

УДК 61:004.651(075.8)

Андрущак І.Є., Шендрик Ю.В.

Луцький національний технічний університет

РЕАЛІЗАЦІЯ МЕТОДУ ПРИЙНЯТТЯ РІШЕННЯ ІЗ ЗАСТОСУВАННЯМ ТЕОРІЇ ВТРАТ

Андрущак І.Є., Шендрик Ю.В. Реалізація методу прийняття рішення із застосуванням теорії втрат. В роботі реалізовано системний підхід до прийняття рішень, що враховує ставлення ОПР до різних стратегій лікування. Встановлені оцінки підтверджено експериментально як із зміною кількості атрибутів, так і обсягу наборів навчальних даних.

Ключові слова: теорія втрат, прийняття рішень, теорія очікуваної корисності.

Андрущак И.Е., Шендрик Ю.В. Реализация метода принятия решения с применением теории потерь. В работе реализован системный подход к принятию решений, учитывает отношение ОПР в различных стратегий лечения. Установленные оценки подтверждено экспериментально как с изменением количества атрибутов, так и объема наборов учебных данных.

Ключевые слова: теория потерь, принятие решений, теория ожидаемой полезности

Andrushchak I., Shendrik Y. Realization of decision-making method is with application of theory of losses. This paper implemented a systematic approach to decision making that takes into account the ratio of ODA to different treatment strategies. Installed evaluation confirmed experimentally as a change in the number of attributes and sets the amount of training data.

Keywords: theory of losses, making decision, theory of utility.

Вступ

Прийняття рішення дуже часто визначається невизначеністю, яка впливає на увесь процес вироблення рішення. В намаганні змодельовати невизначеність фахівці в галузі прийняття рішень пропонують методики, що ґрунтуються на теорії ймовірності і статистиці [Edwards,2007], теорії інформації та ентропії [Shannon,1962], нечіткій логіці [Zimmer, 2000].

В клінічних медичних дослідженнях багато зусиль прикладено розробці систем підтримки прийняття рішень для діагностики та лікування ряду захворювань [Schurink, 2005]. Більшість з цих систем ґрунтуються на ймовірнісних прогностичних моделях. Незважаючи на те, що прогностичні моделі зарекомендували себе як потенційно придатні до клінічних застосувань, все ж вони не відповідають очікуванням осіб, що приймають рішення (ОПР) з метою вдосконалення прийняття рішень. Однією з причин є та, що більшість медичних систем підтримки прийняття рішень (СППР) ґрунтуються на теорії очікуваної корисності, яка дуже часто не відповідає реальній дійсності. До того ж більшість моделей в медицині не враховують побажань ОПР, які є ключовими в раціональному процесі прийняття рішень.

Метою даної роботи є реалізувати підхід до прийняття рішень, що враховує ставлення ОПР до різних стратегій лікування.

Підхід полягає в наступному. По-перше, ми відходимо від традиційної теорії очікуваної корисності в намаганні задовольнити як формальні критерії раціональності, так і людську інтуїцію щодо правильних рішень. По-друге, застосовується теорія втрат (regret theory) [Tsalatsanis, 2010], оскільки втрата – це когнітивна емоція, що поєднує як раціональність, так і інтуїцію, які є ключовими елементами в прийнятті рішень, яке б використовувало особисті переваги ОПР. Отже, ми переформулюємо метод DCA з точки зору теорії втрат з метою оцінки альтернативних стратегій лікування з урахуванням ставлення до них та переваг ОПР.

Для реалізації такого підходу спочатку розраховується порогова ймовірність, при якій ОПР байдуже, які альтернативні стратегії прийняті. Це робиться визначаючи рівень втрати, який відчуває ОПР при прийнятті помилкового рішення. Далі застосовується метод DCA на основі втрат для визначення оптимальної стратегії. Оптимальною стратегією є стратегія, яка приносить найменше втрат.

Прийняття рішень на основі теорії втрат

На рис. 1 зображено типове дерево рішень для призначення лікування на основі прогностичної моделі. Тут представлено дві конкуруючі стратегії (лікувати та не лікувати) і чотири можливі наслідки.

