

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО

Факультет інформатики та обчислювальної техніки
(назва факультету, інституту)

Кафедра автоматизованих систем обробки інформації і управління
(назва кафедри)

"На правах рукопису"
УДК 004.93

«До захисту допущено»
В.о.завідувача кафедри
О.А.Павлов
(підпис) (ініціали, прізвище)
“ ” 20 19 р.

МАГІСТЕРСЬКА ДИСЕРТАЦІЯ

на здобуття ступеня магістра

за спеціальністю 126 Інформаційні системи та технології
(код та назва спеціальності)

ОПП Інформаційні управляючі системи та технології
(код та назва спеціалізації)

на тему: Інтелектуальний аналіз аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань

Виконав: студент VI курсу групи ІС-82мп
(шифр групи)

Григорович Віталій Віталійович
(прізвище, ім'я, по батькові) (підпис)

Науковий керівник доцент, к.т.н., доцент, Баклан І. В.
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Консультант доцент, к.т.н., доцент, Жданова О.Г.
(науковий ступінь, вчене звання, прізвище, ініціали) (підпис)

Рецензент доц каф. АУТС, к.т.н., доц Полторак В.П.
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Засвідчую, що у цій магістерській дисертації немає записань з праць інших авторів без відповідних посилань.

Студент _____
(підпис)

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО»**

Факультет інформатики та обчислювальної техніки
(повна назва)

Кафедра автоматизованих систем обробки інформації та управління
(повна назва)

Рівень вищої освіти другий (магістерський) за освітньо-професійною програмою

Спеціальність 126 Інформаційні системи та технології
(код і назва)

ОПП Інформаційні управляючі системи та технології
(код і назва)

ЗАТВЕРДЖУЮ
Завідувач кафедри
О.А.Павлов
(підпис) (ініціали, прізвище)

« » грудня 2019 р.

**ЗАВДАННЯ
на магістерську дисертацію студенту**

Григорович Віталій Віталійович
(прізвище, ім'я, по батькові)

1. Тема дисертації Інтелектуальний аналіз аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань

науковий керівник дисертації к.т.н., доц. Баклан І.В.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від “ 28 ” жовтня 20 19 р. № 3770-с

2. Строк подання студентом дисертації “ 2 ” грудня 20 18 р.

3. Об'єкт дослідження Інтелектуальний аналіз ритмів електрокардіограм та виявлення аномалій за допомогою спеціалізованих баз знань

4. Перелік завдань, які потрібно розробити виконати огляд існуючих методів та алгоритмів виявлення аномалій; розглянути основні принципи нечіткої логіки та систем нечіткого розбиття; провести аналіз існуючих методів кластеризації, методів їх навчання і самонавчання, що використовуються для вирішення завдань нечіткої кластеризації даних; спроектувати та розробити власну систему виявлення виявлення аномалій

5. Орієнтовний перелік ілюстративного матеріалу Скатерограми центроїди для кожної з груп; Карта зон ризику захворювань; Методика розрахунку; Роз'яснення карти зон ризику захворювань; Схема перетворення вхідних даних;

6. Орієнтовний перелік публікацій

Тези доповіді у матеріалах III всеукраїнській науково-практичній конференції молодих вчених та студентів «Інформаційні системи та технології управління» (ІСТУ-2019)

Стаття у міжнародній конференції «Розвиток науки в XXI столітті»

15 листопада 2019р, Харків

7. Консультанти розділів дисертації

| Розділ | Прізвище, ініціали та посада консультанта | Підпис, дата | |
|--------|---|----------------|------------------|
| | | завдання видав | завдання прийняв |
| | | | |
| | | | |
| | | | |

8. Дата видачі завдання “ 2 ” вересня 20 19 р.

Календарний план

| № з/п | Назва етапів виконання магістерської дисертації | Строк виконання етапів магістерської дисертації | Примітка |
|-------|---|---|----------|
| 1 | <i>Систематизація результатів огляду літератури</i> | <i>9.09.2019</i> | |
| 2 | <i>Порівняльний аналіз існуючих методів розв'язання задачі</i> | <i>16.09.2019</i> | |
| 3 | <i>Постановка та формалізація математичної моделі задачі</i> | <i>23.09.2019</i> | |
| 4 | <i>Модифікація існуючих методів розв'язання задачі</i> | <i>25.09.2019</i> | |
| 5 | <i>Розробка інформаційного та програмного забезпечення</i> | <i>1.10.2019</i> | |
| 7 | <i>Проведення експериментальних досліджень розроблених алгоритмів</i> | <i>30.10.2019</i> | |
| 8 | <i>Оформлення документації</i> | <i>11.11.2019</i> | |
| 9 | <i>Подання роботи на попередній захист</i> | <i>20.11.2019</i> | |
| 10 | <i>Подання роботи на основний захист</i> | <i>02.12.2019</i> | |

Студент

(підпис)

В.В. Григорович

(ініціали, прізвище)

Науковий керівник

(підпис)

В.І. Баклан

(ініціали, прізвище)

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

IoT (Internet of Things). – Інтернет речей.

SDK (Software Development Kit) – набір із засобів розробки, утиліт і документації.

USB (Universal Serial Bus) – універсальна послідовна шина.

IDE (Integrated Development Environment) – інтегроване середовище розробки.

GPIO (General-purpose input/output) – Інтерфейс введення/виведення загального призначення.

UART (Universal Asynchronous Receiver-Transmitter) – універсальний асинхронний приймач/передавач.

ADC (Analog-to-digital) – аналого-цифровий перетворювач.

Data Mining – видобування інформації.

K-means – алгоритм К-середніх.

ЕКГ – електрокардіограма.

РЕФЕРАТ

Магістерська дисертація: 124 с., 27 рис., 1 додаток, 24 табл., 81 джерел.

Актуальність. Дуже великою за масштабами проблемою, з приводу питання людського здоров'я, є невчасне визначення порушень, тобто виявлення їх не на початкових періодах, при яких достатньо всього лиш звернутися до лікаря та використати набагато простіший вид лікування хвороби аби уникнути проблеми зі станом здоров'ям, а в той період коли захворювання вже почало впливати на весь організм та на всю систему людських органів.

Актуальність задачі зумовлена потребою виявлення серцево - судинних захворювань у їх початковий період, щоб забезпечити просте, швидке та набагато діюче лікування пацієнта, а також забезпечити своєчасний контроль хронічних хвороб, що у свою чергу має покращити ступінь охорони здоров'я та зменшити витрати на непотрібні заходи для нього.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась на кафедрі автоматизованих систем обробки інформації та управління Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського» в рамках теми *«Інтелектуальний аналіз аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань»*.

Мета дослідження і завдання дослідження. Метою дисертаційної роботи є покращення якості постановки діагнозів пацієнтам з серцево-судинними захворюваннями за рахунок розробки онлайн системи інтелектуального аналізу ритмів електрокардіограм для виявлення аномалій у роботі серцево-судинної системи.

Задачі дослідження:

- аналізу існуючих підходів та методів кластеризації ритмів ЕКГ;
- вибір оптимального алгоритму кластеризації даних у вигляді часових рядів;
- проектування власної системи виявлення аномалій у ритмах ЕКГ;
- проведення аналізу методів та підходів згортання даних;

– розробка власного програмного забезпечення аналізу ритмів ЕКГ та спеціалізованої бази даних до неї;

– експериментування базовані на основі реальних даних ритмів ЕКГ.

Об’єкт дослідження – процес інтелектуального аналізу даних та аномалій у ритмах ЕКГ.

Предмет дослідження – метод інтелектуального аналізу даних часових рядів за допомогою кластеризації та прогонки через нейронну мережу з навчанням без вчителя.

Методи дослідження, використані в роботі, засновані на методах кластерного аналізу, методу головних компонент, та нейронної мережі.

Наукова новизна одержаних результатів полягає у розробці онлайн системи аналізу ритмів електрокардіограм та виявлення аномалій в них. Розроблений підхід до обробки даних надає можливість проаналізувати вхідні дані у вигляді часових рядів, виявити аномалії в них, та передбачити можливі захворювання пацієнта. Підхід базується на поєднанні методів кластерного аналізу та методу головних компонент.

Публікації. Тезисні матеріали опубліковані у III всеукраїнській науково-практичній конференції молодих вчених та студентів «Інформаційні системи та технології управління» (ІСТУ-2019). Опублікована стаття у міжнародній конференції «Розвиток науки в XXI столітті», 15 листопада 2019р, Харків.

ЕКГ, RR-ІНТЕРВАЛИ, АНОМАЛІЯ, СКАТЕРОГРАМА, КЛАСТЕРИЗАЦІЯ, МЕТОД ГОЛОВНИХ КОМПОНЕНТ, K-MEANS.

ABSTRACT

Master's Thesis: 124 with., 27 Fig., 1 app, 24 tables, 81 sources.

Topicality. A very large scale problem with regard to human health issues is the untimely identification of disorders, that is, the detection of them not in the initial periods, in which it is enough to just consult a doctor and use a much simpler form of treatment of the disease to avoid health problems, and at that time when the disease had already begun to affect the whole organism and the entire system of human organs.

The urgency of the task is due to the need to detect cardiovascular diseases in their initial period, to provide simple, fast and much effective treatment of the patient, as well as to provide timely control of chronic diseases, which in turn should improve the degree of health care and reduce the costs of unnecessary measures. for him.

Relationship of work with scientific programs, plans, themes. The work was performed at the Department of Automated Information Processing and Management Systems of the National Technical University of Ukraine «Kyiv Polytechnic Institute. Igor Sikorsky» within the topic “Intellectual analysis of electrocardiogram rhythm anomalies using specialized knowledge bases”.

The purpose and the objectives of the study. The aim of the dissertation is to improve the quality of diagnosis of patients with cardiovascular diseases by developing an online system for the analysis of electrocardiogram rhythms for detecting anomalies in the cardiovascular system.

Research objectives:

- analysis of existing approaches and methods of ECG rhythm clustering;
- choosing the optimal algorithm for clustering data in the form of time series;
- design of the wax system for detecting anomalies in ECG rhythms;
- conducting analysis of data collapsing methods and approaches;
- development of own software for analysis of ECG rhythms and specialized database for it;
- Experiments are based on real ECG rhythm data;

The object of study is the process of mining the flow of data in the form of time series.

The subject of the study is a method of intellectually analyzing time series data by clustering and routing through a neural network with non-teacher training.

The research methods used in the work are based on cluster analogue method, principal component method, and neural network.

The scientific novelty of the obtained results is the development of an online system for the analysis of rhythms of electrocardiograms and detection of anomalies in them. The developed approach to data processing provides an opportunity to analyze the input data in the form of time series, to detect anomalies in them, and to predict possible diseases of the patient. The approach is based on the combination of cluster analysis and principal component methods.

Publications. Theses have been published in the Third All-Ukrainian Scientific and Practical Conference of Young Scientists and Students "Information Systems and Technologies of Management" (ISTU-2019). Published an article in the international conference "The development of science in the XXI century", November 15, 2019, Kharkiv.

ECG, RR-INTERVALS, ANOMALY, SCATTERGRAM, CLUSTERING, PRINCIPAL COMPONENT ANALYSIS, K-MEANS.

ЗМІСТ

| | |
|--|------------|
| ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ | 4 |
| РЕФЕРАТ | 5 |
| ABSTRACT..... | 7 |
| ВСТУП | 11 |
| 1 ПРОЕКТНІ РІШЕННЯ З РОЗРОБКИ СИСТЕМИ..... | 14 |
| 1.1 Опис бізнес – процесів | 14 |
| 1.1.1 Опис процесу діяльності | 14 |
| 1.1.2 Актори і функції | 24 |
| 1.1.3 Структура бізнес-процесів..... | 29 |
| 1.2 Опис постановки задачі | 37 |
| 1.3 Рішення з інформаційного забезпечення..... | 38 |
| Висновки до розділу | 40 |
| 2 МОДЕЛІ ТА МЕТОДИ РОЗВ'ЯЗАННЯ ЗАДАЧІ | 41 |
| 2.1 Змістовна постановка задачі..... | 41 |
| 2.2 Математична модель | 42 |
| 2.3 Огляд методів розв'язання..... | 56 |
| 2.4 Розробка методу розв'язання задачі..... | 71 |
| 2.5 Розробка алгоритму розв'язання | 77 |
| 2.6 Результати досліджень ефективності методу | 78 |
| Висновки до розділу..... | 79 |
| 3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ | 80 |
| 3.1 Засоби розробки..... | 80 |
| 3.2 Вимоги до технічної частини | 84 |
| 3.3 Вимоги до програмного продукту..... | 84 |
| 3.4 Архітектура програмного забезпечення | 85 |
| 3.5 Інструкція користувача | 85 |
| 3.6 Випробовування програмного продукту | 89 |
| Висновки до розділу..... | 91 |
| 4 РОЗРОБКА СТАРТАП-ПРОЕКТУ | 92 |
| 4.1 Опис ідеї проекту..... | 92 |
| 4.2 Технологічний аудит ідеї проекту..... | 93 |
| 4.3 Аналіз ринкових можливостей запуску стартап-проекту..... | 94 |
| 4.4 Розроблення ринкової стратегії проекту | 101 |
| 4.5 Розроблення маркетингової програми стартап-проекту | 104 |
| Висновок стартап-проекту | 108 |

| | |
|---|-----|
| ВИСНОВКИ | 109 |
| СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ | 111 |
| ДОДАТОК А Графічний матеріал | 118 |
| Скатерограми центроїди для кожної з груп | 119 |
| Скатерограми центроїди для кожної з груп(продовження) | 120 |
| Карта зон ризику захворювання у просторі перших двох компонент та метрика розбиття на кластери | 121 |
| Методика розрахунку | 122 |
| Роз'яснення карти зон ризику пацієнта | 123 |
| Схема перетворення вхідних даних до отримання графіків результатів | 124 |

ВСТУП

Саме найголовніше для кожної людини - це звичайно ж здоров'я. Як говориться у народній приказці: «Здоров'я за гроші не купити». Тому дуже потрібно вберігати його та застосовувати заходів з приводу профілактики виникнення різного виду захворювань.

Більшість із нас у своєму житті обов'язково зіштовхується з порушеннями власного здоров'я. Щорічна медична статистика надає досить тривожні дані про збільшення об'єму летальних випадків у пов'язаних з захворюваннями серця та серцево судинної системи. Приміром, у загальній статистиці летальних наслідків у країнах Європи та Північної Америки відсоткова кількість серцево-судинних захворювань являється 40-50 відсотків, в Україні - більш ніж половина, - 60 відсотків. Результати проведених досліджень говорять про велике поширення факторів ризику серцево-судинних хворіб в Україні.

Не минає й року в якому реєструється щонайменше 50 тис. випадків інфаркту міокарда, інсультів - 100-120 тис., миготливих аритмій - 20 тис., та приблизно 4 тис. набутих вад серцевої системи та народжується більш ніж 3 тис. малят з природженими дефектами серця.

Підхмарний рівень летальності від серцево-судинних захворювань обумовлений:

- неякісним виявленням хвороб на ранніх стадіях та етапах;
- медичні заклади та заклади охорони здоров'я на сьогоднішній день мають досить низький рівень оснащення обладнанням, яке необхідну для своєчасної та якісної діагностики, лікування, профілактики, надання інтенсивної допомоги, та запобіганню захворювань. А також незадовільною організацією процесу лікування та діагностування хвороб;
- Замале бюджетне та державне фінансування для закупівлі витратних матеріалів, обладнання, ліків;
- обмежена доступність новітніх методів лікування загальних серцево-судинних захворювань;

– Замалою проінформованістю людського населення про фактори ризику захворювань та можливість відвертання серцево-судинних хвороб.

Найбільш розповсюдженими є захворювання системи кровообігу. Вони потребують створення чіткого порядку надання лікарської допомоги з моменту першого її виявлення, та аж до здійснення реабілітаційних заходів, - це ішемія, або ішемічна хвороба серця, судинно-мозкові захворювання та природжені недуги серця.

Дуже великою за масштабами проблемою, з приводу питання людського здоров'я, є невчасне визначення порушень, тобто виявлення їх не на початкових періодах, при яких достатньо всього лиш звернутися до лікаря та використати набагато простіший вид лікування хвороби аби уникнути проблеми зі станом здоров'ям. Якщо задуматись над тим, що отримання професіональної медичної допомоги з кожним днем виростає в ціні, одразу стає зрозумілим, що необхідно терміново шукати новий підхід до вирішення проблеми.

На сьогоднішній день не лише в нашій країні серцево-судинні захворювання належать до найбільш складних і життєвоzagрозливих. Так, у 2004 році зареєстровано більше ніж 10 млн хворих на гіпертонію, 7,6 млн. недужих - на ішемію і близько 3 млн. осіб - з цереброваскулярними хворобами. Майже щороку лікувально-профілактичними закладами виявляється біля 2 млн. хворих з такими патологіями, кожний другий з яких на даний момент все ще має змогу працювати. Летальність від хвороби системи кровообігу в Україні посідає надвисоке положення і у близько 2-4 рази вища, ніж у наших сусідів та країн Європейського союзу та в світі загалом. Та насамперед в нашій країні люди помирають від цих захворювань не тільки частіше, але ще й раніше ніж у вище сказаних країнах.

