**ЗАСТОВУВАННЯ АЛГОРИТМУ ЗНАХОДЖЕННЯ МІНІМАЛЬНОЇ ВІДСТАНІ ДЛЯ ВИЗНАЧЕННЯ ГРУПИ РИЗИКУ СТУДЕНТА**

**СИСТЕМА ОЦІНКИ ФУНКЦІОНАЛЬНОГО СТАНУ КРОВООБІГУ СТУДЕНТІВ (ТА ВИКЛАДАЧІВ) ТЕХНІЧНОГО УНІВЕРСИТЕТУ ІЗ ЗАСТОСУВАННЯМ ПРОБИ МАРТІНЕ**

**Войник Б.О., Борисова Г.В.**

*Студенти 5-го курсу НТУУ «КПІ ім. Ігоря Сікорського»*

APPLICATION OF AN ALGORITHM FOR DETECT A MINIMUM DISTANCE FOR DETERMINATION OF THE STUDENT'S RISK GROUP

**Voinyk B.O., Borisova G.V.**

*Students of the 5th year of NTUU «Igor Sikorsky KPI»*

**АНОТАЦІЯ**

Визначено групи ризику студентів за допомогою алгоритму знаходження мінімальної відстані на базі отриманих результуючих таблиць. Отримано мітки для кожного студента, що характеризують субгрупу ризику. Встановлено вектор направлення студента.

**Ключові слова:** кластеризація, патерн, мінімальна відстань, квадрат евклідової відстані, функціональні проби

**ABSTRACT**

The students' risk groups were determined using the algorithm of finding the minimum distance based on the resulting tables. The labels for each student that characterize the subgroup of risk are obtained. The direction of student direction is set.

**KEYWORDS**: clustering, pattern, minimum distance, square of the Euclidean distance, functional tests

**Вступ**

Систематичні самостійні спостереження за станом свого здоров’я є необхідним кроком для поліпшення функціонування систем організму та підтримки його в тонусі. Виконання фізичних вправ позитивно впливає на організм студента і дає змогу провести аналіз фізичних можливостей. За допомогою функціональної проби Мартіне можна дослідити динаміку зміни артеріального тиску та пульсу в період між станом спокою та на кожній хвилині після навантаження протягом п’яти хвилин. Оскільки стан фізичного здоров’я може змінюватися незалежно від фізичних навантажень, то регулярне проведення проби Мартіне дасть змогу спостерігати динаміку змін в організмі за певний період [1,4]

Це потребує розробки програмного продукту, який би реєстрував дані та автоматично проводив би дослідження стану фізичного здоров’я.

**Метою роботи** є визначення групи та субгрупи ризику студентів за допомогою квадрату евклідової відстані на базі результуючих таблиць, побудованих на навчальній базі

**Матеріали та методи дослідження**

Для дослідження було використано базу даних студентів Національного технічного університету України “Київський політехнічний інститут ім. Ігоря Сікорського”, створену у процесі досліджень кафедрою фізичного виховання. База даних містить 599 спостережень, з яких 276 жіночої статі та 323 чоловічої,

а також 353 кількісних показників, що відображають фізичний стан пацієнтів.

Всі студенти мали медичну довідку про можливість проходження занять з фізичного виховання, а також давали згоду на проходження тестування.

Всі студенти проходили пробу Мартіне, яка полягала у виконанні 20 присідань за 30 секунд.

Безпосередньо перед початком проби і щохвилинно протягом п’яти хвилин вимірювались показники систолічного та діастолічного артеріального тиску і частоти серцевих скорочень (ЧСС).

Окремо для юнаків та дівчат для визначення функціональних патернів, тобто співвідношень вказаних вище показників, застосовувався кластерний аналіз методом «k-середніх». На попередньому етапі досліджень було отримано 7 кластерів у юнаків та 8 кластерів у дівчат, які істото відрізнялись один від одного за характером реакції показників артеріального тиску (АТ) та ЧСС на виконання проби Матіне.