Далі використовуватимуться такі позначення для випадкових подій: $D+$ - наявність захворювання, $D-$ - захворювання відсутнє, $Rx+$ - результати обстеження вказують на наявність захворювання та необхідність призначення лікування (для вибраної порогової ймовірності p_t), тобто «тест позитивний», $Rx-$ - тест негативний. Тоді $p = P(D+)$ - ймовірність наявності захворювання, $p = P(D-)$ - ймовірність відсутності захворювання, $U_i, i = \overline{1,4}$ - корисності, що відповідають кожному наслідку. А саме, $U_i = U(Rx_i, D_i)$, де $U(\bullet, \bullet)$ - деяка функція корисності, $Rx_i \in \{Rx+, Rx-\}$, $D_i \in \{D+, D-\}$. Так $U_1 = U(Rx+, D+)$ - корисність призначеного лікування для пацієнта, що має захворювання, $U_2 = U(Rx+, D-)$ - корисність призначеного лікування для пацієнта, що не має захворювання.

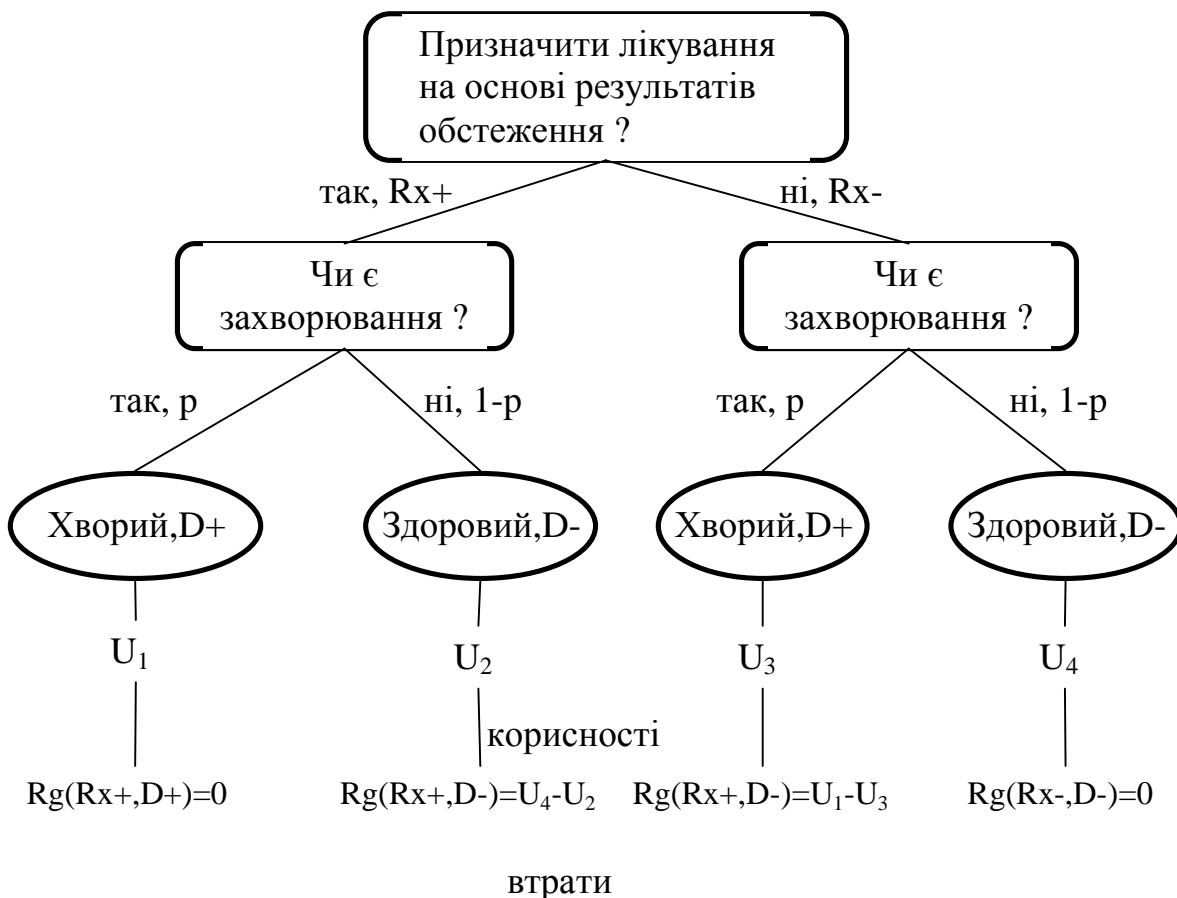


Рис. 1. Дерево рішень на основі теорії очікуваної корисності.

Зауважимо, що термін «лікування» вживається в даній роботі в загальному розумінні і він включає будь-яке медичне втручання, наприклад, терапевтичне лікування, хірургічну операцію або діагностичну процедуру.