У такому випадку, захворювання серцево-судинної системи є дуже великою прогалиною та соціальною проблемою, для розв'язання якої потрібна не тільки державна підтримка, а й координація усіх зусиль центральних та місцевих органів виконавчої влади.

З іншої сторони, у сьогоднішній період є направленість у розвитку обробки великих обсягів інформації та їх інтелектуального аналізу за

допомогою методів кластеризації. Такий вид аналізування даних дає змогу виявити та зрозуміти різноманіття процесів які проходять для подальшого використання отриманої інформації у різних сферах життєдіяльності, що супроводжують людину, наприклад у медичній сфері.

Так, для прикладу, при проведенні медичних досліджень (вимір артеріального тиску, термометрія, запис ЕКГ) зовсім не вирішено задачу, яка пов'язана з інтелектуальним аналізом часових рядів.

Одним із виходів може бути розробка, яка представлена у даній магістерській дисертації - «Інтелектуальний аналіз аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань», який представляє собою систему аналізу даних за допомогою поєднання методів кластерного аналізу та методу головних компонент.

Мета дослідження – є розробка онлайн системи інтелектуального аналізу ритмів електрокардіограм для виявлення аномалій у роботі серцево-судинної системи.

Об'єкт дослідження – процес інтелектуального аналізу даних та аномалій у ритмах ЕКГ.

Предмет дослідження – Інтелектуальний аналіз даних часових рядів за допомогою кластерного аналізу та нейронної мережі за навчанням без вчителя

Задачі дослідження:

- аналізу існуючих підходів та методів кластеризації ритмів ЕКГ;
- вибір оптимального алгоритму кластеризації даних у вигляді часових рядів;
- проектування власної системи виявлення аномалій у ритмах ЕКГ;
- проведення аналізу методів згортання даних та підходів;
- розробка власного програмного забезпечення аналізу ритмів ЕКГ та спеціалізованої бази даних до неї;
- проведення експериментів на основі тестових та реальних даних.

1 ПРОЕКТНІ РІШЕННЯ З РОЗРОБКИ СИСТЕМИ

1.1 Опис бізнес – процесів

1.1.1 Опис процесу діяльності

При вивченні медичних інформаційних систем мною були розглянуті консультативні діагностичні системи (КДС), котрі були розроблені та призначенні для діагностування ненормальних, або патологічних, становищ. При вивченні інформації про такі системи стало зрозуміло що в них прийнято розділяти такі основні частини: база даних (БД) і база знань (БЗ); механізм для логічного висновку (МЛВ); інтерфейс користувача. Також МЛВ прийнято розділяти за способом реалізації на 1) експертні КДС, 2) імовірнісні КДС. В ЕС прийнято реалізовувати логіку прийняття рішення як досвідченим лікарем. ЕС належать до класу систем «штучного інтелекту». Такі підходи як формальна і неформальна (нейромережева) логіка можуть бути використані для їх побудови. У таблиці 1.1 наведено специфіку для кожного з даних напрямків.

Таблиця 1.1 – Експертні системи та специфіка їх реалізації на підходах формальної і неформальної логіки

| | Експертні системи на основі напрямку формальної логіки | Експертні Системи на основі напрямку неформальної логіки Або нейромережевої логіки |
|---------------|---|---|
| Джерело знань | Формалізовані навички та досвід експерта в області, виражені у виді логічних стверджувачів, засад і фактів, які безумовно будуть прийняті системою. | Сукупний загальний досвід вчителя-експерта, що розрізняє та обирає приклади для подальшого навчання а також доповнює свій власний досвід нейронної мережі, яка навчається на таких прикладах. |

Продовження таблиці 1.1

| | Експертні системи на основі напрямку формальної логіки | Експертні Системи на основі напрямку неформальної логіки Або нейромережевої логіки |
|----------------------|---|---|
| Характер знань | Формально–логічне або як його ще називають - "лівопівкульне" знання яке відображено як правила. | Асоціативне, або "правопівкульне" знання. Має вигляд зв'язків між нейронами у нейронній мережі |
| Розвиток знань | За формою розширення сукупності правил а також фактів | За формою донавчання на додатковій послідовності прикладів. Має також уточнення кордонів категорій та формуванням нових типів та категорій |
| Роль експерта | Має змогу задавати на основі певних правил обширний обсяг знань ЕС | Має змогу відбирати характерні та типові приклади, при цьому не формулюючи для них спеціального обґрунтування свого вибору |
| Роль штучної системи | Проведення пошуку послідовності фактів та певних правил щоб довести судження | Формулювання індивідуального досвіду у вигляді певних категорій, які отримані та базовані на основі прикладів, та розділу на категорії, образів |

Такі комп'ютерні системи розділяють на декілька поколінь, за їх призначенням та можливостями. До, так званого першого покоління прийнято відносити системи, які можуть тільки повторити логічний поставлений висновок експерта. Але спеціалісту, котрий повинен вирішувати досить складне за інтелектуальним навантаженням завдання, відверто недостатньо

представлених можливостей такої системи, яка має змогу лише імітувати людську діяльність. Йому(експерту) нагально потрібно, щоб така система мала змогу виступати та виконувати роль повноцінного асистента, консультанта, порадника, який зможе провести повноцінний аналіз будь-яких даних, висувати і відкидати певні гіпотези, мати змогу проводити оцінку достовірності подій. Одним з найкращих переваг була б самостійна наповнюваність новими знаннями, їх контроль, оцінка, та вивід з цих даних висновків на основі прецедентів і, як варіант, породження рішень нових завдань, які до цього ніколи не розглядалися. Присутність таких можливостей у таких системах є характерним для другого покоління ЕС. Розробка таких систем розпочалась всього 9-10 років тому. Друге покоління експертних системи, називають партнерськими, або також інтелектуальними підсилювачами здібностей людини(експерта).

Отже, так як створення будь якої з таких ЕС дуже пов'язано з матеріальними витратами, а також часовими витратами, тому слід чітко обумовити ситуації, в яких створення подібних систем є доцільним та рентабельним. До них відносяться такі ситуації:

Потреба у символічних міркуваннях. Вочевидь, немає жодного сенсу починати розробку експертної систему для чисельних розрахунків, таких як перетворень Фур'є чи подібних. Для інтегрування, рішення систем алгебраїчних рівнянь та ін.. Зумовлене це тим що для цього фігурують спеціальні програмні засоби (для прикладу такі як MathCad)

Наявність експертів, які мають компетенцію в обраному колі запитань, котрі згодні проводити співпрацю при створенні ЕС.

Важливість та актуальність поставленої проблема мусить бути доволі насущною і вагомою. Це можуть бути такі проблеми, що вимагають високого ступеня експертизи, або навіть прості, але мають багату трудомісткість у перевірці, та також багаторазово повторюються. Витрата часу на вирішення таких проблем, які можуть виникати досить рідко мають можливість бути вирішені людиною зі звичною, не експертною, кваліфікацією, не мають жодного сенсу.

Необхідність у чіткому обмеженні кола виконуваних завдань, які вирішуються. Тобто предметна область обирається досить «вузько», аби уникнути «комбінаторного вибуху» в великому обсязі даних, яка дуже необхідна для вирішення поставленого завдання на компетентному рівні.

Необхідність у злагодженості та узгодженості думок експертів на тему того, у якому напрямку слід вирішувати завдання які поставлені, які факти є необхідність використовувати а також які взагалі є загальноживані правила для винесення певного судження. В іншому випадку немає можливості у розширенні бази знань за певні межі досвіду та навичок однієї людини та здійснити сплав, або поєднання експертних знань з двох, або більше областей.

Обов'язково має бути в досталь вихідних даних, які потрібні для перевірки працездатності та спроможностей такої експертної системи в тій чи іншій предметній області, аби розробники мали змогу переконатися в досяжності або недосяжності певного заданого рівня та поведінки.

Повинна постійно забезпечуватися спроможність поступового нарощування можливостей та вмінь системи.

Значущість при використанні ЕС проявляється в багатьох моментах.

У збиранні, ефективному уточненні, кодуванні і розповсюдженні експертних знань.

В дійовому рішенні питань, складність котрих переважає людські можливості, і для яких є необхідність в експертних знаннях з декількох різних областей.

У зберіганні найбільш уразливої цінності колективу – її пам'яті.

Експертні системи та їх структура

До складу типової експертної системи входять такі складові

Інтерфейс, який має забезпечувати можливість простого та зрозумілого спілкування користувача функціями експертної системи в інтуїтивно зрозумілій формі. Він дає можливість передавати до неї інформації, що становить повний контент бази даних, а також звертання до системи з будь-яким запитанням або за роз'ясненням.

Розділ роз'яснень є одним з найважливіших деталей експертної системи, бо це той розділ що надає змогу користувачеві переконатися в правильному обґрунтуванні такої інформації, яка одержана ним від цієї експертної системи, дає змогу поставити їй запитання, і на підставі певних мізковитих відповідей та довіряти їй.

Робоча пам'ять - частина база даних. Зберігає інформацію проблемної області та надає інформаційний зв'язок який задається певними заданими правилами в базі знань системи.

Диспетчер. Він означає певний порядок функціонування даної системи, який розплановує порядок постановки а також упорядкованість у досягненні цілей.

Машина логічного висновку або механізм виводу – формально-логічна система, яка реалізується у вигляді деякого програмного відсіку, та на основі визначених правил та методів у базі знань і робить перетворення конкретної інформації про об'єкти до певного виду, що відповідає правильному покликанню експертної системи, наприклад у діагнозі, плані дій і т. ін..

База знань – сукупність всіх існуючих відомостей та даних про проблемну частину(область), для якої саме й призначена така експертна система, що записані при допомозі певних узгоджених формальних конструкцій представлення знань, а саме набору деяких правил, або фреймів, семантичних мереж, чи тощо).

Машина логічного висновку (МЛВ) – програмний складник експертних систем, котрий реалізовує сам процес міркування системи базованій на базі знань та базі даних інформації. Вона виконує дві найважливіші функції: спершу, переглядання можливих фактів та знань з бази даних і потрібних в даний момент правил з бази знань а також додавання, по мірі її можливостей, в базу даних нових фактів. По-друге, визначення певної організованості перегляду та застосування цих самих правил. Дана підсистема верховодить процесом консультування, зберігає для користувачів інформацію з приводу отриманих висновків, та надає запит, коли для коректного спрацьовування

наступного правила в практичній множині виявляється обмаль інформаційних даних.

Загальна мета таких ЕС – це виведення певного заданого факту, котрий має називається висновком або цільовим ствердженням. Або в іншому випадку спростувати такий факт. Таке цільове ствердження може бути або «закладено» заздалегідь у базі системи, або може витягуватися під час діалогу системи з користувачем.

Така робота системи в загальному вигляді являє собою просту послідовність певних закладених до системи кроків, на кожному з яких з бази обирається декотре правило, яке потім використовується для поточного вмісту практичної множини. Вміст такої робочої множини, або робочої пам'яті, є наслідок тривалого діалогу з користувачем, який постійно поступово поповнюється інформацією з бази даних. Цикл буде закінчується тоді, коли виведено або спростовано головне цільове ствердження. Ланцюг роботи експертної системи по інакшому може називатися логічним висновком. Логічний висновок може здійснюватися дуже багатьма варіантами та способами, з котрих найбільш популярні це – прямий та зворотній порядок виводів.

Прямий порядок виводу – починається від фактів, які розташовується в робочий підмножині, та закінчується висновком. Якщо такий висновок все ж вдається віднайти, то такий вноситься до бази даних системи. Прямий висновок також в більшості називають висновком, котрий управляється даними.

Експертні системи та їх бази знань.

Якість ЕС прийнято задавати розміром, кількістю інформації, і якістю самої бази знань. Система працює в подальшому циклічному розпорядку:

- вибір, або запит даних та наслідків аналізів;
- інтерпретування результатів та коректне засвоєння отриманої нової інформації;
- висунення при допомозі бази знань системи непостійних припущень.

Подібний процес простягається до тих пір, доки не прийде інформація, котрої буде достатньо для остаточного та вирішального висновку.

Так у будь-який з моментів часу в системі побутують три підтипи знань:

- статичні знання – знання про потрібну предметну область, які після того моменту коли їх виявили, не змінюються;
- динамічні знання – перемінливі знання про потрібну предметну область, які постійно поновлюються по мірі вливу нової важливої інформації;
- робочі знання – тимчасовий, непостійний набір знань, котрий використовується для рішення певної конкретної задачі.

Перераховані до цього вище знання мають зберігатися у базі знань системи. Для побудови бази є необхідність проводити опитування між фахівцями, які є експертами в тій чи іншій конкретній предметній області, а потім впорядкувати, класифікувати, організувати та забезпечити ці знання показниками, аби трохи згодом була б можливість їх легко витягти з бази знань.

Джерелом знань для кожної конкретної ЕС можуть стати книжки та підручники, посібники, матеріали наукових досліджень в конкретній області. Також самі розроблювачі можуть мати теоретичні та практичні знання а також певний досвід у даній конкретній області. Але класичним джерелом знань є експерт, досвідчений професіонал в області знань. Для благополучного рішення такого питання потрібні спільні зусилля спеціалістів різних областей таких як математиків, програмістів, психологів, медиків.

На сьогоднішній момент немає підготовлених систем, що дозволяли б повністю виключити людину з циклу, що має причетність до формування бази знань системи, але все-таки ми маємо теоретичні наукові дослідження і дослідні розроблення, факт наявності яких може дозволити визначити три підкласи способів набування знань:

- традиційна та узвичаєна бесіда експерта з інженером про знаннях, в якій усі знання отримуються від експерта;
- автоматизоване створення знань, яке надає змогу частині правил отримуватися в автоматичному режимі;
- будівництво персонального макета дослідження наочної області визначеним експертом, що зволяє впорядити планомірний процес наукового

дослідження в цій області базованій на індивідуальних представленнях саме цього експерта-консультанта.

На початку розробці ЕС є необхідність розпочати роботу зі створення моделі на «папері». Така модель починає формуватися у процесі комунікації з експертом-консультантом. У такий спосіб виділяються визначальні поняття, котрими може оперувати експерт, формується тезаурус, тобто словник системи. Вже після таких дій на кількох нескладних та простих взірцях детально аналізується методика, якою експерт може вирішувати такого типу задачі.

Експертні системи та їх застосування у медичній сфері

Знання у даній сфері мають емпіричний характер. Вони досить слабо структуровані і недобре формалізовані, тому саме завдання такого характеру дуже доцільно вирішувати за допомогою таких ЕС, які при процесу вирішенні небанальних доручень проводять перебір і розбір немалої кількості допустимих варіантів і придатні до навчання застосовуючи нарощений ними досвід. Таку ідеологію ЕС можна виразити формулою: знання + висновки = система.

У 1990 році Вільям Бакст з Каліфорнійського університету в Сан-Дієго використовував нейронну мережу – багат шаровий перцептрон – для розпізнавання інфаркту міокарда у пацієнтів, що надходять в приймальний спокій з гострим болем у грудях. Метою було створення такого інструментарію, здатного допомогти лікарям та медикам, які не в змозі зорудувати з великим потоком даних, які означають стан обстеженого пацієнта з хворобою. Також ще однією ціллю може бути покращення якості діагностики. Доволі сумна статистика має такий результат: лікар в змозі правильно діагностувати інфаркт міокарда у лише 80% хворих і може випадково поставити такий діагноз у 25% випадків. Такий відсоток помилок є завеликим. Історія вживання багатоманітних методів оброблення великої кількості даних для покращення якості діагностування нараховує десятиліття, проте кращий з них зміг допомогти зменшити число інцидентів гіпердіагностики всього лише на 3%.

Своє завдання дослідник переобтяжив, оскільки аналізував дані тільки тих пацієнтів, яких вже зорієнтували до кардіологічного відділу. Бакст

використовував всього лише 20 параметрів, серед них були такі параметри вік, стать, локалізація болю, реакція на препарати, симптоми, такі як нудота і блювота, потіння, непритомність, частота дихання, пришвидшене серцебиття, попередні діагнози такі як інфаркти, гіпертонія, діабет, здуття шийної вени, ряд особливостей ЕКГ та наявність вагомих ішемічних змін.

Мережа змогла продемонструвати точність 92 відсотка при виявленні інфаркту і видала лише 4 відсотка інцидентів сигналів помилкової тривоги, невірно затверджуючи направлення пацієнтів без інфаркту в кардіологічний відділ. Також був присутній факт успішного використання штучних нейронних мереж у діагностиці хвороби. Зараз є необхідність роз'яснити, в яких параметрах проводиться оцінка якості діагнозу в суцільному випадку. Припустимо, що з десяти осіб, у яких дійсно є інфаркт міокарда, діагностичний метод дає змогу виявити хворобу у восьми з них. В такому випадку точність методу складе 80 відсотків. Якщо ж ми оберемо десять персон, у котрих немає захворювання інфаркту міокарда, а методика діагностики запідозрить його у трьох осіб з десяти, то частина неправильних тривог набуде 30 відсотків, при цьому набавна до нього характеристика – специфічність методу – сягне 70 відсотків.

