

2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Москва : Техносфера, 2012. 1072 с.
3. Burger W., Burge M. Principles of Digital Image Processing. Fundamental Techniques. London : Springer-Verlag, 2009. P. 261.
4. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в Matlab. Москва : Техносфера, 2006. 616 с.
5. Власенко В. П., Саваневич В. Е., Погорелов А. В., Брюховецкий А. Б. Яркостное выравнивание астрономических изображений с использованием медианного фильтра. Радиотехника. 2016. № 185. с. 70-80.
6. Лукьяница А. А., Шишкин А. Г. Цифровая обработка изображений. Москва : Ай-Эс-Эс Пресс, 2009. 518 с.
7. CoLiTec - Collection Light Technology. URL: http://www.neoastrosoft.com/category/news/?lang=en_us (дата звернення: 10.12.2021).
8. FrameSmooth software - new tool for the calibration of astronomical images. URL: http://var.kozmos.sk/files/bezovec_2016.pdf (дата звернення: 11.12.2021).

Нагорний П.В., студент групи ПІ-181

Національний університет «Чернігівська політехніка», inn5665@gmail.com

Науковий керівник: Акименко А.М., канд. фіз.-мат. наук

Національний університет «Чернігівська політехніка», anakim2@gmail.com

ВИКОРИСТАННЯ АЛГОРИТМІВ DATA MINING В МЕДИЦИНІ

Сучасне суспільство характеризується стрімким зростанням обсягів інформації, яка циркулює в навколишньому інформаційному просторі. Основним фізичним представленням, засобом вираження інформації є дані. На більш ранніх етапах розвитку людства обсяги даних, які необхідно було обробляти для прийняття рішень, були незначні та дозволяли навіть їх інтуїтивно-логічний аналіз без використання цифрових засобів. В наш час традиційні алгоритми виявилися неефективними, тож актуальними є завдання розробки та інтеграції нових методів аналізу даних. Вирішення цих завдань розглядається в теоретико-прикладній науці Data Mining (фактично, наука про дані та їх обробку). На даний момент розроблено вже значну кількість методів обробки даних, які застосовні в різних областях людського життя. Розглянемо використання алгоритмів Data Mining в медицині.

Основна сфера застосування Data Mining в медицині – підсумування та узагальнення первинних даних симптоматики пацієнтів з метою визначення хвороби та конкретних ефективних методів її лікування. Перш за все, необхідно визначитися з основними джерелами даних для Data Mining в медицині. До таких джерел належать [1]:

- клінічні дані пацієнтів;
- показники давачів медичного обладнання;
- конкретні числові характеристики, встановлені методом експертного аналізу;
- аудіо- та візуально-графічні об'єкти;
- показники ефективності фармацевтичних препаратів;
- дані невідкладної допомоги.

Тепер визначимося з основними алгоритмами обробки великих даних, застосовними в контексті медичних досліджень. Виокремлюють [2] наступні методи: алгоритм С4.5, метод опорних векторів, EM-алгоритм, алгоритм AdaBoost, наївний байєсовський класифікатор, метод k-середніх, алгоритм Apriori, алгоритм PageRank, алгоритм k-найближчих сусідів (kNN), алгоритм CART. Можливості зазначених методів розглянемо на прикладі алгоритму С4.5.

Алгоритм С4.5 дозволяє побудувати класифікатор у вигляді дерева рішень. Метод оперує атрибутами – числовими характеристиками певних параметрів. Приклади атрибутів в

медицині: температура, сила кашлю, гематомегалія тощо. Для кожного пацієнта визначається певна якісна ознака (наприклад, «хворий на ОРВІ»). За відомим набором даних алгоритм дозволяє визначати якісну ознаку для нових пацієнтів шляхом побудови дерева рішень (рис. 1).



Рис. 1. Схема побудови дерева рішень за алгоритмом C4.5

До основних особливостей алгоритму C4.5 можна віднести [3]:

- ✓ використання обновлюваної інформації, що робить алгоритм більш гнучким;
- ✓ замість перенавчання використовується однопрохідне проріджування;
- ✓ підтримка як дискретних, так і неперервних значень.

