

ТЕОРІЯ ДЕМПСТЕРА-ШАФЕРА ДЛЯ ЗАДАЧ КЛАСИФІКАЦІЇ

Лузан М.О.

Науковий керівник – к.т.н., доцент каф. ПІ Бабій А.С.

Харківський національний університет радіоелектроніки

(61166, Харків, пр. Науки, 14, каф. Програмної інженерії, тел. (057) 702-14-46)

e-mail: mykyta.luzan@nure.ua, тел.: +38 (050) 876-04-19

Today classification tasks are highly used in many different cases from detecting fraud in banks and faults on electric stations to sports betting. Such algorithms as Logistic Regression, Naive Bayes, MLP etc. are in demand. But one of the bottlenecks of these methods is problem when appears uncertainty. In this case classifiers based on Dempster-Shafer theory provides good performance. It converts features into Dempster-Shafer mass function that gives normalized plausibilities which are more informative than common output probability distribution from other classifiers. The theory allows to determine the degree of belief (confidence) in the classifier's decision.

Використання теорії Демпстера-Шафера у задачах класифікації є раціональним рішенням, коли в нас є неоднозначність. Всі можливі події задаються множиною θ та функцією маси, яка відображає частки достовірності кожної з подій. У множину θ входять пуста множина (\emptyset), з масою нуль, всі можливі події та їх власні підмножини за умови, що маса події не задає ніяких додаткових тверджень щодо її підмножин, та універсальна множина, яка описує невизначеність. Сума мас цих подій складає одиницю. Також завдяки таким мірам, як довіра (1) та правдоподібність (2) можна задати інтервал, у якому буде знаходитись ймовірність появи події з множини θ .

$$Bel(A) := \sum_{B \in A} m(B), \quad (1)$$

$$Pl(A) := \sum_{B \cap A \neq \emptyset} m(B) = 1 - Bel(\bar{A}) \quad (2)$$

Дві функції маси, які показують незалежні частки достовірності можуть буди скомбіновані завдяки правилу Демпстера:

$$(m_1 \oplus m_2)(A) = \frac{1}{1 - k} \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad (3)$$

де $A \subseteq \theta$, $A \neq \emptyset$ та k – міра конфлікту між наборами мас.

Можна застосувати цю теорію для бінарної класифікації, коли в нас є тільки два класи, тобто $\vartheta = \{\vartheta_1, \vartheta_2\}$. Тоді представимо наші ознаки у вигляді вектору $\phi(x)$ довжиною j , а маси подій будуть розраховуватися як комбінація за правилом Демпстера для події з додатньою вагою достовірності та з від'ємною вагою для протилжної події, ваги яких дорівнюють:

$$w_j := \beta_j \cdot \phi_j(x) + \alpha_j, \quad (4)$$

де α_j та β_j – коефіцієнти; знак вагів відповідає тому відбудеться подія чи ні.

Далі маси різних ознак будуть об'єднані за формулою (3), та завдяки отриманій масі може бути розрахована міра правдоподібності (2) для відповіді моделі, як частка правдоподібності окремої події до суми правдоподібностей всіх подій. Повна схема класифікатора зображена на рисунку 1.

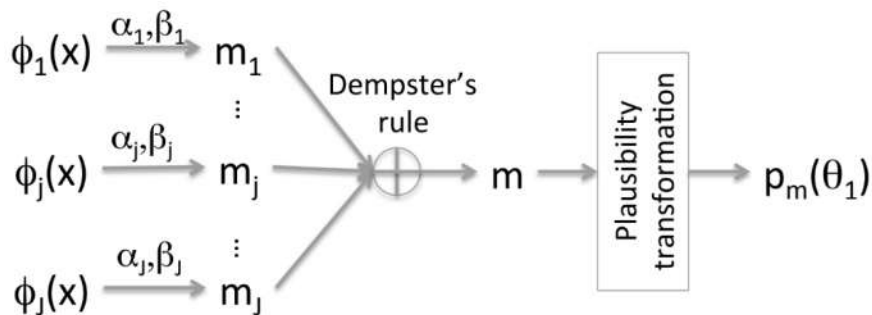


Рисунок 1 – Архітектура класифікатора для бінарного випадку [1]

Отже, завдяки класифікатору на основі теорії Демпстера-Шафера можна отримати нормалізовані значення правдоподібностей для подій, отриманих завдяки функції маси. Ця функція має більше ступенів свободи, тому це дає можливість робити кращі відповіді у випадках нестачі свідчень або їх конфліктності. Цей підхід можна реалізувати для вдосконалення широкого діапазону алгоритмів, включаючи нейронні мережі, які зараз користуються великим інтересом.

Перелік посилань

1. Thierry Denoeux. Logistic Regression, Neural Networks and Dempster-Shafer Theory: a New Perspective, 2018.
2. Denœux, Thierry. A neural network classifier based on Dempster-Shafer theory. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on. 30. 131 - 150. 10.1109/3468.833094, 2000.
3. Whitbrook, Amanda and Chen, Qi and Aickelin, Uwe and Roadknight, Chris. Data Classification Using the Dempster – Shafer Method. Journal of Experimental & Theoretical Artificial Intelligence, Volume 26, Issue 4, 2014.