В ідеальному випадку такий метод діагностування зобов'язаний мати стовідсоткові чутливість і специфічність – спершу, не пропускати жодної реально хворої людини і, по-друге, надавати хибних тверджень аби не лякати здорових людей. Щоб застрахуватися, потрібно намагатися передусім забезпечити стовідсоткову чутливість методу – не можна пропускати захворювання. Але це, як правило, обертається тим що, падає специфічність методу – у багатьох людей лікарі допускають захворювання, якими по-справжньому пацієнти не страждають.

Нейронні мережі показують собою нелінійні системи, що дозволяють набагато істотніше та якісніше класифікувати дані, ніж звичайно використовувані лінійні методи. У додатку в медичній діагностиці вони надають змогу значно збільшити специфічність методу, не знижуючи відсоток його чутливості.

Нагадаймо, що нейронна мережа, яка діагностує інфаркт міокарда, працювала з досить великим пакетом параметрів, вплив котрих на постанову діагнозу людині не постає можливість оцінити. Але тим не менш, нейромережі проявили себе спроможними приймати рішення, ґрунтуючись на виявлених ними схованих закономірностях в багатовимірних даних. Примітна властивість нейромереж полягає в тому, що вони не запрограмовуються – не застосовують якихось правил виводу для постанови діагнозу, а навчаються робити це на прикладах. Саме в цьому сенсі нейромережі взагалі не збіжні на експертні системи, розробка котрих у 70-ті роки здійснювалася після псевдо "перемоги штучного інтелекту" над тим методом моделювання пам'яті, розпізнавання образів і приведення до загального вигляду, який базувався на вивченні нейронної організації мозку.

У медицині віднаходить вживання і інша дуже важлива відзнака нейромереж – їх здібність передбачати непостійні послідовності. До цього вже зазначалося, що експертні системи досягнули успіху в аналізі ЕКГ. Нейромережі тут теж несуть велику користь. КіЧженху, Ю Хенуві і Вілліс Томпкінс з університету штату Вісконсін розробили та підготували нейромережеву систему фільтрів для електрокардіограм, які дозволяють пригнітати нелінійний і нестационарний шум набагато краще та істотніше, ніж методи, які використовувалися до цього. Справа в тому, що нейромережа добре завбачала шум за його вагомністю у попередні моменти часу та періоди. А те, що нейромережі мають дуже велику ефективність для передбачення несталих послідовностей, та аргументовано показали результати і наслідки змагання передбачуваних програм, які проводив університет у Санта Фе – нейромережі змогли зайняти перше місце і домінувати серед найкращих методів.

Також є можливість використовувати нейромережі і для прогнозу дії різних родів розроблених способів лікування. Такі мережі вже успішно використовуються в хімії для прогнозу особливостей сполук на базі їх молекулярної будови. Дослідники-науковці з Національного інституту раку в США почали використовувати нейромережі для пророкування механізму дії препаратів, які застосовуються при хіміотерапії злоякісних пухлин. Потрібно

також зауважити, що фігурують мільйони різних відмінних одна від одної молекул, які також потрібно опрацьовувати на предмет їх анти ракової активності. Експерти Інституту раку розділили відомі онкологічні препарати на шість підгруп відповідно до механізму їх дії на клітини та змогли навчити багат шарові мережі систематизувати нові речовини і розрізнати їх дію. У якості вихідних даних використовували результати проведених експериментів з придушення зросту клітин з розбіжних пухлин. Нейромережева класифікація дозволила визначити, які з багатьох сотень молекул, що апробуються кожного дня, доцільно вивчати далі в дуже коштовних експериментах.

Перераховані вище технології ще далеко не вичерпують увесь діапазон застосування таких ЕС. Діагностування болю в зубній долі, прогнозування генних ускладнень, хвороби слухового апарату, раптова смертність новонароджених – ось далеко не повний спектр вдалого та результативного застосування ЕС у сучасній медицині.

1.1.2 Актори і функції

Розріст інформаційних комп'ютерних технологій нерозривно пов'язаний із розростом інформаційних систем, які в економіці знайшли своє використання для автоматизації машино-людинного розв'язання економічних завдань. Для розв'язування будь-якої з задач за допомогою комп'ютерних технологій потрібно виробити інформаційне забезпечення, тобто забезпечити розрахунки необхідними даними і математичним забезпеченням.

Полегшену схему автоматизованого розв'язання таких задач зображено на рисунку 1.1. Необхідна для розв'язування задачі інформаційна складова має можливість надходити безпосередньо або через інформаційну систему забезпечення, яка також має змогу поповнюватися новою інформацією. Істотною особливістю ІС є те, що вона забезпечує даними користувачів з кількох різних областей та організацій.

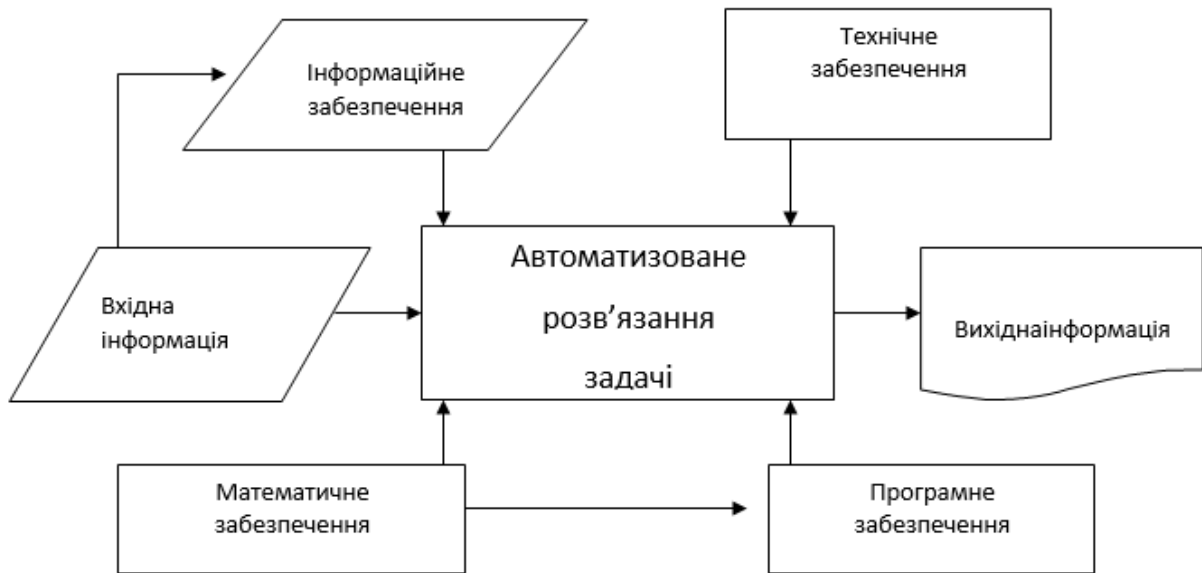


Рисунок 1.1 – Автоматизоване розв'язання задач. Візуальна схема

Алгоритми та математичні моделі можуть надаватися у вигляді, що завбачений для конкретного етапу розробки та програмування, і у конфігурації, придатній для практичного використання для розв'язання завдання. Вихідна інформація може бути подана в досить різних варіантах.

На протязі розросту організаційних ІС структурованість і надмірність інформативних даних та процеси обчислень часто змінювалися, що стало певними «мірлами» виділення поколінь таких систем.

У першому поколінні інформаційних систем, котрі в зарубіжній літературі більш відомі під науковою назвою «DataProcessingSystem» — DPS («системи оброблення даних»), а у вітчизняній — «автоматизовані системи управління (АСУ) — позадачний підхід», для кожної конкретної задачі окремо збиралися дані і народжувалася математична модель. Даний підхід обумовлював математичну і інформаційну крайність. Найтипівішими прикладами таких АСУ є системи керування резервами, виписування рахунків, розрахунків заробітних плат для людей.

В подальшому ріст ІС вже пов'язується із тенденцією баз даних. На такій основі почали з'являтися інформаційні системи які було прийнято називати системами другого покоління. Основним функціоналом даних систем є забезпечення керівництва інформативними даними. Достеменно управлінську інформаційну систему охарактеризовує структурований плин інформації та

даних, вбудовування задач обробки даних, створення або генерування запитів і структурування звітів. При застосуванні MIS вже були визнані очевидні переваги групового застосовування даних. Економічна ефективність АСУ була досить значною: повністю змінився характер поведінки управлінського колективу. Збільшилась оперативність, якість, наукова аргументованість та об'єктивність управлінських рішень; зросла інформованість персоналу управлінського сектору.

Системи підтримки прийняття рішень — СППР (Decision Support Systems — DSS) — це інформаційні системи які окреслили системами вже третього покоління. СППР — інтерактивні комп'ютерні системи, які були призначені для підтримки всіляких видів діяльності в разі прийняття рішень стосовно погано структурованих або зовсім неструктурованих питань. Інтерес до СППР зростає на постійній основі. У більшості досить розвинених країн світу розробка та запровадження СППР змогло перетворитися на сферу бізнесу, що почав швидко розвиватися.

СППР можуть мати суцільно не тільки інформаційне, але й математичне забезпечення — бази моделей, тобто запроваджено ідею поділу обчислень аналогічно до того, як розподілення даних стало ваговитим фактором у ординарних інформаційних системах.

Інформаційна комп'ютерна система СППР використовується для сприяння різних видів діяльності у процесах прийняття рішень: вибору гуртових стратегій дій, дефініції спеціальних завдань, делегування відповідальності, оцінювання наслідків, ініціювання відмін. Запитання підтримки рішень на всіх стадіях такого виду процесів стають все далі актуальнішими.

Арена впливу організаційних адміністративних рішень досить велика. Рішення народжуються індивідуумами повністю на всіх рівнях в організації і грандіозною кількістю гуртів в організаціях. Вельми відома класична класифікація управлінських рішень на чотири підвиди, які асоціюються та поєднуються з організаційними ступенями.

Стратегічне планування (Strategic Planning) — процеси прийому рішень, прив'язані з розподілом ресурсів, контролюванням за ефективністю організації,

дефініцією генеральної політики, оцінкою інвестицій та/або пропозицій щодо злиття корпорацій.

Адміністративне управління (Management Control) — це рішення, які торкаються здобутку і вжитку ресурсів при допомозі адміністративного персоналу; поводження клієнтів і постачальників; започаткування виготовлення нових продуктів; видатків на проектування, дослідження та розроблення.

Оперативний контроль (Operational Control) — це рішення щодо ефективності організаційних дій; слідкуванням за якістю продукції та обслуговуванням; знадоб в оцінці продукції чи обслуговуванні.

Операційне виконання (Operational Performance) — буденні рішення, які приймаються керівниками з ціллю здійснення стратегічних і тактичних постанов та протічних операцій.

Організаційне управління. Класифікація проблем

Вони поділяються на декілька типів. До першого припадають добре структуровані (абсолютно формалізовані, кількісно сформульовані) питання, в котрих істотні залежності визначені настільки всебічно, що вони зможуть бути виражені числами або навіть символами, і тому так легко піддаються стандартизуванню та програмуванню. До таких подібних завдань припадають: обрахунок і контроль; оформлення документообігу, їх розповсюдження тощо. У звичних інформаційних системах (АСУ) такого виду задачі здебільшого автоматизовані, зазвичай, повністю.

Другий тип — це неструктуровані проблеми (завдання), для котрих були описані лише найважливіші ресурси, атрибути і рекомендації, а чисельні залежності між ними залишаються невідомі. Розв'язок таких задач можливий лише у разі використання неформалізованих процесів та процедур, які мають основу на неструктурованій, з панівним ступенем невизначеності інформативних даних. До такого типу задач належить вагома частина проблем що стосуються прогнозування, далекосяжного планування, організаційної трансформації. Більшість неструктурованих питань знаходять розв'язок при допомозі евристичних методик, у котрих не передбачена ніяка упорядкованість

логічних процедур пошуку їх розв'язання, а отже сама методика цілком і повністю залежить від персональних характеристик людини.

До третього типу проблем належать досить слабо структуровані, змішані, або ще можуть бути напів структуровані проблеми. Вони мають як кількісні, так і якісні частини. Причому невизначні та маловідомо акцентовані проблеми користуються тенденцією домінування. Для такого виду завдань своєрідна неприсутність методик розв'язку на базі безпосереднього перетворення даних. Постанова таких завдань потребує від себе прийняття вирішень за умови великої недостатності в інформаційному полі. Були також відомі інциденти, коли на базі теорії нечітких множин а також її застосувань були впроваджені формальні схеми розв'язку подібних задач. До слабо структурованих задач ми можемо віднести такі задачі як розподіл капіталовкладень, вибір проекту, проведення наукових опрацювань і розроблень, тощо.

До найтипівіших погано структурованих задач характерні такі риси:

- рішення, які збираються бути прийнятими, стосуються лише майбутнього;
- великий обсяг альтернатив;
- від неповноти знань залежать рішення з приводу нинішніх досягнень;
- потреба у великих витратах внутрішніх ресурсів для реалізації запропонованих рішень, бо пов'язані з елементами ризику;
- вимоги, які стосуються вартості й тривалості розв'язку визначені неповністю;
- необхідність комбінування різних ресурсів для розв'язання що ускладнює проблему.

Опишемо функціональне поводження системи та її взаємозв'язок з користувачем при сприянні функціональної моделі, яка описана за допомогою варіантів використання.

Система додає доступ до даних одному актору – **користувачу/лікарю**.
Означимо перелік дій, що може виконувати цей актор:

- внесення нових даних;

- перегляд актуальних даних;
- анотування запису та визначення рівня ризику захворювання
- зміна даних у кожному окремому записі;
- завантаження звіту про запис.

Надаймо короткий опис кожному з варіантів використання:

Внесення нових даних – завантаження нових даних до бази знань або внесення нової інформації про запис.

Перегляд актуальних даних – користувач переглядає дані за певним обраним записом ЕКГ. Для нього відображається уся інформація, що зараз зберігається у базі даних про цей запис.

Анотування запису та визначення рівня ризику захворювання – система надає функціонал для анотування, тобто короткого опису кожного засобу та визначення рівня ризику того чи іншого серцево-судинного захворювання.

Зміна даних про кожний запис – внесення коректив щодо запису, його анотації чи рівня ризику захворювання.

Завантаження звіту про запит – можливість вивантаження з бази знань звіту який має в собі всі дані про запис наявні у базі на цей момент.

1.1.3 Структура бізнес-процесів

Поява персональних комп'ютерів котрі з'явилися в кінці 80-х років минулого століття стало досить бурхливим поштовхом для інформатизації та автоматизації дуже багатьох галузей техніки і науки, в тому числі, і медичної сфери діяльності. Ми не можемо сказати, що до появи настільних обчислювальних систем в медичній сфері обчислювальна техніка взагалі ніколи не використовувалася. Для прикладу, один з найбільш потужніших методик діагностування – рентгенівська томографія – принципово базована на використанні досить складних суцільно математичних методах обробки приладової інформації і навіть встигла отримати іншу, більш сучасну назву – комп'ютерна томографія. Проте, в нашому випадку, лікаря цікавить тільки сам

результат – комплект діагностичних зображень, отриманих на екрані приладу. Безпосередня праця з комп'ютерною системою покладається на інженерний колектив томографа. Компактність, відносна доступність, легкість і простота використання персонального комп'ютера значно приблизили сучасну обчислювальну техніку до всіх, хто має хоч невелике відношення до роботи з великим об'ємом інформації, в тому числі, і до непрофесіоналів в області вживання математичних методик оброблення даних.

Медицина, а особливо діагностування, тісно пов'язана з обробкою і накопиченням інформаційних даних. Відповідно, від якості методик роботи з інформацією, залежить надійність діагностування і в подальшому, ефективність лікування. У задумі роботи з інформацією діагностика опирається на отримання показників про життєвий стан пацієнта та їх інтерпретацію. Що торкається інтерпретації даних які ми отримуємо, то це запитання не настільки просте. Зазвичай лікар, при аналізі отримуваних даних робить висновки з наявності виразних ознак захворювань, які виявляються тим чи іншим обстеженням. Найбільш не складним та очевидним є випадок, коли хворобі відповідає однозначний комплекс ознак. Але людський організм це дуже складна і заплутана система, отже такий випадок зустрічається дуже рідко. Частіше всього виявляються ознаки, які з певною вірогідністю відповідатимуть декільком яким-небудь недугам. В такому випадку добре, коли є можливість виконати додаткове дослідження для отримання більш точної інформації та збільшення кількості ознак хвороби. Однак така можливість не завжди є. Саме у такому зв'язку з великою неповнотою даних і доволі частою відсутністю виразного порога в інтерпретації параметра частіше за все виникають проблеми у постанові надійного та якісного діагнозу, а на перше місце, чим керується лікар виходить досвід отриманий від попереднього досвіду.

