

Застосування методу логістичної регресії в медичних дослідженнях

Скарга-Бандурова І.С., доц.; Фейгіна Д.І., студ.
Технологічний інститут Східноукраїнського національного
університету ім. В. Даля, м. Северодонецьк

Незважаючи на тисячоліття існування медицини, проблема організації збору, обробки та аналізу інформації, отриманої у процесі медичної діяльності, залишається однією з найбільш актуальних. Сучасна медична інформаційна система представляє собою великі розгалужені бази даних, в яких зберігаються терабайти інформації, далеко не завжди придатних для прийняття ефективних рішень. Вирішення цієї проблеми лежить у сфері інтелектуальної обробки великих масивів даних.

В роботі проведено огляд методів інтелектуального аналізу даних і виявлення знань в базах даних, проаналізовані програми та моделі процесів, застосовні до задач охорони здоров'я. Одним з таких методів є логістична регресія. Основною метою логіт-регресії є виявлення зв'язку між декількома незалежними змінними і залежною змінною. В аналізі даних принцип логістичної регресії дозволяє не тільки класифікувати спостережувані події, а також моделювати зв'язок між ними. Рівняння логіт-регресії має наступний вигляд (1):

$$p_i = 1 / (1 + e^{-z_i}), \quad (1)$$

де p_i – імовірність події, z_i - лінійна комбінація прогностичних параметрів (предикторів) (2).

$$z_i = a_0 + a_1x_{i1} + a_2x_{i2} + \dots + a_kx_{ik}, \quad (2)$$

де x_{ij} – значення j -го предиктора для i -го спостереження ($j = 1, \dots, k$; $i = 1, \dots, n$).

Метод логіт-регресії в задачах медичних досліджень дає можливість отримати модель, побудовану на зв'язку залежності ознаки захворювання з прогностичними характеристиками певної хвороби, що і є одним з найважливіших аспектів медицини, а саме, своєчасної діагностики захворювання у пацієнта. Іншим застосуванням методу, традиційно, є оцінка ризиків у медичній галузі.