

ЗАСТОСУВАННЯ МЕТОДУ КОРЕЛЯЦІЙНИХ ПЛЕЯД ДЛЯ КЛАСТЕРИЗАЦІЇ ДІАГНОСТИЧНИХ ОЗНАК

Максюта Н.В., Поворознюк А.І., Стешкін С.В.

Національний технічний університет

„Харківський політехнічний інститут”, м. Харків

Робота присвячена актуальному питанню – відбору діагностично цінних ознак відносно визначених захворювань з метою підвищення надійності розпізнавання діагнозу при побудові комп'ютерних систем медичної діагностики.

У роботі приведений порівняльний аналіз існуючих методів кластеризації ознак та методів відбору діагностично цінних з них. Розроблено структуру підсистеми кластеризації діагностичних ознак, алгоритм її функціонування та її програмну реалізацію. Кластеризація ознак виконана на основі методу кореляційних плеяд, за допомогою якого отримані групи ознак з урахуванням внутрішнього зв'язку між ними. При цьому метод кореляційних плеяд дозволив наочно представити структуру взаємозв'язків ознак у вигляді графу. Вершинами графа є самі ознаки, а дуги характеризують ступінь зв'язку між ними, у якості ваги дуги виступає значення коефіцієнта кореляції. Діагностична цінність ознак порахована за допомогою теоретико-інформаційного підходу. При розробці програмної реалізації підсистеми кластеризації діагностичних ознак авторами запропонований алгоритм відображення результатів кластеризації, а саме зручне розміщення вершин графу на екрані монітора.

Перевагою роботи є комплексне використання методу кластеризації і методу відбору діагностично цінних ознак. Тобто спочатку вихідний простір ознак розподілюється по групах, які містять сильно корельовані ознаки, потім у кожній групі обирається діагностично цінна ознака. Це дозволило сформувати простір незалежних діагностично цінних ознак відносно визначених захворювань та, в свою чергу, підвищити надійність розпізнавання діагнозу, що підтверджується результатами експериментальної перевірки на реальних даних. Вихідна вибірка містить 407 пацієнтів з результатами обстеження: значення показників клінічного аналізу крові, реологічних показників та діагноз (один із восьми). При цьому виконано порівняльний аналіз надійності розпізнавання діагнозів, використовуючи ознаки, обрані лише за теоретико-інформаційним підходом, та ознаки, обрані всередині сформованих груп. У останньому випадку надійність розпізнавання вище на п'ять відсотків, що свідчить про ефективність розробленого підходу.