Опираючись на таку ситуацію є потреба у накопиченні і статистичному аналізі інформаційних даних для підвищення якості діагностики, отриманих при спостереженні багатьох пацієнтів з різними видами захворювань, глибокому обстеженні та слідкуванню за станом пацієнта протягом тривалого періоду, використання комп'ютерних експертних систем – програм, які

надаватимуть консультативну допомогу у постановці діагнозу на основі порівняння поточних даних з накопиченими до цього.

Для відстежування стану пацієнта, та запису даних про його хвороби протягом життя використовується медична карта пацієнта – паперовий документ з великою кількістю сторінок, що заповнюється, зазвичай, рукописно. Протягом багатьох десятиліть такий документ був єдиною можливістю для накопичення даних та систематизованої інформації про кожного пацієнта.

На сьогоднішній день у такому вигляді така медична карта взагалі не відповідає сучасним вимогам у 21-столітті, а особливо в умовах страхової та комерційної медицини, коли найменша неточність діагнозу, нечіткість або некоректність запису, може спричинити складний та довгий судовий розгляд з позовами на величезні суми. Поки що в нашій в країні подібне – екзотика, але як показує світова практика, якість ведення документообігу надає якісь гарантії юридичного статусу лікаря або самого клінічного закладу в цілому. Сучасний положення медичної карти, як правознавчого документу, дуже далекий від ідеальності. Що торкається безпосередньо медичного вжитку, досить великим питанням, є аналіз часових залежностей зумовлене слабкою структурованістю карти і неохайністю самих записів. У великому ряді випадків зберігається лише кінцевий варіант аналізів, а значна кількість приладових вивчення має характер виписок, тобто приміток лікарів-діагностів, які виконали інтерпретацію діагностичних даних, а самі дані, якщо і зберігаються, то лише архіві діагностичного відділу, а отримати до них оперативний доступ, якщо й складається можливим то цей процес є досить складним.

Відповіддю до рішення вищесказаних проблематик є спеціалізовані медичні інформаційні системи.

Як вже зазначалося до цього, інформаційні процеси наявні у всіх областях медичної сфери й галузі охорони здоров'я. Однією з найважливіших складових інформаційних процесів є інформаційні потоки. Ці потоки укладаються з окремих інформаційних звісток, які втілені в сигналах і документах, та рухаються в обширі простору і часі від джерела інформації до одержувача.

Для праці з такими інформаційними потоками спеціально призначені інформаційні системи (ІС).

Інформаційна система – організаційно впорядкована сукупність документів та технологічних інформаційних систем, у тому числі з вживанням способів обчислювальної техніки та зв'язку, які запроваджують інформаційні процеси.

На даний момент існують багато різних підходів до класифікації вищеписаних медичних інформаційних систем. Рисунку 1.2 зображена класифікація МІС в залежності від рівнів керування організацією.

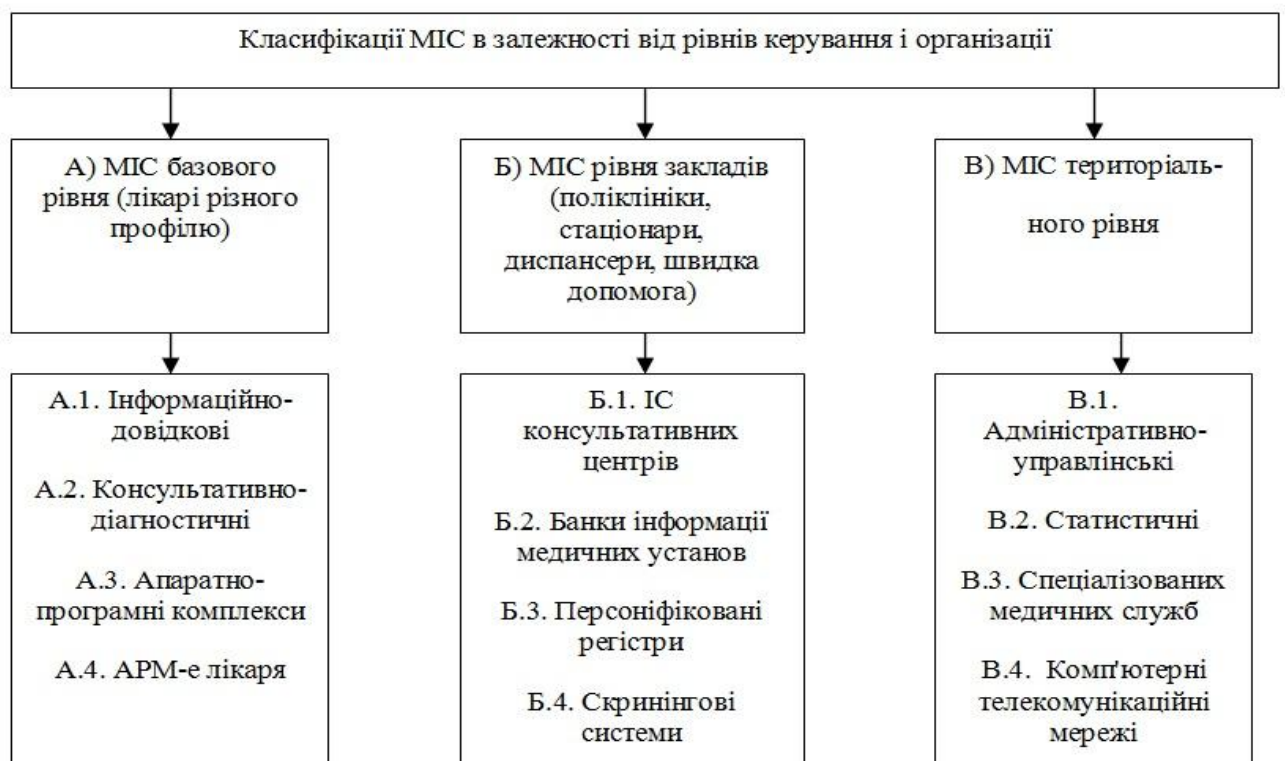


Рисунок 1.2 – Медична інформаційна система та їх класифікації

Базовий рівень МІС

МІС базового рівня має мету у комп'ютерній підтримці у роботі лікаря-діагноста, гігієніста, лаборанта та ін. До них відносяться:

Медичні інформаційно-довідкові системи призначення яких у пошуку та видачі інформації медичного характеру за запитом користувача, тобто лікаря. Потреба у накопиченні великих обсягів професійно цінної для лікаря інформації і операції над нею – одна з проблем, з якою приходиться натикатися

медичним працівникам різного роду. Такі системи зможуть полегшити вирішення цієї проблеми.

Масиви даних в таких системах (банки і бази даних) розміщують медичну довідкову-інформаційну характеристику різного типу. Системи даного підкласу не виконують обробку інформації, але можуть забезпечувати швидкий доступ до необхідних відомостей. Це може бути як наукова інформація з різних медичних сфер, так і будь-яка довідкова статистична чи обліково-документальна інформація.

В даний момент часу є дуже велика кількість різноманітних комерційних систем. Найпростішим прикладом таких систем є довідники лікарських препаратів, класифікатори, електронні атласи, тощо.

Медичні консультативно-діагностичні системи. Історично такого виду системи були одними з перших інформаційних систем які почали застосовувати. Актуальність у практичному використанні систем в медицині визначилася такими обставинами. В останні роки в усьому світі відзначається зростаюча розбіжність між соціальною потребою в гарному здоров'ї і виразно проявленому процесі деформації біологічної природи людського населення по мірі заглиблення науково-технічного прогресу, що супроводжується височезними швидкостями соціальних, економічних, кліматичних, техногенних, та інших змін, що несуть за собою катастрофічний характер. Всі ці події призводять до зниження функціональності та можливостей органів та систем організму в цілому, ослаблення спадковості, зміни характеру і зростання різноманітності проявів патологій у здоров'ї людини. Іноді виникає така ситуація, коли кожен «вузький» спеціаліст одному й тому ж пацієнту аргументовано ставить свій, відрізний від іншого фахівця діагноз. Звідси стає зрозумілим, що найефективнішим є шлях ґрунтовного покращання якості діагностики а також подальшого лікування постає в рівночасному, системному аналізі усіх видів симптомів захворювання, що і реалізує КДС.

Крім цього, в останній час у медичній сфері все більш значливо формується впевненість про необхідність врахування існування проміжного стану між станами здоров'я людини і захворювання, в той час коли адаптаційні

можливості людського організму вже знижені, але маніфестних виявів хвороби ще немає. У такій ситуації використання КДС при загальнодоступних обстеженнях практично здорового населення змогло б принести незмірну користь за рахунок завчасної діагностики патологічних процесів.

Застосування такої системи на практиці зможе надати можливість:

- запропонувати лікарю-діагносту кваліфікованого технічного-продвинутого консультанта, який швидко, якісно і на високому професійному рівні обробить усю інформацію, що може надходити про об'єкт і зможе видати поділ ймовірностей допустимих станів;

- мати змогу оперативно приймати рішення з приводу невідкладних заходів у критичні моменти;

- скорочення загального числа процедур та зменшення часу попереднього аналізу;

- надати змогу у навчанні та забезпечити тренуванням по цій методиці персонал, для прийняття відповідних рішень.

ЕС досить «дружно» налаштовані до користувача, що пояснюється їх можливостями відтворювати логічні висновки і закінчення, аргументуючи при цьому користувачеві, з якої підстави видаються такі чи інші висновки чи рекомендації, що різко підвищує довіру користувача-лікаря до свого «електронного порадики»; закликати користувача виконувати ті чи інші конкретні дії, яких потребує система для отримання необхідної відповіді; швидко і точно надавати відповідь на запитання користувача, причому, по мірі акумуляції нових знань, досвіду і навичок, швидкість відповіді прискорюється і стає більш якісною, точною і повноціннішою; також представляти користувачеві сумарні результати в найбільш практичній і доступній формі.

Висока «кваліфікація» ЕС і забезпечувана ними всеосяжна доступність будь-котрого користувача до банку знань, накопичених професіоналами найвищого рівня, дозволяють, понизити потреби до ступеня професійної грамотності користувача, виключити чи взагалі мінімізувати допустимі

неточності і некоректні висновки, та підвищити продуктивність праці і якість прийняття рішень.

Зокрема в останні роки все більший і більший інтерес викликає вивчення і використання для рішення важких комбінаторних задач систем штучного інтелекту непримиримо нового типу – штучних нейронних мереж.

Медичні апаратно–програмні комплекси МАПК використовуються для інформаційної допомоги та/або автоматизації діагностичного та лікувального процесу, які здійснюються при безпосередньому контакті з насамперед організмом недужого або об'єктом медичного дослідження.

В сьогоденню пору одним з напрямків інформатизації медицини є комп'ютеризація медичної та лікувальної апаратури. Використання в парі комп'ютера та вимірювального і керуючого пристрою в медичній практиці надало змогу впровадити нові ефективні способи для забезпечення автоматизації збору інформації про стан недужого, та її обробки в реальному масштабі часу а також управління станом його здоров'я.

Найтипівішими представниками МАПК є медичні системи моніторингу(слідкування) за станом пацієнтів. Наприклад, при проведенні важких операцій або в умовах високої ризикованості розвитку важких ускладнень; системи комп'ютерного аналізу медичних даних, томографії, ультразвукової діагностики, ЕКГ, ЕЕГ, радіографії; системи автоматизованого аналізу даних мікробіологічних та вірусологічних дослідів, аналізу людських клітин і тканин.

Типові МАПК можуть бути розділені на ряд класів розподілених за їх призначенням:

- системи для виконання функціональних і морфологічних дослідів;
- системи моніторингу;
- системи для управління процесом лікувальним пацієнта;
- системи діагностики в лабораторних відділеннях;
- системи для біологічно-медичних досліджень.

До систем для проведення функціональних і морфологічних досліджень прийнято відносити ті системи котрі займаються дослідженнями системи

кровообігу, органів дихання, відділу головного мозку, нервової системи, органів зору та слуху, рентгенологічні дослідження, УЗД – діагностування та ін.

Моніторні системи дають змогу відслідковувати стан здоров'я хворого в кімнатах інтенсивної терапії та операційних палатах. До систем керування такими процесами як лікування і реабілітація відносять автоматизовані системи інтенсивної терапії, біологічного зворотного зв'язку, протези і штучні органи.

До систем керівництва процесами лікування співвідносяться також системи біологічного зворотного взаємозв'язку – системи, призначені для подання пацієнту життєвої інформації про функціонування його внутрішніх і систем та органів, що дозволяє шляхом демонстративного вольового впливу пацієнта добиватися терапевтичного впливу при певному типі патології. Загальна структуризація медичної приладно-комп'ютерної системи має вигляд, який показано на рисунку 1.3.

В такій апаратурі для знімання медично-біологічної інформації здійснюється трансформація фізичних характеристик становища хворого у форму звичайних електричних сигналів. Одним із шаблонних пристроїв перетворення неперервного електричного сигналу в серію поодиноких цифрових сигналів для уведення інформації в комп'ютерну систему або мікропроцесорний пристрій відводиться аналогово-цифровому перетворювачу (АЦП). Під цифровою формою в даному випадку розуміється представлення сигналу у двійковій системі числення, де присутність електричного сигналу це відповідність цифрі 1, а відсутність – 0. Дані, які були отримані через інтерфейс такого АЦП, проходять обробку комп'ютером а вже потім пересилаються у формі двійкового шифру в інтерфейс цифрово-аналогового перетворювача (ЦАП).

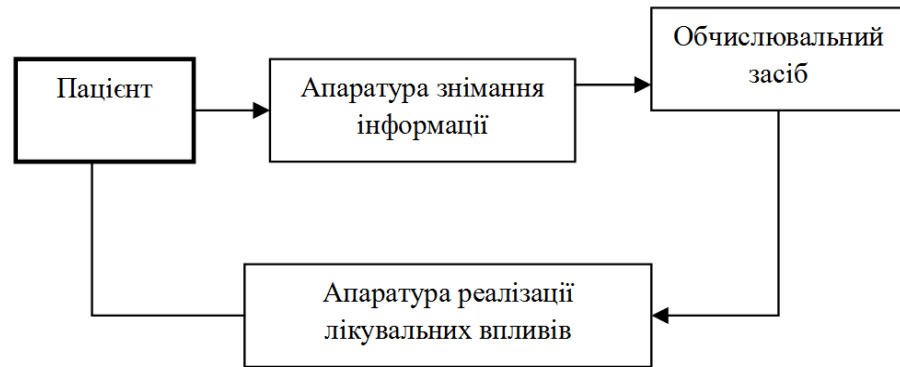


Рисунок 1.3 – Структура медичної приладно-комп'ютерної системи

Програмне забезпечення МАПК не менш потрібно та важливо, ніж апаратне, або технічне. Найбільш досконалі апарати оснащені, «інтегрованим» ПЗ, дякуючи якому лікар має змогу отримати повноцінну схему, яка охоплює увесь процес дослідження, який включає в себе етап підготовки, проведення дослідження та обробки інформаційних даних. У такому виді програмного забезпечення зазвичай виділяють шість головних функціональних підрозділів:

- підготовка дослідження;
- проведення дослідження;
- перегляд отриманих записів;
- обчислювальний аналіз;
- підсумовування та формування висновку;
- архівна робота.

1.2 Опис постановки задачі

В існуючих на даний момент умовах динамічної обробки інформації та у зв'язку з потребою аналізу часових рядів в інтелектуальному аналізі потоків даних, необхідно розібрати стан проблеми кластеризації часових рядів та наявні підходи до вирішення такого роду задач. Пропонується аналіз існуючих публікацій з використанням стандартних метрик кластеризації часових рядів, їх обробки за умов недостатньої кількості даних у ряді, аномалій, а також при неможливості конструювання моделей або при використанні методик на базі

статистики. Ставляться науково-практичні задачі дослідження в межах інтелектуального аналізу даних.

Потрібно проаналізувати стан питання нечіткої кластеризації часових рядів та існуючі методи вирішення. Дослідити різноманітні метрики, які використовуються у методах нечіткої кластеризації часових рядів, на основі підручних методів та метрик.

Мета – є розробка онлайн системи інтелектуального аналізу ритмів електрокардіограм для виявлення аномалій у роботі серцево-судинної системи та покращення якості постановки діагнозів пацієнтам з серцево-судинними захворюваннями за рахунок впровадження системи у клініки.

Завдання: а) аналіз існуючих підходів та методів кластеризації ритмів ЕКГ; б) обрання оптимального алгоритму кластеризації даних у вигляді часових рядів. в) проектування власної системи виявлення аномалій у ритмах ЕКГ. г) проведення аналізу методів та підходів згортання даних. д) розробка власного програмного забезпечення аналізу ритмів ЕКГ та спеціалізованої бази даних до неї. е) проведення експериментів на основі тестових та реальних даних.

1.3 Рішення з інформаційного забезпечення

Реалізація додатку відбувається з використанням Python та SQLite3.

