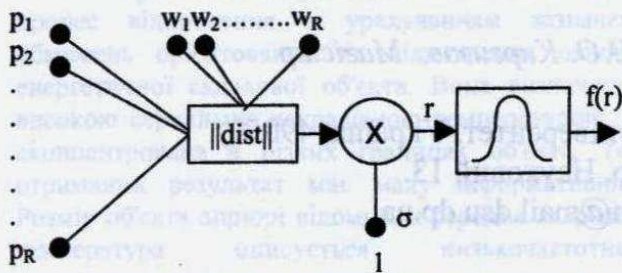


Рис.2. Приклад архітектури радіального нейрона

структуру. Для вхідних в елемент шару нейронів зв'язків вагові значення встановлюються рівними елементам відповідного образу. На рис. 2 представлений радіальний нейрон із входом  $p$  і вагами в розмірності  $R$ .

Якщо різниця між  $p$  і  $w$  зменшується, то значення функції збільшується. У такий спосіб нейрон діє як детектор відображаючий ступінь близькості вхідного вектора  $p$  до вектора  $w$ , зсув  $\sigma$  дозволяє регулювати чутливість нейрона [2].

Шар підсумовування має по одному елементу для кожного класу з навчальної області даних. До будь-якого елемента шару підсумовування йдуть зв'язку тільки від нейронів, що належать відповідному класу. Елемент підсумовування підсумовує вихідні значення нейронів. Ця сума дає оцінку значення функції щільності розподілу ймовірностей для сукупності екземплярів відповідного класу та апіорної ймовірності кожного класу. Вихідний елемент являє собою дискримінатор, що вказує елемент шару підсумовування з максимальні значенням активності. Для мережі ІММ не потрібно навчання в тім змісті, яке потрібно для мереж із зворотним поширенням помилок, тому що всі параметри мережі ІММ (число елементів і значення ваг) визначаються безпосередньо навчальними даними.

Мережі ІММ обмежується задачами класифікації, допускають наявність помилкових даних і забезпечують корисні результати навіть на малих наборах навчальних даних.

## 3 БАЄСОВСЬКІ НЕЙРОННІ МЕРЕЖІ

Нейронна мережа Баєса - графічна модель  $S$ , що моделює спільний розподіл ймовірності  $P(X)$  для множини  $N$  змінних  $X = \{X_1, \dots, X_N\}$ , своєю структурою, у виді орієнтованого ациклічного графа, відображає реальну взаємодію змінних моделі і для задач класифікації як висновки представляє ймовірність приналежності для кожного з класів даних [2]. Вершини в  $S$  відповідають змінним  $X$ , й

також можна представити як мережу зв'язних нейронів. Нехай  $X_i$  - означає як змінну, так і відповідну вершину у  $S$ , а  $Pa_i$  - означає батьків вершини  $X_i$ , тобто змінних від яких залежить  $X_i$ , тоді спільний розподіл ймовірностей для змінних  $X$  структури  $S$ :

$$P(X) = \prod_{i=1}^N P(X_i | Pa_i) \quad (2)$$

де  $P(X_i | Pa_i)$  - умовний розподіл ймовірності. Розглянемо довільний орієнтований (ациклічний) граф, рис. 3:

Для даного графа з правила спільного розподілу ймовірностей маємо:

$$P(A, B, D, E) = P(A)P(B)P(D|A) \cdot P(E|A, B) \quad (3)$$

Ймовірність вершини  $A$ , враховуючи локальні розподіли ймовірностей, обчислюється:

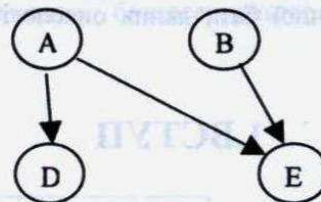


Рис 3 Приклад Баєсовської мережі.

$$P(A|B, D, E) = \frac{P(A)P(B)P(D|A)P(E|A, B)}{\sum_{A'} P(A')P(B)P(D|A')P(E|A', B)} \quad (4)$$

$$= \frac{P(A)P(D|A)P(E|A, B)}{\sum_{A'} P(A')P(D|A')P(E|A', B)}$$

Області застосування Баєсовських мереж: оцінка щільності розподілу змінних, класифікація, регресія, кластеризація, експертні мережі, визначення ступеня впливу складових моделі і результатів її зміни.