Для формулювання формального опису особливостей кожного кластера було проаналізовано 353 кількісних показників, що відображають психофізичний стан пацієнтів.

Крім того, для кожного кластера обчислювався його середній радіус:

*i-*гопоказника

В якості критерію близькості показників окремого студента до центра певного кластера ми використовували евклідову відстань:

– значення *i*-го показника студента

Обчисленню відстані передувало нормування даних (на максимальний розмах) для забезпечення рівноправності участі показників у процедурах обчислення відстані.

**Результати дослідження**

Для дослідження було використано базу, що містить показники артеріального тиску та ЧСС студентів, які пройшли пробу Мартіне більше одного разу [2].

Взаємне розташування центроїдів кластерів після виконання проби Мартіне, які використовувались в якості функціональних патернів реакції на навантажувальну пробу, представлене на рис. 1.

353 кількісних показників, що відображають фізичний стан пацієнтів.

Рисунок 1. Зміни показників артеріального тиску до та протягом 5 хвилин після проби Мартіне. Чоловіки

Рисунок 2. Зміни показників артеріального тиску до та протягом 5 хвилин після проби Мартіне. Чоловіки

Рисунок 3. Зміни показників артеріального тиску до та протягом 5 хвилин після проби Мартіне. Жінки

Рисунок 4. Зміни показників артеріального тиску до та протягом 5 хвилин після проби Мартіне. Жінки

Був розроблений та застосований наступний алгоритм.

1. Визначалась мінімальна евклідова відстань до декотрого кластеру.
2. Визначалась субмінімальна евклідова відстань до іншого кластеру, яка перевищувала мінімальну відстань, але була меншою за відстані до всіх інших кластерів. Тобто для оцінки стану студента визначались дві найменші відстані до найближчих кластерів.
3. Визначалось, чи розташувались показники даного тесту всередині середнього радіуса кластера («так», або «ні»).
4. Якщо «так», то на термінал виводилась інформація щодо функціональних та психофізичних особливостей цього, найближчого, кластера.
5. Якщо «ні», то на термінал виводилась інформація щодо функціональних та психофізичних особливостей кластера з субмінімальною відстанню до нього.

Блок-схема наведеного вище алгоритму представлена на рис. 5.



Рисунок 5. Блок-схема алгоритма генерації заключення щодо поточного функціонального стану студента.

Таким чином, розроблений нами алгоритм дозволяє враховувати додаткові можливі особливості функціонального стану організму при суттєвому відхиленні показників проби Мартіне від певного функціонального патерна, тобто центроїда кластера.

Процедура є достатньо економічною в обчислювальному сенсі.

Представлений алгоритм дозволяє визначати стани кровообігу, що відповідають зниженню регуляторних резервів та істотному підвищенню артеріального тиску. Це дозволяє контролювати індивідуальний рівень фізичних навантажень, змінювати програму тренувань та виявляти стани, що вимагають додаткового медичного контролю.

Крім того, можливе застосування системи для скринінгового контролю великих груп студентів для виявлення осіб з високим ризиком розвитку раптових розладів кровообігу.

Для кластеризації студентів і визначення їх субгрупи ризику ми використали програмний продукт «ClusterBox», попередньо модифікувавши його, додавши функцію глобального режиму, здатну розставити кластери та мітки (субгрупи ризику) не для одного студента, а для всіх одразу, що присутні в базі даних.



Рисунок 1 – Вікно досліджень програмного продукту «ClusterBox»

Робота програмного продукту побудована на базі алгоритму знаходження мінімальної відстані. Цей алгоритм передбачає використання квадрату евклідової відстані, для якого необхідним є побудова результуючих таблиць та наявність в базі даних показників артеріального тиску та пульсу студента в стані спокою і протягом п’яти хвилин після навантаження.