SQLite - це автономний, що працює без сервера транзакційний механізм бази даних SQL. Python отримав модуль sqlite3 в версії 2.5, що означає що можна створювати базу даних SQLite в будь-якої справжньої версії Python, без необхідності завантаження додаткових інструментів.

Легко вбудовується в додатки база даних. Так як SQLite базується на файлах, то вона надає досить широкий набір інструментів для роботи з нею, в порівнянні з мережевими СУБД. При роботі з цією СУБД звернення відбуваються безпосередньо до файлів (в ці файлах зберігаються дані), замість портів і гнізд в мережеских СУБД. Саме тому SQLite дуже потужна завдяки технологіям обслуговуючих бібліотек.

Хоча SQLite - це компактна база даних, вона призначена для серйозних програм! Компанія Airbus, виробник літака, навіть використовує її для деяких польотних програм для своєї лінійки літаків A350. Деякі переваги SQLite3 включають:

Невеликий і самодостатній. Для його запуску не потрібні додаткові програми або компоненти. Двигун бази даних може працювати на будь-якому сучасному ПК або смартфоні. SQLite не вимагає запуску окремого комп'ютерного процесу. IT не покладається на послуги Windows, фонові демон та окреме комп'ютерне обладнання.

Портативний. Ділитися базами даних дуже просто. Просто скопіюйте один файл для цього. Саме це полегшує мені надання вам основної вибіркової бази даних SQL. Файл бази даних SQLite знаходиться вдома в Windows стільки, скільки в OS X.

Стандарти обізнані. SQLite базується на стандартній мові SQL. Деякі функції опущені, але ті, що реалізуються, тісно відповідають стандартам. Це означає, що те, що ви дізнаєтесь, буде легко перекладати в інші двигуни бази даних.

Мінуси SQLite. У неї є недоліки, оскільки вона настільки компактна і покликана працювати на невеликому просторі, були зроблені деякі компроміси.

Обмежені типи даних. SQLite не підтримує типи даних дати або часових позначок. Це недолік, оскільки багато головоломок SQL, таким чином, можливості для навчання, на початку недоступні.

Підтримка мови SQL. Є кілька функцій, які не підтримуються в SQLite. Праві зовнішні приєднання не підтримуються, а також не є повними зовнішніми приєднаннями. Оскільки ліві зовнішні з'єднання підтримуються і схожі на Right Outer Joins, можливість вчитися не втрачається. Також нелегко змінити існуючу таблицю, а також немає перевірок на цілісність. Це досить важливі функції для управління виробничою базою даних.

Обмеження розміру. SQLite не призначений для підтримки надзвичайно великих баз даних. Він не в змозі масштабувати сотні користувачів і не зберігати гігабайти даних. Він не підходить там, де ви очікуєте, що велика

кількість користувачів одночасно змінюватиме дані (висока одночасність), а також там, де великий обсяг транзакцій.

Висновки до розділу

У цьому розділі були розглянуті види сучасних медичних інформаційних систем. Також були розглянуті консультативні діагностичні системи (КДС), котрі були розроблені та призначенні для діагностування ненормальних, або патологічних, становищ. При вивченні інформації про такі системи стало зрозуміло що в них прийнято розділяти такі основні частини: база даних (БД) і база знань (БЗ); механізм для логічного висновку (МЛВ); інтерфейс користувача. Також МЛВ прийнято розділяти за способом реалізації на 1) експертні КДС, 2) імовірнісні КДС. Такі комп'ютерні системи розділяють на декілька поколінь, за їх призначенням та можливостями. Також були розглянуті та класифіковані медичні інформаційні системи (МІС) різного виду.

Була описана постановка задачі, визначена мета та завдання роботи. Були прийняті рішення з інформаційного забезпечення щодо системи, тобто була обрана мова програмування – Python, та визначені його переваги та недоліки.

2 МОДЕЛІ ТА МЕТОДИ РОЗВ'ЯЗАННЯ ЗАДАЧІ

2.1 Змістовна постановка задачі

Задача інтелектуального аналізу аномалій ритмів електрокардіограм формулюється таким чином: після проведеного виміру електрокардіограми з неї виділяються дані варіабельності серцевого ритму, а саме RR-інтервали ударів серця. На основі цих даних будуються теплові карти або так звані скатерограми які далі піддаються аналізу методом кластеризації. Проаналізовані дані формуються в кластерну карту зон ризику захворювання у просторі перших двох головних компонент. Основним завданням виявлення аномалій у роботі серця пацієнта та їх аналіз для диференціації стану здоров'я пацієнта.

Диференціація планувалася виконуватися за допомогою нейронної мережі за рахунок кластеризації (навчання без вчителя). Процес дослідження має такі етапи:

- з набору записів електричної активності серця виділяються дані по серцевим ритмам, як показником, найбільш відображаючим загальний стан здоров'я;

- будуються набори скатерограмм в логнормальній шкалі для ефективної візуалізації режимів роботи серця;

- методика головних компонент виділяються головні компоненти для подальшої кластеризації;

- групування даних у кластери методикою k-means.

На сьогоднішній день тенденція в обробці великих обсягів інформації та їх аналізу за допомогою кластеризації надає спромогу збагнути різноманіття процесів для її наступного вживання у різних сферах життєдіяльності, що супроводжують людину. Розглянуті в проаналізованій літературі задачі не поєднують в собі одночасно вирішення всіх задач.

На основі отриманих даних лікар матиме змогу зробити передбачення щодо ризику захворювання. Застосування системи при лікуванні серцево-

судинних хвороб та при їх профілактиці може призвести до позитивних результатів у вигляді зменшення кількості несвоєчасно виявлених порушень у роботі серцево-судинної системи.

2.2 Математична модель

У різних медичних задачах (збір інформації про хворого, діагностика, вибір тактики рішення) лікар стикається із дуже глобальною проблемою – проблемою прийняття рішень. При цьому з кожним роком виростають вимоги до точності, якості діагнозу і його достовірності. Іншими словами – до його істинності.

Математика як наука є головним інструментом у процесі пошуку істини, і, отже, в процесі прийому рішень.

Під прийняттям рішень розуміється особливий процес людської роботи, орієнтований на вибір найбільш припустимого варіанту рішення питання. Прикладом може служити процес прийому рішення про тип (форму) хвороби за відомою вихідною інформацією такою як результати аналізів чи зовнішні прояви хвороби або рішення проблеми, так званого, гуртового вибору рішень, коли головне завдання полягає в тому, щоб показати «справедливі» положення обліку персональних відборів, що зумовлюють до розумного групового рішення. Дане завдання вирішується, наприклад, способом проведення консилиуму, коли кожен з його учасників висловлює свою особисту думку щодо плану лікування, і, в заключному підсумку, вибирається один, найкращий варіант.

Медицина являє собою слабоструктуровану область знання, що створює серйозні труднощі для процесу прийняття рішень. В одних інцидентах, що характеризуються класичними проявами недуги, гіпотеза або навіть остаточне рішення з'являється вже в процесі огляду, в інших – тільки після професійного обстеження та медичного дослідження. Важливо зауважити, що послідовність діагностичних досліджень може покорятися корекціям, а іноді й докорінній трансформації, залежно від одержуваних у процесі обстеження наслідків.

Швидкість прийняття рішення залежить як від фаху та діагностичного "чуття" лікаря-діагноста, так і від особливостей прояву захворювання у конкретного пацієнта.

За попередні роки відбувся надзвичайний стрибок у розвитку персональних комп'ютерів і програмного забезпечення. Було створено багато програмних продуктів, які мають змогу реалізувати складні методики прийняття рішень. Це надає перспективу перевіряти прийняті рішення, перед цим завчасно побудувавши модель медико-біологічного процесу.

Свобідний доступ до програмного забезпечення дозволяє розв'язати медичні завдання в інтерактивному або діалоговому, а найважливіше в онлайновому режимі. Інтерактивний – це режим роботи, який відтворює взаємодію між людиною і комп'ютером. Застосовуючи інтерактивні процедури користувач/лікар може в діалоговому розпорядку знаходити якнайкращі рішення задачі, переміняючи умови – обмеження задачі або параметри цільових функцій. На кожній ітерації лікар, як особа, яка приймає рішення (ОПР) для наступного дослідження може порадити нові умови в задачі.

Інтерактивні процедури дають змогу ефективно розбити працю: комп'ютер виконує те, що він робить краще за все, а саме обробляє дані, а ОПР на основі вже нової інформації, наданої комп'ютером розробляє та обирає методи для отримання кращого вирішення. При цьому головне рішення та головна роль завжди залишається за людиною. Прийняти рішення – означає здійснити важливий вибір з деякого комплексу альтернатив. Не надаючи уваги різноманітності існуючих проблем, виділяють такі основні фази процедури прийняття рішення:

- дефініція мети;
- формування переліку альтернатив або визначення купи допустимих вирішень;
- створення оцінки, яка дає змогу порівнювати альтернативи, тобто проводити задачу оцінювання;
- обрання найкращого рішення з купи допустимих вирішень тобто проведення оптимізаційного завдання;

У теорії прийняття рішень загальна сукупність перерахованих задач являє собою суцільну проблему прийняття рішень.

Теоретичною базою рішення перших трьох завдань є системний аналіз, а четверта являє собою теорію математичного програмування.

Якщо вирішення задачі не загальновідомо, тобто відсутній аналог або вирішення неоднозначно, то на головний план виступають проблеми дефініції методу пошуку вирішення.

Переважає більшість цих методик фундаментується на стратегіях повного перебору, імпліцитного, або неявного перебору, та евристичного пошуку.

Стратегія повного перебору використовується при відсутності достатньої кількості початкової інформації про задачу і відносно незначній множині альтернатив (до 10^3 частин при ручному підрахунку, і до 10^9 для «ЕОМ»).

Імпліцитний перебір залучає велику групу, так названих, градієнтних методів, таких як симплекс метод, метод мінімальної вартості, динамічне програмування і т. ін. Усі вони засновані для розгляду на кожному кроці розшуку не всього простору задачі, а тільки деякого її фрагмента чи сектора.

Евристичні методи – методи вирішення задач, засновуються на евристиці, або евристичному міркуванні. А саме на використанні певних правил і прийомів, узагальнюючих булий досвід і інтуїцію вирішального. Евристичні міркування – в широкому сенсі підрозділ психології, котрий вивчає природу розумової діяльності індивіда, інтелектуальних операцій, при вирішенні ним різних задач. До евристичних методів відносять лише ті, які прив'язані безпосередньо зі здібностями людського організму, з несподівано запропонованими вирішеннями, тобто безпосередньо з терміном еврика, осяяння.

Їх застосування також має сенс при жорстких ресурсних обмеженнях, наприклад при діях в екстремальних або невизначених ситуаціях).

Як вже згадувалося вище, прийняття рішень, по змісту, є ні що інше як звичайний вибір. Прийняти рішення – значить обрати якийсь конкретний

варіант дій з певної безлічі варіантів та різновидів. Такі варіанти вибору перейнято називати альтернативами.

Безліч альтернатив залежить від декількох факторів. По-перше, від підручної бази знань: або алгоритм вирішення завдань вже міститься в базі, або алгоритму в базі на даний момент немає, але присутній аналог, або ж задача не лічить аналогів в базі знань. По-друге, безліч альтернатив має залежність від проблемної ситуації: або вирішується вже нова задача, або змінюється одна з умов функціонування системи, або ще могла з'явитися нова інформація, або відбувся збій у системі або в її структурних елементах.

Плодом прийняття рішення називається явище, або подія, можливість появи якого продиктована даним рішенням.

Система переваг – правила, критерії, при дотриманні яких порівнюються альтернативи і приймаються відповідні рішення.

Рішення – рішення, або альтернативи, що задовольняють правилам, що містяться в системі переваг.

Загальну задачу прийняття рішень чи як її називають завданням вибору, можна сформулювати наступним чином .

Нехай в даному випадку X – безліч альтернатив (рішень), Y – маса можливих результатів (наслідків). Очікуватиметься існування причинного зв'язку між вибиранням деякої альтернативи і настанням відповідного результату. Крім цього, передбачається наявність механізму оцінки доброякісності вибору – звичайно ж шляхом оцінки якості результату. Необхідно обрати найліпшу альтернативу, для якої належний результат має якнайкращу оцінку якості.

Опираючись на зв'язки між рішеннями і наслідками прийнята подальша класифікація задач прийняття рішень.

Детермінована задача прийняття рішень

Їй відповідає якнайбільш простий вид взаємозв'язку – детермінований, коли кожна альтернатива приводить до певного але єдиного результату. У такому випадку існує функціональна залежність між альтернативою X_i і наслідком Y_i .

У разі, коли кожній альтернативі відповідає не лише один єдиний результат, тобто значиться недетермінований тип зв'язку, то задачі прийняття рішень розщеплюються на два підкласи:

- а) задача прийняття рішень в умовах ризику;
- б) задача прийняття рішень в умовах стохастичної або імовірнісної, невизначеності.

У першому випадку (а), на переміну від випадку (б), для кожної альтернативи x_i відвічає функція щільності ймовірностей на множині результатів Y .

На рисунку 2.1 кожна стрілка охарактеризовується вагою, тобто числом P_{ij} - ймовірністю випадання результату y_j при виборі альтернативи x_i .

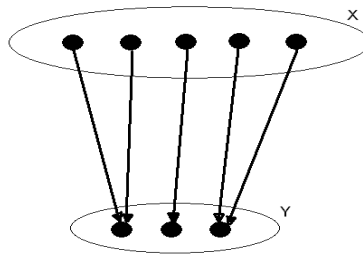


Рисунок 2.1 – Детермінований тип зв'язку

В умовах невизначеності недетермінованого зв'язку альтернатива-результат виникають два типи завдань:

- задачі прийняття рішень в таких умовах пасивної взаємодії ОПР та зовнішнього середовища, тобто зовнішній осередок веде себе інертно;
- задачі прийняття рішень в умовах конфлікту або гри, у якій зовнішнє середовище поводить себе активно щодо ОПР, що проявляється певними діями іншої людини.

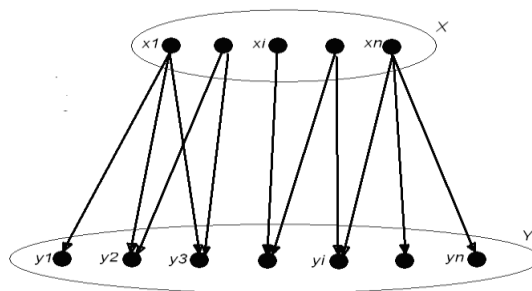


Рисунок 2.2 – Імовірнісний тип зв'язку

На рисунку 2.3 інцидент 1 відповідає прийому рішення в умовах визначеності: точками y_1, y_2, y_3 позначені результати, які відповідають вибору альтернатив x_1, x_2, x_3 (три альтернативи і три певних результати). Випадок 2 охарактеризовує задачу прийняття рішення в умовах невизначеності: після обрання будь-якої з альтернатив x_1, x_2 або x_3 може бути показаний лише інтервал розташування відповідного результату y_i . Випадок 3 відтворює ситуацію відбору в умовах ризику. Показані графіки відповідних щільностей ймовірностей подій в залежності від обрання альтернативи x_1, x_2 або x_3 .

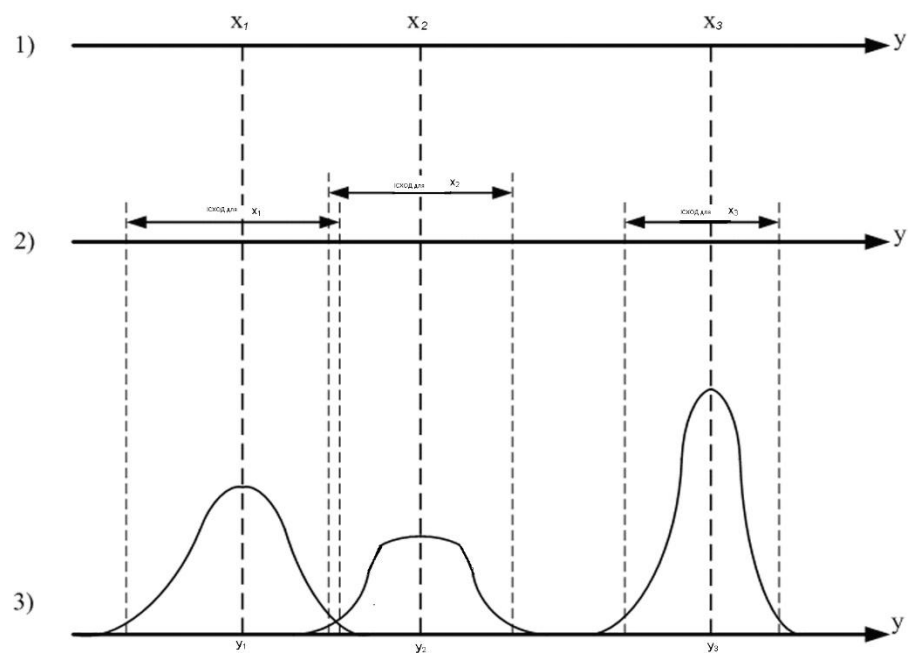


Рисунок 2.4 – Види зв'язків між альтернативами та результатами

Всі вимоги, які були сформульовані в реальних задачах та записані у вигляді математичних висловів, складають, так звану, математичну постанову задачі. Процес математичної постанови задачі і подальшого її вирішення можна відобразити у вигляді декількох етапів.