### 3.1 БАЄСОВСЬКІ КЛАСИФІКАЦІЙНІ МЕРЕЖІ

Баєсовська мережа (БМ) класифікації  $B = \langle N, A, \Theta \rangle$  - спрямований ациклічний граф  $\langle N, A \rangle$ , де кожна вершина  $n \in N$  являє собою ознаку набору даних, а кожна дуга  $a \in A$  між вершинами, являє собою ймовірнісну залежність. З кожною вершиною  $n$ , зв'язана її умовна щільність розподілу ймовірності  $\theta_i$  (УЦПІ). БМ може використовуватися для обчислення умовної ймовірності однієї вершини, з огляду на значення привласнені іншим вершинам, отже, БМ може використовуватися як класифікатор, що дає апостеріорну ймовірність вершини класу (4). Відповідно, класифікаційна система повертає номер класу з найбільшою апостеріорною ймовірністю. Одна з переваг БМ у порівнянні з іншим типам класифікаторів, типу нейронної мережі, - те, що БМ враховують взаємозв'язку серед набору ознак. Застосування БМ для задач класифікації включає дві



підзадачі:

- визначення графічної моделі БМ;
- визначення апіорних ймовірностей та умовних розподілів ймовірностей;
- виконання класифікації;

У задачах, що є оптимальними, тобто для повних наборів даних – для визначення УЩРІ використовує емпірично умовну частоту даних, інакше необхідно робити її відновлення. Дану задачу можна вирішити, скориставшись методом вікон Парзена.

Для визначення структури БМ маються два підходу:

- БМ - структура, що кодує спільний розподіл ознак. Це означає, що найкраща БМ - мережа, яка найкращим способом описує спільне розподілення даних моделі [2].
- БМ - структура, що кодує групу умовних відносин незалежності (УН) серед вершин. Використовуючи статистичний критерій взаємної інформації, можемо знаходити усі УН та використовувати їх як критерій побудови БМ [3].

## 3.2 ВИДИ КЛАСИФІКАЦІЙНИХ БМ

Нижче представлено п'ять видів БМ класифікаторів, які відповідають апіорному представленню взаємодії змінних системи: Проста БМ, Деревоподібна БМ, Розширена БМ, Мульти БМ і Звичайна БМ.

### 3.2.1 Проста БМ

Проста БМ [2] є структурою, що має вершину класу  $C$  як батьківську вершину всіх інших вершин змінні  $X_1, X_2, X_3, X_4$  (Рис. 4). Ніякі інші з'єднання в Простий БМ не розглядаються.

На відміну від нижче перелічених класифікаторів, Проста БМ легка в побудові, оскільки структура дається апіорно і не треба використовувати процедуру побудови мережі.

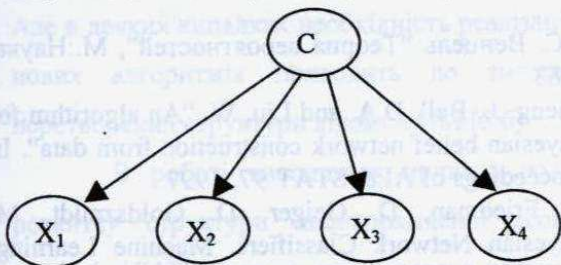


Рис. 4 Проста БМ

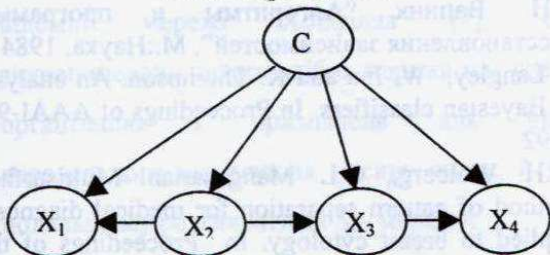


Рис. 5 Деревоподібна БМ

Проста БМ припускає, що всі змінні незалежні один від одного, та представляє непогані результати в порівнянні зі складними класифікаторами на великій кількості даних, які не мають сильної кореляції [5].