Побудова результуючих таблиць відбувається на базі проведення дисперсійного аналізу. Після нього дані групуються до компактної таблиці за кожною змінною (АТС0, АТД0, ЧСС0 і т.д.) і для кожного кластеру окремо. Слід зазначити, що результуючі таблиці для дослідження і кластеризації будуються один раз на тих даних, для яких патерни вже відомі. Таким чином, база даних з відомими патернами, для якої будуються результуючі таблиці, є навчаючою. В нашому випадку навчаючою була база даних студентів 1-2 курсу Національного технічного університету України “Київський політехнічний інститут ім. Ігоря Сікорського”, що містить 1495 спостережень, з яких 669 жіночої статі та 826 чоловічої. Кластерний аналіз був проведений Настенко Є.А. та Носовець О.К., згідно з їхніми даними всіх чоловіків було поділено на 7 кластерів, а жінок – на 8 [3].

Отже, маючи необхідні дані, ми можемо застосувати програмний продукт для визначення групи ризику студента. Алгоритм наведено на рисунку 2.



Рисунок 2 – Блок-схема алгоритму знаходження мінімальної відстані до кластеру [5]

Алгоритм передбачає знаходження мінімальної відстані, що в свою чергу характеризує кластер, до якого відноситься студент. Субмінімальна відстань (наступна відстань після мінімальної) в свою чергу буде характеризувати мітку студента. Знаючи кластер та мітку ми можемо встановити вектор направлення студента.

Результатом глобального режиму програми є створення текстових файлів формату «\*.txt», в яких зберігаються дані кластерів та міток для кожного студента відповідно. Дані додаються до загальної бази даних для подальшого дослідження. Таким чином дослідження показали скільки студентів відноситься до певного кластеру.

Таблиця 1 – кількість студентів в кластері. Чоловіки

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Кл-р 1 | Кл-р 2 | Кл-р 3 | Кл-р 4 | Кл-р 5 | Кл-р 6 | Кл-р 7 |
| 40 | 67 | 45 | 50 | 6 | 11 | 103 |

Таблиця 2 – кількість студентів в кластері. Жінки

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Кл-р 1 | Кл-р 2 | Кл-р 3 | Кл-р 4 | Кл-р 5 | Кл-р 6 | Кл-р 7 | Кл-р 8 |
| 58 | 16 | 35 | 13 | 8 | 41 | 51 | 46 |

**Висновки**

Нами було модифіковано програмний продукт «ClusterBox», додавши до нього функцію глобального режиму.

Отримані за допомогою дисперсійного аналізу результуючі таблиці дали змогу реалізувати алгоритм знаходження мінімальної відстані до кластеру на базі квадрату евклідової відстані.

За допомогою програмного продукту було кластеризовано базу даних та визначено вектор напрямку для кожного студента. Модифіковано початкову базу даних студентів, додавши до неї колонки з кластерами та мітками.

Майбутні дослідження будуть присвячені пошуку більших відмінностей між кластерами для розширення їх характеристик. Нашим завданням також буде пошук закономірностей між кластерами та мітками студентів, які пройшли пробу Мартіне більше одного разу для можливості прогнозування їх групи ризику.

**Список літератури**

1. Брехман И.И. Валеология – наука о здоровье. – М.: ФиС, 1990. – 208с.
2. Оценка функционального состояния и адаптивных возможностей организма у студентов вуза в процессе занятий физической культурой: методические указания для преподавателей физической культуры и студентов / сост. С.Е. Бебинов. – Омск: Изд-во СибАДИ, 2004. – 16 c.
3. Побудова результуючих таблиць для реалізації алгоритму знаходження мінімальної відстані до патернів на базі проведення дисперсійного аналізу [Електронний ресурс] – Режим доступу до ресурсу: http://tsh-journal.com/wp-content/uploads/2017/08/VOL-1-No-14-14-2017.pdf
4. Самоконтроль у фізичному вихованні школярів [Електронний ресурс] – Режим доступу до ресурсу: http://ua.textreferat.com/referat-552-1.html
5. Создание простой блок-схемы [Електронний ресурс] – Режим доступу до ресурсу: https://support.office.com/en-US/article/Basic-tasks-in-Visio-B58073C2-12C8-4981-AD7E-235066FA910D