Дослідження об'єкта являє собою аналіз властивостей функціонування об'єкта. На цій стадії виявляються чинники, що впливають на об'єкт і позначається ступінь їхнього впливу; вивчаються та аналізуються характеристики об'єкта при різних умовах та умовах; обираються оптимізуючі критерії, тобто цільові функції.

Описове моделювання має в основі встановлення і фіксацію основних зв'язків і залежностей між характеристиками процесу або події згідно оптимізуемому критерію.

Математичне моделювання

Вибір і виготовлення методу рішення

Для більш кращого розуміння наведемо приклад математичної моделі задачі про харчування. Бодай в раціон пацієнта входить 3 повністю різних ситних речовини (все тіж білки, жири, вуглеводи) і необхідно їх, не менше b_1 , b_2 , b_3 одиниць, відповідно. Вони знаходяться в 5 різних продуктах, які можна придбати по вартості c_1, c_2, \dots, c_5 за одиницю.

Одиниця i -го продукту містить a_{ij} одиниць j -ої поживної речовини, тобто, наприклад, a_{23} показує, що в одиниці другого товару третього живильної речовини буде a_{23} одиниць.

Отже яку кількість продуктів кожного з видів потрібно купити, щоб їх ціна була мінімальна, а раціон пацієнта обіймав всі необхідні речовини в належній кількості?

Цільова функція цього завдання – мінімізувати за кількістю продуктів (x_1, \dots, x_5) витрати на харчування:

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 = \sum_{i=1}^5 c_i x_i \rightarrow \min_{x_1, \dots, x_5}$$

Умови у обмеженні задачі наступні: кількість первинної поживної речовини повинно бути не менше ніж b_1 , тобто

$$a_{11}x_1 + a_{21}x_2 + a_{31}x_3 + a_{41}x_4 + a_{51}x_5 = \sum_{i=1}^5 a_{i1}x_i \geq b_1$$

Також подібно для інших поживних речовин:

$$a_{12}x_1 + a_{22}x_2 + a_{32}x_3 + a_{42}x_4 + a_{52}x_5 = \sum_{i=1}^5 a_{i2}x_i \geq b_2$$

$$a_{13}x_1 + a_{23}x_2 + a_{33}x_3 + a_{43}x_4 + a_{53}x_5 = \sum_{i=1}^5 a_{i3}x_i \geq b_3$$

Очевидно, що кількість продуктів – це невід'ємна величина

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0.$$

Допустимим вирішенням називають такий набір значень шуканих змінних, який буде задовольняти поставленим умовам, тобто обмеженням самого завдання.

Вирішенням завдання буде те рішення з купи допустимих рішень, при якому цільова функція досягне свого найбільшого або найменшого, дивлячись від постановки обмеження, значення.

Вирішення задачі на ЕОМ. Завдання, які обрисовують поведінку конкретних об'єктів, як правило, мають дуже багато несталих величин і велику кількість залежностей між ними. Тому в розумні терміни вони можуть бути вирішені лише за допомогою ЕОМ.

Аналіз отриманого рішення. Аналіз рішення може бути двох типів формальним та змістовним. При формальному аналізі перевіряють відповідність отриманого рішення побудованої математичної моделі. При змістовному розборі перевіряють відповідність отриманого рішення тому реальному об'єкту, який моделювався на початку. В результаті змістовного аналізу в модель можуть бути внесені певні зміни і увесь процес може повторюватися. Високу важливість змістовного аналізу можна підтвердити на прикладі. Коли вперше вирішили задачу про харчування, то в якості фактору оптимізації обрали мінімум витрат, а в умови обмеження вводили тільки вимоги по калорійності їжі. Вирішення задачі було таким: харчуватися слід оцтом, який входить до складу різноманітних продуктів харчування, в такому випадку буде забезпечена і калорійність, і вартість буде мінімальна.

Аналіз стійкості рішень. Для перевірки стійкості рішення у вихідні дані вносяться деякі зміни в кордонах можливих погрішностей або інтервалів існування прикмет, а потім досліджується поведінка рішення або аналітичними або чисельними методами.

Розробкою методик вирішення завдань, які містять цільову функцію і умови, тобто обмеження, займаються в підрозділі математики, названому математичним програмуванням.

Математичне програмування – це математична дисципліна, в якій вчать теорію та методи вирішення задач про віднаходження екстремумів функцій у множинах, що визначаються рівностями та нерівностями.

Задачі з кількома цільовими функціями, або з однією цільовою функцією, але приймаючої векторні значення або як варіант значення ще більш складної природи, прийнято називати багатокритеріальними задачами. Їх вирішують за при допомозі відомостей до завдань з єдиною цільовою функцією або на базі використання «теорії ігор», в якій очікується, що особа, яка приймає рішення грає в деяку азартну гру, силкуючись добитися максимально гарного та вдалого результату.

Застосування «теорії ігор» в медичній сфері пов'язане, наприклад, з наявністю конфліктів між пацієнтом і лікарем. Бувають різні випадки, найчастіше це: не виправдалися сподівання на позбавлення від захворювання, невдоволення суб'єктивних очікувань люб'язного і попереджального ставлення, підвищені запити хворого по відношенню до персоналу і т.ін. Очевидні конфлікти і в медичному менеджменті – між урядовцями і практичними лікарями, страховими коопераціями та звичайними лікарнями і т.ін. Конфлікт сторін є одним з найважливіших елементів гри і абсолютно нормальним явищем суспільного існування. Конфлікт може розгортатися на внутрішньо-особистісному рівні, ступені міжособистісних відносин, між соціальними групами, чи навіть державами. Формування конфлікту найчастіше пояснюється об'єктивними умовами: будь-який розріст обумовлює формування конфлікту, якого зазвичай не можна минути. Вивчаючи питання розвитку конфліктів, постає необхідність зосередитися на прийомах виходу з них, переходу у безпечну обстановку, яка може бути контрольованою і, отже, зміненою самою людиною. Таким чином і постає необхідність вирішення конфліктів, в тому числі з застосуванням теорії прийняття рішень математичною методикою «теорії ігор».

Якщо в процесі ухвалення рішення ОПР не втрачає і не набуває інформацію, то ухвалення такого рішення ми можемо розглянути як миттєвий акт. Подібні завдання іменуються статичними. І навпаки, якщо ОПР в процесі

ухвалення рішення отримує чи втрачає якусь інформацію, то таке завдання вже прийнято називати динамічним. У динамічних задачах дуже доречно приймати рішення поетапно.

У своїй професійній діяльності кожен лікар безумовно стикається з ситуаціями, в яких інформація виявляється не повною або не закінченою, а тільки побічно пов'язаною з тим, що йому дійсно потрібно знати про хворого. У таких інцидентах лікар вимушений приймати своє рішення про діагноз і лікування в умовах неконкретності. Але насамперед додаткова кількість інформації, далеко не завжди допоможе у зменшенні невизначеності, а інколи й навпаки, тільки погіршить ситуацію.

Діагноз і вибір дії – це терміни, що використовуються у досліджах процесу прийняття рішення в різних сферах діяльності людини. У медицині їх можна зіставити з поняттями - діагностика і лікування. Процеси прийняття рішення про діагноз і про вибір лікування найтіснішим чином взаємопов'язані і зобов'язані розглядатися вкупі.

Як вже було описано вище, додаткова інформація не завжди є необхідністю для зняття невизначеності, з якою стикається медичний працівник при роботі з конкретним пацієнтом. Тому дуже актуальним є підбір правильних методів, які зможуть допомогти лікарю найбільш ефективним способом прийняти рішення про діагноз і обрати оптимальне та найкраще лікування.

Наступний етап - вибір нульової гіпотези. Це вирішальний момент при перевірці значущості отриманих даних і результатів. «Значимий» результат рівносильний тому, що нульова гіпотеза має бути відкинута, оскільки вона виявилася неправильною. Подібним чином, ОПР як би припустився помилки при обранні гіпотези, можливо внаслідок незнання об'єкта який досліджувався або в результаті недостатньої компетентності, чи недостовірності даних. Коли отримані результати не відповідають початковим припущенням, тоді значущі результати є «хорошими». Неодноразово повідомляється, що отримані результати значимі при рівності значимості $p = 0,05; 0,01$ (ступінь значимості $p = 1 - \alpha$).

Підсумовуючи, можна висловитись, що сама суть перевірки гіпотези полягає в тому, що очікуваний об'єкт порівнюється з істинним еталоном і в результаті зіставлення виноситься правильне або помилкове твердження, яке далі і називається рішенням задачі. Сенс перевірки статистичних гіпотез полягає в тому, щоб за даними випадкової вибірки перейти або відвернути гіпотезу з найменшим ризиком помилки. Зазвичай гіпотезу яка перевіряється прийнято називати «нульовою гіпотезою» і позначають H_0 .

Нульовими вважаються ті гіпотези, котрі стверджують, що відмінності між порівнюваними величинами відсутні, а спостережувані аномалії пояснюються лише випадковими коливаннями у вибірках. Решта гіпотез, котрі різняться від H_0 , та протиставляються їй, називаються альтернативами і позначаються H_1 .

Правильність чи хибність рішення залежить від того справжня гіпотеза чи неправильна. Якщо гіпотеза істинна і ОПР її зараховує, то, з його боку, це правильне вирішення питання. Статистично вірне рішення чисельно характеризується довірчою ймовірністю, яку позначають як α . Але, якщо гіпотеза істинна, а ОПР з відхилив її з якихось причин, то це – його неправильне рішення. Його прийнято називати «помилкою першого роду» і вона оцінюється ймовірністю, яку позначають як $1 - \alpha$. У разі, якщо гіпотеза хибна, але ОПР її приймає, то це, звичайно ж знову, помилкове рішення. Таке рішення прийнято називати «помилкою другого роду», і оцінювати ймовірністю, яку позначають як $1 - \beta$. Якщо гіпотеза помилкова, то вірним рішенням для ОПР буде її відхилення, з імовірністю β . У наукових досліджах переважають ті способи перевірки гіпотез, за допомогою котрих мінімізується можливість прийняти помилкової гіпотези як вірної, тобто помилок другого ряду. Властивість якогось з цих способів перевіряти безпомилково та відхилити помилкові гіпотези називають потужністю цього способу, або потужністю критерію. Кількісно потужність можна охарактеризувати умовною ймовірністю β , а $1 - \beta$ -ймовірність помилок другого роду.

Таким чином, при прийнятті статистичного рішення завжди залишається ймовірність припустити помилку. Такі помилки бувають двох типів:

– відхилення H_0 в той час, як H_0 насправді вірна (тобто помилка першого роду або помилкою α -типу);

– прийняття H_0 в той час як H_0 справді являється помилковою (тобто помилка другого роду, або помилкою β -типу).

Все вищесказане описано у таблиці 2.1.

Таблиця 2.1 – Опис гіпотези

| Характеристика гіпотези і відповідна дія ОПР | Статистична характеристика рішення | Визначення ймовірності |
|--|------------------------------------|------------------------|
| гіпотеза вірна, і ОПР її приймає | Коректно прийняте рішення | α |
| гіпотеза вірна, але ОПР її відхиляє | Допущена помилка I роду | $1 - \alpha$ |
| гіпотеза є помилковою, але ОПР її приймає | Допущена помилка II роду | $1 - \beta$ |
| гіпотеза є помилковою і ОПР її відхиляє | Коректно прийняте рішення | β |

Ймовірність похибки першого роду називають рівнем значущості $1 - \alpha$.

Ймовірність прийняття вірного рішення H_0 , рівну α , називають надійністю, тобто ймовірність не зробити помилку першого роду.

Ймовірність не зробити похибку другого роду, рівну β , називають потужністю критерію.

До наукових гіпотез відносяться такі, які в майбутньому можуть бути випробувані. Емпірична перевірка гіпотези носить назву верифікації.

Для статистичної перевірки медичних гіпотез застосовують такі критерії як: t-Ст'юдента, λ -Колмогорова, F-Фішера, χ^2 -Пірсона, G-Кохрана та ін.

Такі критерії як λ , X^2 , є непараметричними критеріями, тому що слугують для перевірки гіпотез про поділи в цілому. Критерії t, F, G – параметричні, так як вони використовуються для перевірки гіпотез про параметри розподілу.

При використанні формули Байєса вдається акумулювати інформацію, що надходить з всіляких різних джерел з метою підтвердження або не підтвердження тієї чи іншої гіпотези (діагнозу). Формула Байєса дозволяє за допомогою умовних ймовірностей спільно використовувати спостережувані дані та інформацію, відому раніше, для вирішення завдання диференціального діагностування.

Наведемо приклад. Нехай у хворого підозрюється грип. Тобто маємо деяку гіпотезу H , яка полягає в тому, що у хворого буде саме грип, а не щонебудь інше. Будемо вважати, що в медичних установах на основі раніше отриманих статистичних даних відома апріорна (первісна) ймовірність $P(H)$ того, що пацієнт в даний час року і в даній місцевості захворіє на грип. Нехай ознака D означає наявність високої температури у даного конкретного пацієнта. Формула Байєса дозволяє отримати ймовірність грипу при наявності у пацієнта високої температури $P(H/D)$ (кінцева або апостеріорна ймовірність). Тобто ми хочемо на основі наявної інформації уточнити апріорну ймовірність істинності гіпотези H . Відповідно до формули Байєса маємо:

$$P(H/D) = \frac{P(D/H)P(H)}{P(D)} = \frac{P(D/H)P(H)}{P(D/H)P(H) + P(D/\bar{H})P(\bar{H})}$$

Для того щоб скористатися формулою Байєса для прикладу необхідно знати ймовірності: $P(D/H)$ – ймовірність високої температури при грипі; $P(D/\bar{H})$ – ймовірність високої температури при відсутності грипу. Ми припускаємо, що обидві ці ймовірності нам відомі. Вони здобуваються при обробці накопичених раніше статистичних даних. Ясно, що всі три числа $P(H)$, $P(D/H)$, $P(D/\bar{H})$, можуть бути отримані заздалегідь і не залежать від даних конкретного пацієнта. Знаючи що $P(\bar{H}) = 1 - P(H)$, ми можемо скористатися формулою Байєса. Нехай відомо, що:

$$P(H) = 0,001; \quad P(\bar{H}) = 1 - P(H) = 0,999$$

Нехай відомо також:

$$P(D/H) = 1,0;$$

$$P(D/\bar{H}) = 0,01$$

Тоді за формулою Байєса отримуємо:

$$P(H / D) \cong 0,009$$

Таким чином, ймовірність захворювання на грип при надходженні свідчення про високу температуру збільшилася і складає 0,009 у порівнянні з 0,001 (вихідна апіорна ймовірність).

Теорема Байеса застосовується при прийнятті рішень в експертних системах. Схема роботи байесовської експертної системи полягає в наступному. Спочатку ми маємо апіорну ймовірність $P(H)$ (на прикладі – у хворого грип), яка зберігається в базі знань. Але, отримавши свідчення D (висока температура) і перерахувавши цю ймовірність за формулою Байеса, ми можемо записати її на місце $P(H)$. Отримання чергового свідчення призводить до нового оновлення (збільшення або зменшення) цієї ймовірності. Кожен раз поточне значення цієї ймовірності буде вважатися апіорною для застосування формули Байеса. Зрештою, зібравши всі відомості, що стосуються всіх гіпотез (наприклад, діагнозів хвороб), експертна система приходять до остаточного рішення, виділяючи найбільш ймовірну гіпотезу в якості результату експертизи.

Стосовно до завдання діагностування формула Байеса надає змогу вибрати одну з декількох діагностичних гіпотез, фундаментуючись на обчисленні ймовірностей захворювань за ймовірностями виявлених у пацієнтах симптомів.

Теорема Байеса говорить, що кінцева ймовірність гіпотези $P(H/D)$ пропорційна її вихідній ймовірності $P(H)$, помноженої на її правдоподібність $P(D/H)$.

Дуже важливу роль при обрахунках за формулою Байеса відіграє відношення правдоподібності, а саме відношення двох правдоподібностей.

$$\frac{P(D/H_i)}{P(D/H_j)} (i \neq j)$$

Ця величина може висловлювати, наприклад, відношення ймовірності комплексів симптомів при захворюванні астмою чи іншим захворюванням до ймовірності того ж комплексу симптомів у контрольній групі. Наведемо приклад.

Завдання з області «прийняття рішення» виникають у тих випадках, коли завдання настільки важке, що для його постанови і вирішення не має можливості відразу знайти визначений відповідний апарат формалізації або коли процес постанови завдання починає вимагати участі фахівців з декількох різних областей знань. Для таких складних ситуацій технологія «прийняття рішення» має спеціалізовані підходи, прийоми та методики. На початку йде визначення області проблеми прийняття рішення, далі виявляють фактори, що впливають на її рішення, підбирають методи і прийоми, які дозволять сформулювати задачу таким чином, аби рішення могло бути прийнятим. Далі отримують вираз, що прив'язує мету із засобами її досягнення. Все це втілюється в математичних моделях – різних критеріях.