### 3.2.2 Деревоподібна БМ

Деревоподібні БМ класифікатори розширюють Прості БМ, в припущенні, що змінні мережі формують дерево (Рис.5). Такі мережі є частковим випадком Розширених БМ [2].

### 3.2.3 Розширена БМ

Розширені БМ класифікатори доповнюють Деревоподібні БМ, дозволяючи змінним мережі формувати довільний ациклічний граф (Рис.6).

Використання Розширених і Деревоподібних БМ менш ефективне в порівнянні з нижче викладеними БМ, оскільки для кожного класу

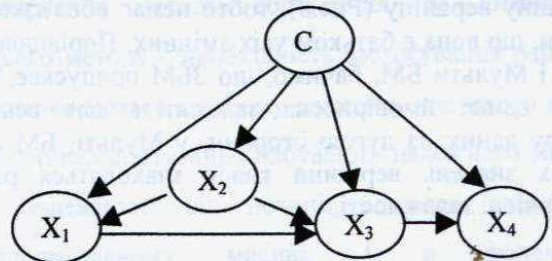


Рис. 6: Розширена БМ

передбачається однаковий вид залежності між змінними і вершиною класу даних.

### 3.2.4 Мульти БМ

Мульти БМ [3] складається з множини локальних мереж, кожна з яких відповідає значенню, що може приймати вершина класу (Рис 7). Мульти БМ може розглядатися як узагальнення Розширених БМ. У Розширених БМ співвідношення серед змінними є однаковим для всіх значень вершини класу; з іншої сторони в Мульти БМ співвідношення серед змінними для різних класів можуть бути різні - тобто, для різних значень вершини класу, змінні

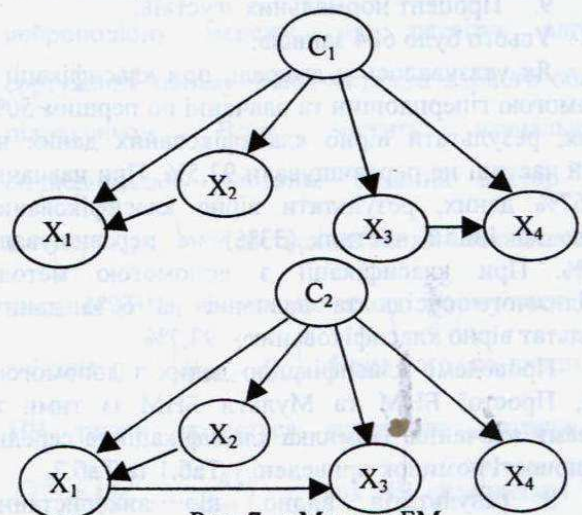


Рис. 7 : Мульти БМ



можуть формувати різні локальні мережі з різною структурою.