Даний алгоритм можна реалізувати за допомогою різних мов програмування. В літературі наводиться наступна реалізація на мові програмування python [4]:

```

import math
import utils
def cnt(table, column, variant):
    #повернення кількісних варіантів з таблиці
    return table[column].count(variant)
def entrop(table, res_column):
    #розрахунок початкової ентропії таблиці
    s = 0
    for variant in utils.deldup(table[res_column]):
        p = cnt(table, res_column, variant) / float(len(table[res_column]))
        s += p*math.log(p, 2)
    return -s
def entropAft(table, column, res_col):
    #розрахунок ентропії таблиці після поділу на підтаблиці
    s = 0
    for subt in utils.get_subtables(table, column):
        s += (float(len(subt[column])) / len(table[column])) * info(subt, res_col)
    return s
def gain(table, □, res_column):
    #критерій розбиття
    return entrop(table, res_column) - entropAft(table, □, res_column)
    
```

Таким чином, використання алгоритму C4.5 є ефективним в медицині в цілях розробки класифікатора даних. Основні переваги алгоритму: гнучкість, швидкість роботи, підтримка різних форматів даних. C4.5 – не єдиний алгоритм Data Mining, ефективний в медицині. Серед інших: метод опорних векторів, алгоритм Apriori, метод k-середніх тощо.

Перелік посилань

1. Шумейко А. А., Сотник С. Л. Интеллектуальный анализ данных (Введение в Data Mining): учеб. пособ. Днепропетровск: Белая Е. А. 2012. 212 с.
2. Костюкова Н. И., Кудинов А. Е. Автоматизация научных исследований в области медицины с применением технологии mining. *Альманах современной науки и образования*. 2010. № 3 (34). Ч. 1. С. 22-24.
3. Зелинский С. С., Удуд Е. А., Новрузова В. И. Big Data в медицине: направления использования. Стимулирование инновационного развития общества в стратегическом периоде : сб. статей Междунар. науч.-практ. конф. (Челябинск, 5 ноября). Уфа. 2018. С. 176-178.
4. Зелинский С. С., Зелинская С. А., Базян Д. Е., Удуд Е. А. Возможности использования алгоритмов Big Data Mining в медицине. Научное пространство: актуальные вопросы, достижения и инновации : сб. статей XVII Междунар. науч.-практ. конф. (Анапа, 18 апреля). Анапа. 2020. С. 155-160.

Нагорний П.В., студент групи ПІ-181

Національний університет «Чернігівська політехніка», inn5665@gmail.com

Науковий керівник: Злобін С.В., викладач циклової комісії дисциплін професійної підготовки спеціальностей 122, 123

ВСП «Фаховий коледж транспорту та комп'ютерних технологій» Національного університету «Чернігівська політехніка», s.zlobin75@gmail.com

МЕТОДИКА ВИБОРУ ВИДУ ТЕСТУВАННЯ НА ОСНОВІ ЯКІСНОЇ ОЦІНКИ РОЗРОБЛЮВАНОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Сучасний стрімкий розвиток цифрових технологій, активне їх впровадження в усі сфери життя людини актуалізують питання якості пропонованого програмного забезпечення (ПЗ). З точки зору користувача, якість ПЗ – відповідність його призначенню. З точки зору розробника – відповідність вимогам. В той же час, внаслідок існування великої кількості екзогенних та ендогенних чинників, якість ПЗ не завжди відповідає очікуваному рівню. Єдине можливе рішення поліпшення якості ПЗ – корекція знайдених проблем. Але для цього необхідно достатню увагу приділяти тестуванню ПЗ.

В наш час найбільш популярними видами тестування є ручне, автоматизоване, змішане. З першого погляду може здатися, що автоматизоване тестування у всіх аспектах краще за ручне та змішане, адже має чіткі переваги: економія часу, більше тестове покриття, зменшення ризику виникнення помилок тестування. Проте автоматизоване тестування потребує значно більших ресурсів та часто окупається лише після кількох релізів ПЗ. Саме тому в багатьох випадках, особливо коли мова йде про невеликі проекти, розробники обирають ручне чи змішане тестування [1].

Розглянемо методику обрання виду тестування на основі якісної оцінки ПЗ [2]. Основна ідея методики полягає в експертній оцінці різноманітних видів тестування. Підстава оцінювання – розроблений для конкретного ПЗ перелік конкретних питань, кожне з яких передбачає чіткий варіант відповіді «так» чи «ні». Питання розподілені по трьом групам. В першій групі позитивна відповідь на питання надає більше підстав для автоматизованого тестування. В другій групі – для ручного тестування. В третій групі – для змішаного тестування. В [3] детально охарактеризовано принцип побудови таких питань. Тому обмежимося лише наведенням 3-х прикладів таких питань з кожної групи.

1-ша група:

- ❖ чи наявний в ПЗ функціонал, який часто повторюється?
- ❖ чи передбачається регулярний вихід нових версій ПЗ?
- ❖ чи передбачається експлуатація ПЗ за максимального навантаження?