Ймовірнісна природа прогностичних моделей значно ускладнює процес прийняття рішень. Наприклад, якщо прогностична модель оцінює ймовірність захворювання для пацієнта в 40%, то не зрозуміло, чи призначати пацієнту лікування чи ні. Відповідь з точки зору класичної теорії прийняття рішень полягає у застосуванні поняття порогової ймовірності p_t , яка визначається як ймовірність, при якій особі, що приймає рішення байдуже, яку стратегію обрати (призначити

лікування чи ні). Використовуючи поняття порогової ймовірності пацієнту слід призначити лікування, якщо $p \geq p_t$, і навпаки в протилежному випадку.

Однак оскільки в більшості випадків рішення приймаються в умовах невизначеності, то після прийняття рішення може виявитися, що більш прийнятною є інша альтернатива. Така інформація вводить для ОПР поняття втрат - англійською «regret», що дослівно перекладається як «жаль», «смуток» [Tsalatsanis, 2010].

Формально втрати можуть бути виражені як різниця між корисностями від вчиненої дії і від дії, яку зважаючи на ретроспективу слід було здійснити [Tsalatsanis, 2010]. Розрахунок втрат представлено на дереві рішень (рис.1).

Втрати можуть відчуватися будь-якою стороною, залученою в процес прийняття рішень. Ми для визначеності вважатимемо, що ОПР є лікуючий лікар.

Для початку застосуємо теорію втрат для оцінки порогової ймовірності p_t , при якій альтернативні стратегії (призначити лікування або не призначити) є рівносильними.

Виходячи з дерева рішень (рис.1) бачимо, що очікувані втрати від призначеного лікування становлять $(1 - p)(U_4 - U_2)$, втрати від ненаданого лікування - $p(U_1 - U_3)$. Виходячи з означення порогової ймовірності приходимо до співвідношення для визначення p_t :

$$(1 - p_t)(U_4 - U_2) = p_t(U_1 - U_3).$$

Звідси:

$$p_t = \frac{1}{1 + \frac{U_1 - U_3}{U_4 - U_2}} \quad (1)$$

- співвідношення для визначення порогової ймовірності на основі втрат.

Зауважимо, що співвідношення (1) невизначене для $U_4 - U_2 = 0$, тобто у випадку, коли немає жодних втрат від призначення непотрібного лікування. В такому разі $p_t = 0$ - це практично нереальна ситуація.

Метод DCA на основі втрат

Нагадаємо, що метод DCA може бути застосований для визначення стратегії, якої слід дотримуватися виходячи із значення порогової ймовірності (наприклад, використати модель і призначити лікування тільки у випадку $p \geq p_t$, лікувати всіх, не лікувати нікого).

У класичному методі DCA розраховується сукупна функція корисності [Vickers, 2006]. При застосуванні теорії втрат обчислюється очікувана втрата при використанні певної стратегії [Tsalatsanis, 2010]. Так на основі дерева рішень маємо:

$$ERg[\text{модель}] = P(p < p_t / D+)(U_1 - U_3) + P(p \geq p_t / D-)(U_4 - U_2), \quad (2)$$

де $Rg[\bullet]$ - втрати, пов'язані з використанням певної стратегії. Таким же чином з рис.2 маємо, що:

$$ERg[\text{лікувати всіх}] = (1 - p)(U_4 - U_2), \quad (3)$$

$$ERg[\text{не лікувати нікого}] = p(U_1 - U_3). \quad (4)$$

При аналізі стратегій на основі втрат використовується величина різниці сукупних очікуваних втрат (Net Expected Regret Difference, NERD), яка може бути розрахована [Tsalatsanis, 2010] для NERD[нікого не лікувати, модель], NERD[нікого не лікувати, лікувати всіх], NERD[нікого не лікувати, нікого не лікувати] за формулою:

$$NERD = \frac{\#TP}{n} - \frac{\#FP}{n} \frac{p_t}{1 - p_t}, \quad (5)$$

а для NERD[лікувати всіх, модель] за формулою:

$$NERD = \frac{\#TN}{n} \frac{p_t}{1-p_t} - \frac{\#FN}{n}. \quad (6)$$

Тут $\#TP, \#FP, \#TN, \#FN$ - кількості вірно та помилково позитивних, вірно та помилково негативних пацієнтів, розраховані для певної порогової ймовірності. Далі наведемо формулювання алгоритму методу DCA на основі втрат.

Вхідні дані: інформація про пацієнтів: кількість n , вектор діагнозів $d = (d_1, \dots, d_n)^T$ та матриця результатів обстежень $I = (I_1, \dots, I_n) \in R^{m \times n}$. Тут

$$d_i = \begin{cases} 1, & \text{захворювання є} \\ 0, & \text{захворювання немає} \end{cases}, \quad I_i \in R^m - \text{діагностичні дані для кожного пацієнта, що}$$

можуть бути перетворені у ймовірнісне значення \hat{p}_i , наприклад на основі логістичної регресії.