### 3.2.5 Звичайна БМ

Звичайна БМ (ЗБМ) - вид необмеженого БМ

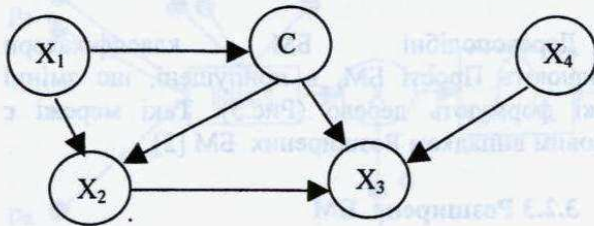


Рис. 8 : Звичайна БМ

класифікатора, що має особливість, не властиву перерахованим вище мережам. Загальна особливість перелічених БМ - те, що вершина класу трактується як спеціальна вершина - вона є батьком усіх змінних. Однак, ЗБМ трактує вершину класу як звичайну вершину (Рис.8), тобто немає обов'язкової умови, що вона є батьком усіх змінних. Порівнюючи ЗБМ і Мульти БМ, бачимо, що ЗБМ припускає, що існує єдина ймовірнісна залежність для всього набору даних; із дугою сторони, у Мульти БМ для різних значень вершини класу знаходяться різні ймовірнісні залежності.

## 4 РЕЗУЛЬТАТИ

Експерименти проводились на медичних даних ракових захворювань [6]. Записи описували зображення пункцій пацієнтів та мали наступні параметри:

1. Середня товщина згустків.
2. Однорідність розмірів згустків.
3. Однорідність форм згустків.
4. Густина згустків.
5. Епітеліальний розмірів згустків.
6. Процент згустків без оболонки.
7. Хроматин.
8. Мітоз.
9. Процент нормальних згустків.

Усього було 684 записів.

Як указувалось у джерелі, при класифікації з допомогою гіперплощин та навчанні по першим 50% даних, результати вірно класифікованих даних на іншій частині не перевищували 93,5%. При навчанні на 67% даних, результати вірно класифікованих даних на іншій частині (33%) не перевищували 95,9%. При класифікації з допомогою методу найближчого сусіда та навчанні на 67% даних, результат вірно класифікованих - 93,7%

Проведемо класифікацію даних з допомогою ІНМ, Простої БНМ та Мульти БНМ із тими ж умовами навчання. Помилка класифікації та середні ймовірності помилок приведені у Таб.1 та Таб.2.

З результатів видно, що використання баєсовського правила та врахування залежностей

між параметрами приведе до суттєвого підвищення точності класифікації навіть при навчанні на половині набору даних.

Таблиця №1.

Навчання - 50% даних, тестування - інші 50%.

	Проста БНМ:	Мульти БНМ:	ІНМ:
Середня Ймовірність Помилки	0.0091	$2.9 \cdot 10^{-6}$	0.0388
Помилка Класифікації, %	8.7977	1.7595	16.715

Таблиця №2.

Навчання - 67% даних, тестування - інші 33%.

	Проста БНМ:	Мульти БНМ:	ІНМ:
Середня Ймовірність Помилки	0.0067	$0.3 \cdot 10^{-7}$	0.0425
Помилка Класифікації, %	7.8221	1.7699	15.929

## 5 ВИСНОВОК

В статті було продемонстровано новий метод класифікації, заснований на Баєсовській нейронній мережі, з допомогою якого проведено класифікацію даних онкологічних захворювань.

З порівнянь результатів видно, що використання взаємозв'язків між параметрами підвищує точність класифікації даних.

Таким чином, врахування залежностей між параметрами та щільностей їх розподілення, значно підвищує точність класифікації онкологічних захворювань.

## ЛІТЕРАТУРА

1. Л.С. Венцель "Теория вероятностей", М.:Наука, 1983
2. Cheng, J., Bell, D.A. and Liu, W. "An algorithm for Bayesian belief network construction from data". In Proceedings of AI & STAT'97, 1997
3. N. Friedman, D. Geiger, D. Goldszmidt, M. Bayesian Network Classifiers. Machine Learning, 29, 1997, P. 131-161.
4. В.Н. Вапник "Алгоритмы и программы восстановления зависимостей", М.:Наука, 1984
5. P. Langley, W. Iba and K. Thompson. An analysis of Bayesian classifiers. In Proceedings of AAAI-9,2 1992
6. W.H. Wolberg, O.L. Mangasarian. Multisurface method of pattern separation for medical diagnosis applied to breast cytology. In Proceedings of the National Academy of Sciences, 87,9193--9196