

# Застосування методу «дерева рішень» для побудови системи встановлення діагнозу

Басюк Тарас  
Кафедра ІСМ  
НУ “Львівська політехніка”  
Львів, Україна  
Taras.M.Basyuk@lpnu.ua

Дуб’як Ірина  
Кафедра ІСМ  
НУ “Львівська політехніка”  
Львів, Україна  
IraDubyak@gmail.com

*This article is dedicated to description of applying features of the “decision tree” method in a process of making a diagnosis. The current state of subject area studied has been analyzed, the main problem areas have been defined and functions of the designed diagnostic system have been determined. Analysis of the task solutions showed that in this case the effective one is the “decision tree” method, in particular classification tree – the subspecies according to the type of predicted indicator. The decision tree has been built and the conclusions about the feasibility of its application have been made.*

**Ключові слова:** діагностування, симптоми, захворювання, дерево рішень.

Проблема діагностування стану здоров'я є актуальною для кожного індивідууму з огляду як на значний рівень забруднення навколишнього середовища так і катаклізми й появу низки нових захворювань. При цьому, інтенсивний розвиток техніки сприяє впровадженню множини новітніх методів й технологій медичного діагностування, що вимагає використання інноваційних технологій [1]. В процесі діагностування стану здоров'я, який протягом багатьох століть був завданням лікарів, надбанням медичної галузі знань, все більшого значення набувають технічні рішення. Пояснюється це тим, що технічно стало можливо автоматизувати множину етапів діагностичного процесу. Стрімкий розвиток технологій надає можливості із створення системи встановлення діагнозу, яка надасть засоби із спрощення роботи лікарів, тим самим одержавши певну перемогу в боротьбі за точність діагнозу, заощадження часу і коштів [2].

В процесі створення системи встановлення діагнозу труднощі можуть виникати під час збору

експертних даних, які мають узгоджуватися та затверджуватися спеціалістами в галузі медицини та охорони здоров'я. Адже база знань повинна містити найбільш релевантні та достовірні дані, щоб встановити правильний діагноз та сформувати відповідне лікування. З огляду на те, необхідно створити сховище даних, що міститиме інформацію як про симптоми захворювань так і відповідні препарати [3].

У зв'язку із необхідністю постійної взаємодії користувача з значною кількістю даних та з метою чіткого розподілу ролей оптимальним рішенням побудови є клієнт-серверна архітектура, де в якості клієнта виступатиме мобільний пристрій з наявним мережевим підключенням. Основними функціями проектованої системи є: інкапсуляція окремих з'єднань в межах однієї сесії; встановлення діагнозу захворювання відповідно до вибраних симптомів; надання рекомендацій щодо лікування із переліком препаратів та курсу прийому; створення системи нагадувань щодо прийому ліків; збереження нагадувань в історії сповіщень. При цьому, основною функцією є встановлення правильного діагнозу. Проведений аналіз показав, що в якості методу вирішення для автоматичного аналізу поточних і альтернативних даних є їх представлення у формі дерева рішень. Даний метод передбачає значні можливості опрацювання та успішно застосовується для вирішення завдань у різних галузях, зокрема в медицині [4].

Дерево рішень являє собою послідовну ієрархічну структуру і включає в себе: *гілки з атрибутами*, від яких залежить результат – цільова функція; *вузли* – випадкові вершини, в яких визначено можливі варіанти розвитку подій;

кінцеві (листяні) вузли із значеннями цільової функції, які являють собою остаточні результати вибору певного значення атрибута і об'єднують кілька об'єктів [5].

Дерева рішень за типом прогнозованого показника поділяються на два види: дерева класифікації і дерева регресії. Розробляючи систему встановлення діагнозу доцільно використати дерева класифікації, оскільки вони застосовується у випадках, якщо необхідно провести дослідження по визначених атрибутах, а саме віднести об'єкти (симптоми) до наперед відомого класу (певного захворювання). Дерева рішень розбивають дані на групи, в результаті чого виникає ієрархія операторів "якщо..., то...", з допомогою яких класифікуються дані. Для побудови дерева класифікації доцільно використати алгоритм C4.5, в якого кількість нащадків вузла є не обмеженою. Алгоритм C4.5 використовує вхідну інформацію при створенні дерева рішень та потребує попереднього навчання. Для коректної роботи необхідно дотримуватись таких вимог: один з атрибутів набору даних повинен бути міткою класу, тобто ознакою, яка визначатиме клас об'єкта; мітки класу повинні бути дискретними [6].

Дерево будується на підставі навчальної вибірки, що містить інформацію про захворювання та відповідні їм симптоми (рис.1). На рисунку представлено спрощене дерево рішень, за допомогою якого діагностуються певні захворювання дихальної системи. В даній системі кореневий вузол – це температура, в ролі об'єктів виступають узагальнені симптоми, згруповані в певну категорію, в ролі атрибутів – конкретизовані симптоми, вибрані по місцю ураження або характеру перебігу захворювання. Підмножини, які не мають відношення до потрібного вирішення завдання, відсікаються, тим самим наближуючись до листкових вузлів – конкретних діагнозів.

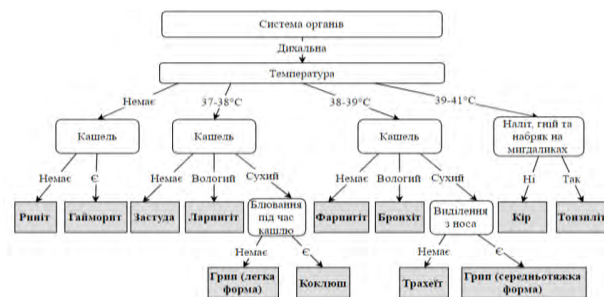


Рис. 1. Приклад дерева рішень

Проектування системи визначення діагнозу із використанням методу дерева рішень як альтернативного варіанту медичної консультації допоможе людям в будь-який момент часу отримати як інформацію щодо можливого захворювання так і отримати рекомендації щодо методів лікування.

## ЛІТЕРАТУРА

- [1] Перспективи розвитку медицини початку XXI століття [Електронний ресурс] // Режим доступу: <http://healthconomics.ru/item/3185-perspektivy-razvitiya-meditsiny-nachala-xxi-veka/3185-perspektivy-razvitiya-meditsiny-nachala-xxi-veka>. Останній перегляд: 2017. – Назва із заголовку екрану;
- [2] V. Gopinath, B. Naik, Y. Akhila, J. Pavani, "e-HDAS: e-Healthcare Diagnosis & Advisory System," IJRET: International Journal of Research in Engineering and Technology, vol. 2, pp.255-256, November 2013.
- [3] Абдікеев Н. М. Управление знаниями корпорации і реінженірінг бізнесу: Підручник/ Під наук. ред. д-ра техн. наук, проф. Н. М. Абдікеева. – М.: ІНФРА-М, 2011. – С. 382.
- [4] Левітін А. В. Обмеження потужності алгоритмів: Дерева прийняття рішень// Introduction to The Design and Analysis of Algorithms. – М.: «Вільямс», 2006. – С. 409-411.
- [5] S. Murthy, "Automatic construction of decision trees from data: A Multi-disciplinary survey," Data Mining and Knowledge Discovery, vol. 2, pp. 345-389, December 1998.
- [6] L. Breiman, and J. Friedman, R. Olshen, and C. Stone, "Classification and Regression Trees," Belmont, CA: Wadsworth International Group, 1984