Вихідні дані: різниці сукупних очікуваних втрат для усіх моделей та стратегій: лікувати всіх, лікувати нікого як функції від порогової ймовірності.

Метод:

1. Вибираємо порогову ймовірність p_t .
2. Вважаємо, що пацієнт є позитивним, якщо $\hat{p}_i \geq p_t$.
3. Обчислюємо кількості вірно та помилково позитивних.
4. Обчислюємо значення NERD(нікого не лікувати, модель) на основі (5).
5. Обчислюємо NERD(лікувати всіх, модель) на основі (6).
6. Обчислюємо значення NERD(нікого не лікувати, лікувати всіх) на основі (5).
7. Повторюємо кроки 1-6 на інтервалі значень p_t .
8. Креслимо графіки усіх NERD, обчислених на кроках 4-6 залежно від значення порогової ймовірності.

Використовуючи метод DCA на основі втрат оптимальне рішення для кожної порогової ймовірності вибирається на основі принципу транзитивності, а саме, якщо $NERD(\text{стратегія 1, стратегія 2}) > NERD(\text{стратегія 2, стратегія 3}) > 0$, то стратегія 2 краща від стратегії 1, а стратегія 3 краща від стратегії 2. Отже, стратегія 3 – оптимальна для даної порогової ймовірності.

Зауважимо, що метод може бути застосований не лише до чисельних вхідних даних про пацієнта, але й мультимедійних, наприклад, ЕКГ, ЕЕГ, УЗД, КТ, МРТ і ін., які повинні попередньо бути виражені в ймовірнісних значеннях.

Висновки

В роботі представлено програмну реалізацію методу DCA на основі втрат з метою підтримки прийняття рішень щодо вибору оптимальної стратегії лікування. Порівняно із традиційним методом DCA підхід на основі втрат враховує переваги ОПР щодо стратегій на основі попереднього досвіду та використовує інтуїтивно більш зрозуміле поняття втрати.

Метод реалізовано у вигляді Netbeans-додатку з вхідними даними у вигляді xml-файлу, що робить можливим його використання в різноманітних Інтернет-проектах, пов'язаних з прийняттям рішень.

1. A.J. Vickers, A.M.Cronin, E.B.Elkin, M.Gonen Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers, BMC Medical Informatics and Decision Making 2008, 8:53 doi: 10.1186/1472-6947-8-53
2. Vickers AJ, Elkin EB: Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006, 26(6):565-574.
3. Tsalatsanis et al.: A regret theory approach to decision curve analysis: A novel method for eliciting decision makers' preferences and decision making. BMC Medical Informatics and Decision Making 2010 10:51.

4. Марценюк В.П. Математичні моделі в системі підтримки прийняття рішень страхового забезпечення лікування онкологічних захворювань: підхід на основі динаміки Гомперца / В.П. Марценюк, І.Є. Андрушак, І.С. Гвоздецька, Н.Я. Климук // Доповіді Національної академії наук України. –2012. – №10. – С. 34-39.
5. Марценюк В.П. Підхід на основі актуарних математичних моделей до задач страхової медицини / В. П. Марценюк, І.Є. Андрушак, Н.Я. Климук // Медична інформатика та інженерія. Науково-практичний журнал. – 2010. – №4. – С. 85-87.
6. Марценюк В.П. О модели онкологического заболевания со временем пребывания на стадии в соответствии с распределением Гомперца / В.П. Марценюк, Н.Я. Климук // Проблемы управления и информатики. Международный научно-технический журнал . – 2012. – № 6. – С. 137-143.
7. Марценюк В.П., Семенець А.В. Медична інформатика. Інструментальні та експертні системи. – Тернопіль: Укрмедкнига, 2004. – 222 с.
8. Edwards W, Miles RFJ, von Winterfeldt D: Advances in decision analysis. From foundations to applications. New York: Cambridge University Press 2007.
9. Shannon C, Weaver W: The mathematical theory of communication. Urbana: The University of Illinois Press 1962.
10. Zimmer man H: An application-oriented view of modeling uncertainty. European Journal of Operational Research 2000, 122: 190-198.
11. Schurink CAM, Lucas PJF, Hoepelman IM, Bonten MJM: Computer-assisted decision support for the diagnosis and treatment of infectious diseases in intensive care units. The Lancet Infectious Diseases 2005, 5(5):305-312.