SECTION 11.
SYSTEM ANALYSIS, MODELING AND OPTIMIZATION

Лобода Юлія Геннадіївна
кандидат педагогічних наук, доцент, доцент кафедри інформаційних технологій
Національний університет «Одеська юридична академія», Україна

Ключник Роман Євгенович
здобувач вищої освіти *факультету кібербезпеки та інформаційних технологій*
Національний університет «Одеська юридична академія», Україна

ПРОЦЕС ПОБУДОВИ МОДЕЛІ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ТА ОЦІНКА ПРОДУКТИВНОСТІ

Сьогодні інтелектуальний аналіз даних стає дедалі популярнішим у галузі медицини. Він відкриває великий потенціал, дозволяючи системам охорони здоров'я систематично використовувати дані та аналітику для виявлення недоліків та передових методів, які покращують прогнозування захворювань, надання медичної допомоги та знижують витрати страхових компаній.

Інтелектуальний аналіз даних є процес виявлення значних кореляцій, закономірностей та тенденцій шляхом дослідження великих обсягів даних, що зберігаються в репозиторіях. Інтелектуальний аналіз даних використовує технології розпізнавання образів, а також статистичні та математичні методи [1].

Для вирішення завдань діагностики та прогнозування у медицині, зокрема для прогнозування серцево-судинних захворювань, активно використовуються різні математичні методи та підходи. Метою інтелектуального аналізу даних є створення моделі, яка може ефективно інтерпретувати та використовувати дані, якими зараз володіють дослідники, та дані, які будуть отримані в майбутньому. Оскільки інтелектуальний аналіз даних включає кілька методів, основним етапом створення моделі даних є вибір методів аналізу, які будуть використовуватися в моделі.

Моделі інтелектуального аналізу даних можуть застосовуватись до конкретних сценаріїв, а саме: визначення ризику та ймовірностей діагнозу чи інших результатів, створення рекомендацій, прогнозування наступної можливої події, групування, аналіз та прогнозування спільних ознак. Створення моделей для покращення прогнозування є досить своєчасним.

Побудова моделі інтелектуального аналізу даних є лише частиною більшого процесу. Цей процес включає в себе ряд завдань: від формулювання питань, що стосуються даних, до створення моделі для відповіді на ці питання, і, нарешті, до розгортання цієї моделі в робочому середовищі. Процес побудови моделі інтелектуального аналізу даних можна як послідовність наступних базових кроків: збір даних про стан здоров’я пацієнтів, визначення проблеми та формулювання завдання, попередня обробка даних, вивчення даних, створення моделі, дослідження та перевірка моделі, розгортання та оновлення моделі.

Процес, є циклічним, тобто створення моделі даних є динамічним і повторюваним процесом. Виконавши перегляд даних, користувач може виявити, що даних недостатньо

для створення необхідних моделей інтелектуального аналізу даних, що веде до пошуку додаткових даних. Також може виникнути ситуація, коли після побудови кількох моделей виявиться, що вони не дають адекватної відповіді на поставлене завдання, і тому необхідно поставити завдання по-іншому. Тоді може виникнути потреба в оновленні вже розгорнутих моделей. Для створення гарної моделі може знадобитися багаторазово повторити кожний крок процесу.

Процес побудови моделі інтелектуального аналізу даних та оцінка продуктивності:

1. Попередня обробка даних. Етапи попередньої обробки даних включають: очищення даних, створення даних та скорочення даних. Дані можуть мати декілька розділів, які є незначними та відсутніми. Очищення даних виконується для обробки даних, які зашумлені, відсутні та інше. Перетворення даних виконується для перетворення даних у відповідну форму, придатну для інтелектуального процесу аналізу. Це включає генерацію нормалізації, збір атрибутів, дискретизацію і генерацію ієрархії. Інтелектуальний аналіз даних – це метод, який використовується для управління величезними обсягами даних. У разі аналіз стає складніше під час роботи з величезними обсягами даних. Щоб позбутися цього, використовується техніка скорочення даних, яка спрямована на підвищення ефективності зберігання та зниження витрат на зберігання та аналіз даних [2].

2. Вибір функції. Вибір функції складається в основному з чотирьох етапів: розробка підмножини, оцінка підмножини, критерій вибору та остаточна ознака підмножини.

3. Створення моделі. Модель отримує дані зі структури інтелектуального аналізу даних, а потім аналізує дані за допомогою алгоритмів інтелектуального аналізу даних. Модель інтелектуального аналізу даних зберігає дані статистичного процесу, як закономірності знайдені внаслідок аналізу. Кожен тип моделі створює свій набір шаблонів, наборів елементів, правил чи формул, які можна використовувати при прогнозуванні. Алгоритми, які можна використовувати у процесі розробки моделі: дерево рішень, нейронні мережі, логістична регресія тощо.

4. Оцінка продуктивності моделі. Існують різні підходи до оцінки продуктивності моделі. Найчастіше використовуваними критеріями виміру є: точність, чутливість, специфічність. Точність окреслюється відношення правильно класифікованих випадків. Чутливість вимірює співвідношення фактичних позитивних результатів, які правильно ідентифіковані. Специфіка правильно вимірює співвідношення фактичних негативних значень [3].

Підсумовуючи, можна сказати, що математичні моделі мають важливе значення та роль у галузі медицини. За допомогою математичних моделей ми можемо глибше зрозуміти природу та механізм захворювань, відкрити нові методи лікування та ліки, покращити профілактику та діагностику захворювань, а також мати широкі перспективи застосування в оптимізації розподілу та управління медичними ресурсами.

Список використаних джерел:

1. Gartner Glossary. Data Mining. URL: https://www.gartner.com/en/information-technology/glossary/data-mining
2. Лобода Ю.Г (2022). Особливості підготовки даних з метою подальшого аналізу. *Інформаційне суспільство: проблеми та перспективи*: матер. VII Всеукр. наук.-практ. конф. (м. Одеса, 20 травня 2022 р.). Одеса, 2022. С. 32-34. URL: https://doi.org/10.32837/11300.17736
3. Damen J A A G, Hooft L, Schuit E, Debray T P A, Collins G S, Tzoulaki I et al. Prediction models for cardiovascular disease risk in the general population: systematic review *BMJ* 2016; 353: i2416. DOI: https://doi.org/10.1136/bmj.i2416