Якщо вдається отримати вираз, що пов'язує мету із засобами, то завдання майже стовідсотково вирішується. Отримати такі вирази досить легко, якщо нам відомий закон, котрий пов'язує мету з засобами. Якщо закон не знайомий, то постає необхідність обрати інакший спосіб відображення проблемних ситуацій. Можна визначити закономірності базуючись на статистичних дослідженнях або функціональних залежностях. Якщо і це не піддається зробити, то вибирають або розробляють теорію, в якій міститься ряд певних тверджень та правил, які дозволять сформулювати концепцію і сконструювати на її основі хід прийняття рішення. У випадку якщо і теорії не існує, то висувається гіпотеза і на її основі народжуються імітаційні моделі, за допомогою яких досліджуються можливі різновиди дослідів.

Для того, щоб допомогти у короткі терміни побудувати завдання, проаналізувати цілі, встановити можливі засоби, підібрати необхідну інформацію, а в ідеальному випадку – отримати вираз, який буде пов'язувати мету із засобами, то тоді застосовують системні уявлення, прийоми і методи системного аналізу.

2.3 Огляд методів розв'язання

У сьогоднішній у медицині стоїть потреба в інтелектуальному аналізі медичних даних. Сучасні лікарські моніторингові системи зобов'язані проводити централізований контроль становища пацієнтів. Моніторинг електрокардіограми(ЕКГ) пацієнтів є важким і не вирішеним завданням на даний момент.

Сигнал ЕКГ, зображений на рисунку 2.5 представляє собою часовий ряд, який є нестационарним та схильним до численних видів різноманітних перешкод.

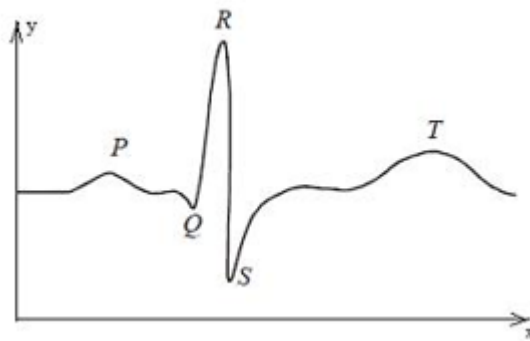


Рисунок 2.5 – Приклад ЕКГ: інтервали P, Q, R, S, T

Метод поверхневої реєстрації електрокардіограм.

Найпоширенішим методом довготривалого реєстрування поверхневої ЕКГ є метод Холтера або холтерівське моніторування (ХМ). У медицині монітор Холтера (частіше просто Холтер) - це тип пристрою амбулаторної електрокардіографії, портативний прилад для моніторингу серця (моніторинг електричної активності серцево-судинної системи) протягом щонайменше 24 до 48 годин.

Найбільш поширене застосування Холтера - це для моніторингу серцевої діяльності ЕКГ (електрокардіографія або ЕКГ). Його тривалий період запису іноді корисний для спостереження за випадковими серцевими аритміями, які важко було б визначити за короткий період. Для пацієнтів, які мають більш мінущі симптоми, можна використовувати монітор серцевих подій, який можна носити протягом місяця або більше.

Монітор Холтера був розроблений в дослідницькій лабораторії Холтера в Гелени Монтані експериментальними фізиками Норманом Дж. Холтером та Біллом Гласкоком, який розпочав роботу над радіотелеметрією в 1949 році. на

початку 1950-х вони спрямовували свої зусилля на розробку носячого приладу моніторингу серця. Монітор Холтера був випущений для комерційного виробництва в 1962 році.

Коли він використовується для дослідження серця, подібно до стандартної електрокардіографії, монітор Холтера записує електричні сигнали від серця через серію електродів, прикріплених до грудної клітки. Електроди розміщуються над кістками, щоб мінімізувати артефакти від м'язової діяльності. Кількість та положення електродів залежать від моделі, але більшість моніторів Холтера використовують від трьох до восьми. Ці електроди з'єднані з невеликим предметом обладнання, який кріпиться до поясу пацієнта або підвішується на шиї, ведучи журнал електричної активності серця протягом усього періоду запису. Система Холтера 12 електродів також доступна, коли потрібна точна інформація ЕКГ-сигналу для аналізу точної природи та походження сигналу ритму.

Старі пристрої використовували магнітофонні стрічки або стандартну аудіокасету C90 або C120 і працювали зі швидкістю 1,7 мм / с або 2 мм / с для запису даних. Після того, як було зроблено запис, його можна було відтворити та проаналізувати зі швидкістю 60х, тому 24 години запису можна було проаналізувати за 24 хвилини. Більш сучасні блоки записують файл EDF на цифрові флеш-пам'яті. Дані завантажуються в комп'ютер, який потім автоматично аналізує вхід, підраховуючи ЕКГ комплекси, обчислюючи підсумкові статистичні дані, такі як середній серцевий ритм, мінімальний та максимальний частоти серцевих скорочень, і знаходить кандидатурні області в записі, гідному подальшого вивчення фахівцем.

Кожна система Холтера складається з двох основних частин - апаратного забезпечення (називається монітором або рекордером) для запису сигналу та програмного забезпечення для огляду та аналізу запису. Розширені рекордери Holter здатні відображати сигнал, що дуже корисно для перевірки якості сигналу. Дуже часто на передній ділянці є також "кнопка пацієнта", яка дозволяє пацієнту натискати на неї в конкретних випадках, таких як хвороба, лягання спати, прийняття таблеток тощо; Потім буде записано спеціальний

знак, який дозволить лікарям або технікам швидко визначити ці ділянки під час аналізу сигналу.

Коли запис ЕКГ-сигналу закінчено (як правило, через 24 або 48 годин), лікар повинен виконати аналіз сигналу. Оскільки для перегляду такого довгого сигналу вкрай потрібен час, в програмному забезпеченні кожного пристрою Холтера є інтегрований процес автоматичного аналізу, який автоматично визначає різного роду серцебиття, ритми тощо. Однак успіх автоматичного аналізу полягає в дуже тісно пов'язаний з якістю сигналу. Сама якість головним чином залежить від прикріплення електродів до тіла пацієнта. Якщо вони неправильно прикріплені, електромагнітні порушення можуть вплинути на сигнал ЕКГ, що призведе до дуже шумних записів. Якщо пацієнт швидко рухається, викривлення буде ще більшим. Такий запис тоді дуже важко обробити. Крім приєднання та якості електродів, існують і інші фактори, що впливають на якість сигналу, такі як тремор м'язів, швидкість дискретизації та роздільна здатність оцифрованого сигналу (пристрої високої якості пропонують більшу частоту дискретизації).

Автоматичний аналіз зазвичай надає лікарю інформацію про морфологію серцебиття, вимірювання інтервалу серцебиття, мінливість серцебиття, огляд ритму та щоденник пацієнта (моменти, коли пацієнт натискав кнопку пацієнта). Вдосконалені системи також виконують спектральний аналіз, оцінку ішемічного навантаження, графік активності пацієнта або аналіз сегмента PQ. Ще одна вимога - можливість виявлення та аналізу імпульсів пейсмейкера. Така здатність може бути корисною, коли лікар бажає перевірити правильність основної функції кардіостимулятора.

Підсумковою частиною дослідження є фінальний протокол, завданням якого є подання лікарю-кардіологу максимально інформативного висновку, з обов'язковим відображенням тих параметрів ритму серця, що здатні вплинути на тактику лікування пацієнта, та прогнозів щодо одужання.

Завдання інтелектуального аналізу ритмів ЕКГ у моніторингових системах даних допомагає в обробці та завчасному виявленні аномалій у людей для діагностики проблем серцево-судинної системи. Для оцінки сегменту

зубців ST і T була використано базу даних під назвою MIT-BIH Arrhythmia, яка часто використовується при виявленні випадків інфаркту міокарда та для ґрунтовних досліджень динаміки роботи серця.

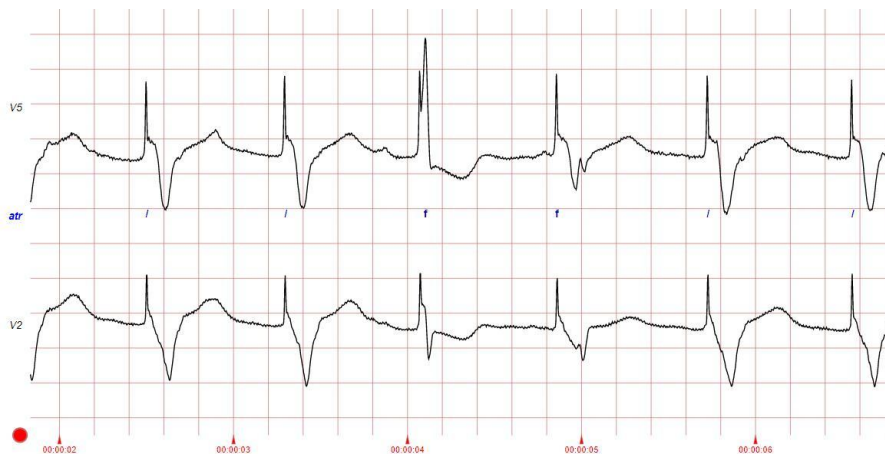


Рисунок 2.6 – Приклад ЕКГ бази даних аритмії MIT-BIH

Вхідним файлом для аналізу бази даних аритмії MIT-BIH може бути файл .beat. Структура такого файлу значиться стандартною для медичних баз даних часових рядів.

Для реалізації інтелектуального аналізу таких часових рядів ми використаємо медичний набір даних часових послідовностей електрокардіограм (ЕКГ) серцебиття.

Сигнали відповідають фігурам електрокардіограми (ЕКГ) серцевих розширень та скорочень для нормального випадку і випадків ураження різними видами захворювання серцево судинної системи. Ці сигнали попередньо оброблялися і сегментувалися. Більшість медичних баз даних, використовують коди приміток. Коди показані в таблиці 2.2.

Таблиця 2.2 – Кодування серцевих ударів

| Код | Опис |
|-----|----------------------------------|
| N | Звичайний удар |
| L | Білий блок у гільці лівого пучка |
| R | Розгалуження правого боку |
| A | Передчасний удар передсердя |

Продовження таблиці 2.2

| Код | Опис |
|-----|--|
| J | Передчасний нодальний удар |
| V | Шлуночки скорочуються передчасно |
| r | Передчасне скорочення шлуночків R-на-T |
| F | Злиття шлуночків і нормального биття |
| i | Втеча з передсердя |
| / | Збивка в темпі |
| Q | Удар не класифікувався |
| ? | Удар не прокласифікувався під час навчання |
| j | Перебіг удару |
| n | Надшлуночковий бічний потік |

Діаграми розкиду або кореляційні ритмографії

Діаграма розкиду (також його називають розсіювачем, діаграмою розсіювання) - це тип графіку або математичної діаграми, що використовує декартові координати для відображення значень для типово двох змінних для набору даних. Сутність методу кореляційної ритмографії полягає в графічному відображенні послідовних пар кардіоінтервалів (попереднього і наступного) в двомірній координатній площині.

Діаграми розсіювання подібні до лінійних графіків тим, що вони використовують горизонтальну та вертикальну осі для побудови точок даних. Однак вони мають цілком конкретне призначення. Діаграми розсіювання показують, на яку зміну впливає інша. Зв'язок між двома змінними називається їх кореляцією.

Діаграми розсіювання зазвичай складаються з великого масиву даних. Чим ближче дані точки наближаються, коли будується пряма лінія, тим вище кореляція між двома змінними або тим сильніший зв'язок.

По осі абсцис відкладається величина $R-R_n$, а по осі ординат - величина $R-R_n + 1$. Графік і область точок, отриманих таким чином, називається кореляційної ритмограмми або скаттерограмми (scatter-розсіювання). При

побудові скаттерограмми утворюється сукупність точок, центр яких розташовується на бісектрисі. Відстань від центру до початку осей координат відповідає найбільш очікуваній тривалості серцевого циклу. Величина відхилення точки від бісектриси вліво показує, наскільки даний серцевий цикл коротше попереднього, вправо від бісектриси - наскільки він довший за попередній. Пропонується обчислювати наступні показники скаттерограмми:

- Довжина основного (без екстрасистол і артефактів) «хмари» (довга вісь еліпса - L) відповідає варіаційному розмаху. За фізіологічною змістом цей показник не відрізняється від $SDNN$, тобто відображає сумарний ефект регуляції BCP, але вказує на максимальну амплітуду коливань тривалості інтервалів R-R;

- Ширина скаттерограмми (перпендикуляр до довгої осі, проведений через її середину - w);

- Площа скаттерограмми обчислюється за формулою площі еліпса:

$$S = (\rho \cdot L \cdot w) / 4.$$

Якщо точки даних проводять пряму лінію, що йде від початку до високих значень X і Y , то тоді кажуть, що змінні мають позитивну кореляцію. Якщо лінія переходить від високого значення на осі y вниз до високого значення на осі x , змінні мають негативну кореляцію.

Ідеальній позитивній кореляції присвоюється значення 1. Досконалому негативному співвідношенню задається значення -1. Якщо кореляції абсолютно немає, наведене значення дорівнює 0. Чим ближче число до 1 або -1, тим сильніша кореляція або сильніша зв'язок між змінними. Чим ближче число до 0, тим слабша кореляція. Тож щось, що, здається, співвідноситься у позитивному напрямку, може мати значення 0,67, тоді як щось із надзвичайно слабкою негативною кореляцією може мати значення -0,21. На рисунку 2.7 показані види кореляції скатерограмм.

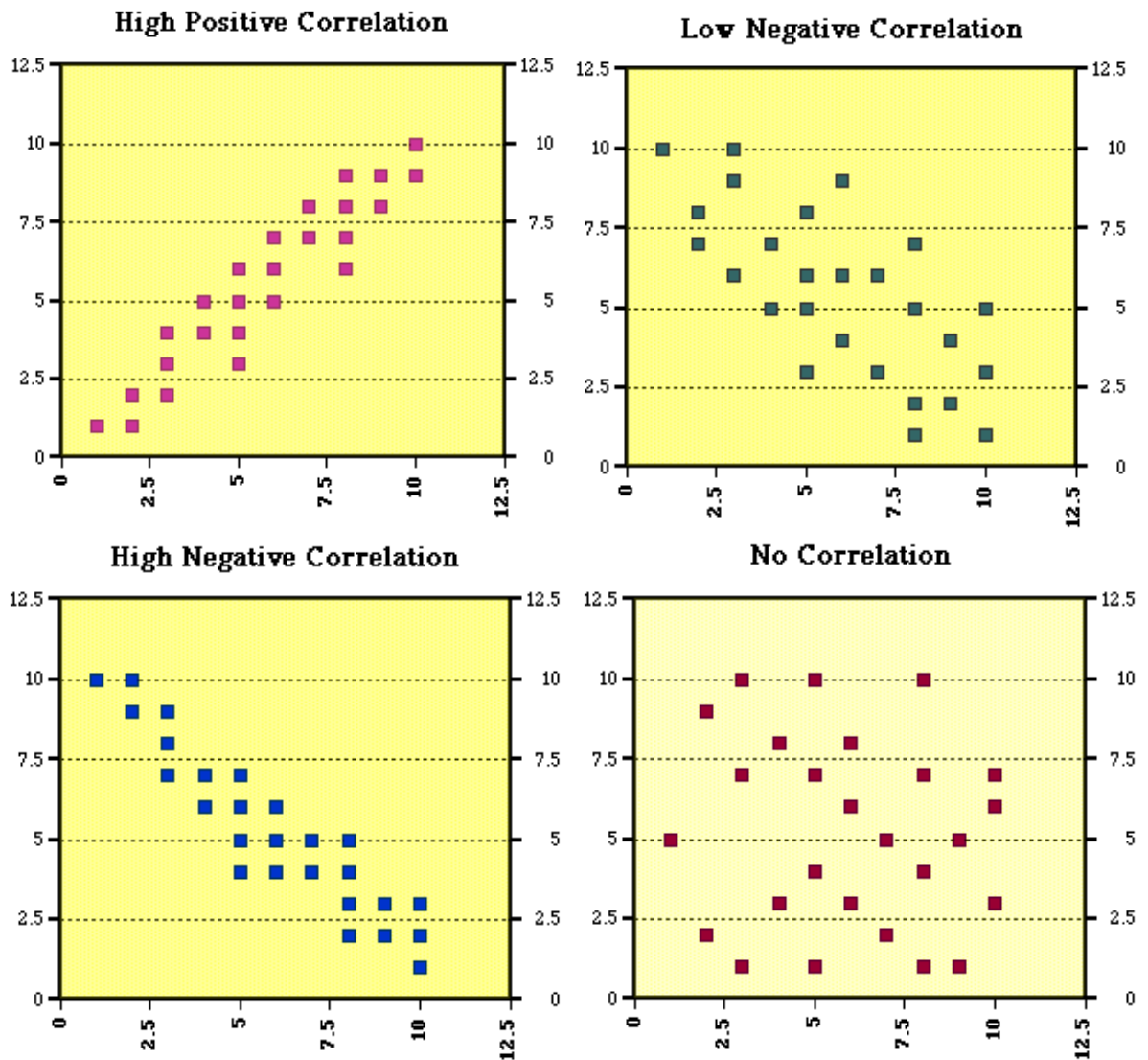


Рисунок 2.7 – Види кореляції скатерограмм

Нормальна форма скаттерограмми є еліпс, витягнутий уздовж бісектриси. Саме таке розташування еліпса означає, що до дихальної додана деяка величина недихальних аритмій. Форма скаттерограмми у вигляді кола означає відсутність недихальних компонентів аритмії. Вузкий овал (рис. 2.8) відповідає переважанню недихальних компонентів в загальній варіабельності ритму, яка визначається довжиною «хмари» (скаттерограмми).

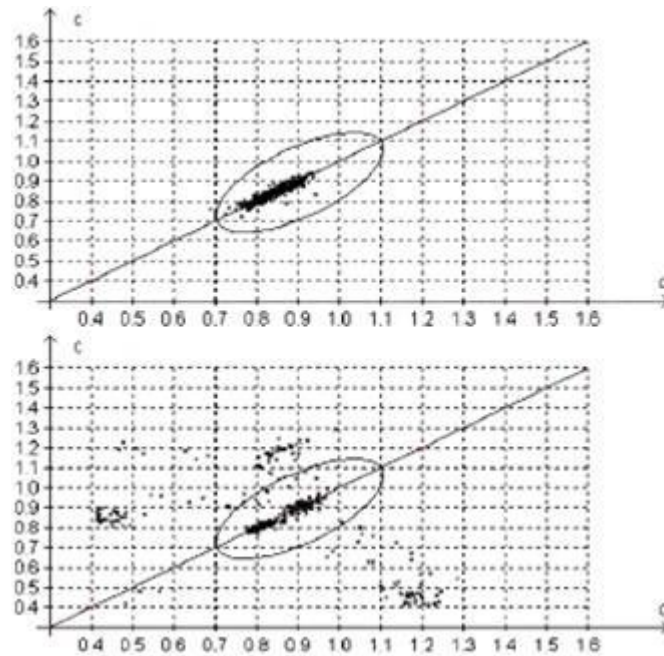


Рис. 2.8 – Зразки кореляційних ритмограмм (КРГ) - скатерограмм, вгорі - нормальна, внизу - у пацієнта з аритмією.

Довжина овалу добре корелювала з величиною HF, а ширина з LF. При аритміях, коли методи статистичного і спектрального аналізу варіабельності серцевого ритму малоінформативними або неприйнятні, доцільно використовувати оцінку кореляційної ритмограмм.

Зменшення розмірності даних.

У проблемах класифікації машинного навчання часто існує занадто багато факторів, на основі яких складається остаточна класифікація. Ці фактори в основному є змінними, які називаються ознаками. Чим більша кількість функцій, тим складніше візуалізувати навчальний набір, а потім працювати над ним. Іноді більшість цих особливостей є співвіднесеними, а значить, і зайвими. Тут вступають у дію алгоритми зменшення розмірності. Зменшення розмірності - це процес зменшення кількості розглянутих випадкових змінних шляхом отримання набору основних змінних. Його можна розділити на вибір функції та вилучення функцій.

Інтуїтивно зрозумілий приклад зменшення розмірності можна обговорити через просту проблему класифікації електронної пошти, де нам потрібно класифікувати, чи електронна пошта є спамом чи ні. Це може включати велику кількість функцій, наприклад, чи має електронний лист загальну назву, вміст електронної пошти, чи використовує шаблон електронної пошти шаблон тощо.

Однак деякі з цих функцій можуть перетинатися. В іншій умові проблема класифікації, яка покладається як на вологість, так і на кількість опадів, може обрушитися лише на одну основну особливість, оскільки обидві вищезгадані дуже корелюють. Тому ми можемо зменшити кількість особливостей у таких проблемах. Проблему 3-D класифікації може бути важко уявити, тоді як 2-D можна віднести до простого двовимірного простору, а 1-D завдання - до простої лінії. На рисунку 2.9 ілюструється ця концепція, коли тривимірний простір функцій розділений на два 1-D просторові простори, а пізніше, якщо буде виявлено їх кореляція, кількість особливостей може бути зменшена ще більше.

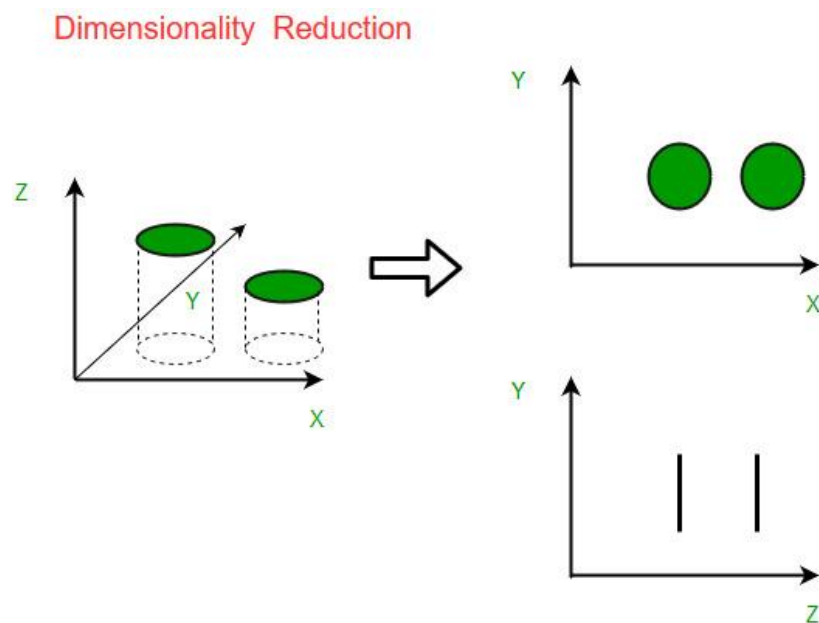


Рисунок 2.9 – Зменшення розмірності в різних проекціях

Для зменшення розмірності є два компоненти:

Вибір функцій. У цьому ми намагаємося знайти підмножину вихідного набору змінних чи функцій, щоб отримати меншу підмножину, яка може бути використана для моделювання проблеми. Зазвичай це включає три способи:

- фільтр
- обгортка
- вбудовування

Вилучення функцій. Це зменшує дані у просторі з високим розміром до простору нижчого розміру, тобто простору з меншим значенням розмірів.

Найпопулярніші методи для зменшення розмірності:

- метод головних компонент (Principal Component Analysis)
- лінійний розділювальний аналіз (Linear Discriminant Analysis)
- узагальнений дискримінантний аналіз (Generalized Discriminant Analysis)

Метод Головних Компонент

Основна лінійна методика зменшення розмірності, аналіз основного компонента, виконує лінійне відображення даних у нижньомірний простір таким чином, щоб дисперсія даних у низькомірному поданні була максимальною. На практиці будується коваріаційна (а іноді і кореляційна) матриця даних і обчислюються власні вектори на цій матриці. Власні вектори, що відповідають найбільшим власним значенням (основні компоненти), тепер можуть бути використані для реконструкції великої частки дисперсії вихідних даних. Більше того, перші кілька власних векторів часто можна трактувати з точки зору масштабної фізичної поведінки системи, оскільки вони часто вносять переважну більшість енергії системи, особливо в маломірних системах. Тим не менш, це потрібно доводити в кожному конкретному випадку, оскільки не всі системи проявляють таку поведінку. Первісний простір (з розмірністю кількості точок) був зменшений (із втратою даних, але, сподіваємось, зберігаючи найважливішу дисперсію) до простору, охопленого кількома власними векторами.

Лінійний Розділювальний Аналіз.

Лінійний дискримінантний аналіз (LDA) – це узагальнення лінійного дискримінанта Фішера, методу, який використовується у статистиці, розпізнаванні візерунків та машинному навчанні для пошуку лінійної комбінації ознак, яка характеризує або відокремлює два чи більше класи об'єктів чи подій.

Узагальнений Дискримінантний Аналіз

GDA займається нелінійним дискримінантним аналізом з використанням оператора функції ядра. Основа цієї теорії близька до машин, що підтримують вектор (SVM), наскільки метод GDA забезпечує відображення вхідних векторів у просторовий простір високих розмірів. Подібно до LDA, мета GDA полягає в тому, щоб знайти проєкцію ознак на простір нижнього розміру,

максимізувавши відношення між класовим розкидом до розсіювання всередині класу.

Переваги зменшення розмірності:

– це допомагає стиснути дані, а отже, зменшити простір для зберігання даних;

– це скорочує час на обчислення;

– це також допомагає видалити зайві функції, якщо такі є.

Недоліки зменшення розмірності:

– це може призвести до деякої кількості втрати даних;

– PCA прагне знайти лінійні кореляції між змінними, що іноді небажано;

– PCA виходить з ладу у випадках, коли значення та коваріація недостатньо для визначення наборів даних;

– ми можемо не знати, скільки основних компонентів потрібно зберігати на практиці, застосовуються деякі правила великого пальця.

Кластерний аналіз

Кластерний аналіз - це багатofакторний метод, який має на меті класифікувати вибірку предметів (або об'єктів) на основі набору вимірюваних змінних на ряд різних груп, таких що подібні предмети розміщуються в одній групі. Можемо навести приклад, де це може бути використано - вгалузі психіатрії, де характеристика пацієнтів на основі скупчення симптомів може бути корисною при виявленні відповідної форми терапії. У маркетингу може бути корисно визначити окремі групи потенційних клієнтів, щоб, наприклад, реклама могла бути належним чином націлена.

Існує ряд різних способів та методів, які можна використовувати для проведення кластерного аналізу.

Ієрархічні методи

Поділяються на :

– Агломераційні методи

– Методи поділу

Агломераційні методи, при яких суб'єкти починаються у своєму окремому кластері. Дві «найближчі» (найбільш подібні) кластери потім

поєднуються, і це робиться повторно, поки всі об'єкти не будуть в одному кластері. Наприкінці вибирається оптимальна кількість кластерів із усіх кластерних рішень.

Методи поділу, в яких всі суб'єкти починаються в одному кластері, і вищевказана стратегія застосовується у зворотному порядку, поки кожен суб'єкт не знаходиться в окремому кластері. Агломераційні методи використовуються частіше, ніж подільні методи, тому цей податок буде зосереджений на першому, а не на другому.

Неієрархічні методи (часто відомі як k-засоби кластеризації)

Дані, що використовуються при кластерному аналізі, можуть бути інтервальними, порядковими або категоричними. Однак наявність суміші різних типів змінної ускладнить аналіз. Це тому, що при кластерному аналізі потрібно мати певний спосіб вимірювання відстані між спостереженнями та типом використовуваного вимірювання, залежатиме від того, який тип даних у вас є. Для вимірювання "відстані" для двійкових та категоричних даних було запропоновано ряд різних заходів. Для інтервальних даних найпоширенішим вимірювачем відстані є евклідова відстань.

Евклідова відстань

Загалом, якщо у вас є p змінні X_1, X_2, \dots, X_p , виміряна на вибірці з n суб'єктів, спостережувані дані для предмета i можуть бути позначені через $x_{i1}, x_{i2}, \dots, x_{ip}$ та спостережувані дані для предмета j від $x_{j1}, x_{j2}, \dots, x_{jp}$. Евклідова відстань між цими двома предметами визначається через

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2}$$

При використанні такої міри, як евклідова відстань, масштаб вимірювання розглянутих змінних є проблемою, оскільки зміна шкали, очевидно, вплине на відстань між предметами (наприклад, різниця в 10 см може бути різницею 100 мм). Крім того, якщо одна змінна має набагато ширший діапазон, ніж інші, ця змінна, як правило, домінуватиме. Наприклад, якби вимірювання тіла проводилися для декількох різних людей, діапазон (у мм) висоти був би набагато ширшим, ніж діапазон окружності зап'ястя,

скажімо. Щоб подолати цю проблему, кожна змінна може бути стандартизована (перетворена на z-бали). Однак це саме по собі представляє проблему, оскільки вона має тенденцію до зменшення мінливості (відстані) між кластерами. Це відбувається тому, що якщо певна змінна добре розділяє спостереження, то, за визначенням, вона буде мати велику дисперсію (оскільки мінливість між кластером буде великою). Якщо ця змінна стандартизована, то розділення між кластерами стане менше. Незважаючи на цю проблему, багато підручників рекомендують стандартизувати. Якщо ви сумніваєтесь, однією стратегією було б провести аналіз кластеру двічі - один раз без стандартизації та один раз - щоб побачити, яка різниця, якщо така є, це впливає на отримані кластери.

В рамках цього підходу до кластерного аналізу існує ряд різних методів, що використовуються для визначення того, які кластери повинні бути об'єднані на кожному етапі. Основні методи узагальнені нижче.

Метод найближчого сусіда (метод одного зв'язку). У цьому способі відстань між двома кластерами визначається як відстань між двома найближчими членами або сусідами. Цей метод порівняно простий, але його часто піддають критиці, оскільки він не враховує структуру кластерів і може спричинити за собою проблему, яку називають ланцюжком, завдяки чому кластери закінчуються довгими і невпинними. Однак це краще, ніж інші методи, коли природні скупчення не мають сферичної чи еліптичної форми.

Метод найвіддаленішого сусіда(метод повного зв'язку). У цьому випадку відстань між двома кластерами визначається як максимальна відстань між членами - тобто відстань між двома предметами, які віддалені один від одного. Цей метод має тенденцію створювати компактні кластери подібного розміру, але, як для методу найближчого сусіда, не враховує структуру кластера. Він також досить чутливий до людей, що вижили.

Метод середнього (між групами) зв'язку. Відстань між двома кластерами обчислюється як середня відстань між усіма парами об'єктів у двох кластерах. Це вважається досить надійним методом.

Центроїдний метод. Тут обчислюється центроїд (середнє значення для кожної змінної) кожного кластеру та використовується відстань між центроїдами. Кластери, центроїди яких найближче один до одного, об'єднуються. Цей метод також досить надійний.

Метод Уорда. У цьому методі об'єднуються всі можливі пари кластерів і обчислюється сума квадратних відстаней у кожному кластері. Потім це підсумовується для всіх кластерів. Вибирається комбінація, яка дає найменшу суму квадратів. Цей метод має тенденцію до створення кластерів приблизно однакового розміру, що не завжди бажано. Він також досить чутливий до людей, що вижили. Незважаючи на це, це один із найпопулярніших методів, поряд із методом середнього зв'язку.

Неієрархічні або k-середні методи кластеризації.

У цих методах заздалегідь задається бажана кількість кластерів і вибирається найкраще рішення. Кроки такого методу такі:

- вибір початкових центрів кластерів (по суті це сукупність спостережень, які знаходяться далеко один від одного - кожен суб'єкт утворює кластер одного, а його центр - значення змінних для цього предмета);

- призначення кожного суб'єкта до його найближчого кластеру, визначеного у відстані до центру;

- знайдення центроїдів скупчених скупчень;

- повторне обчислення відстань від кожного суб'єкта до кожного центроїда та перемістити спостереження, які не знаходяться в кластері, який їм найближчий;

- продовження, до поки центроїди не залишаться відносно стабільними.

Неієрархічний кластерний аналіз, як правило, застосовується при залученні великих наборів даних. Іноді це є переважним, оскільки дозволяє суб'єктам переходити з одного кластеру в інший (це неможливо в ієрархічному аналізі кластерів, коли суб'єкт, щойно призначений, не може переміститися до іншого кластеру). Два недоліки неієрархічного кластерного аналізу: (1) часто важко дізнатися, скільки кластерів ви, ймовірно, маєте, і тому аналіз може мати

повторюється кілька разів і (2) це може бути дуже чутливим до вибору початкових центрів кластерів.

2.4 Розробка методу розв'язання задачі

Завдання що вирішуються, полягають у виявленні аномалій роботи серця за допомогою перетворення ритмів електрокардіограм у скатерограми, та їх інтелектуального аналізу для виявлення зони ризику захворювання пацієнта на основі якого лікар зможе зробити передбачення щодо захворювання.

Для розрахунку скатерограми були використані дані по варіабельності серцевого ритму.

Варіабельність серцевого ритму (BCP) - засіб вимірювання змін в серцевому ритмі. Зазвичай вона розраховується шляхом аналізу тривалості інтервалів між серцевими скороченнями, ґрунтуючись на ЕКГ або кривих артеріального тиску. Варіабельність серцевого ритму є показником різниці в тимчасових інтервалах між двома послідовними серцевими скороченнями, так званих кардіоінтервалів (інтервали R-R) і вимірюваних в мілісекундах (мс). Кардіоінтервали виходять від сигналу ЕКГ, як показано на наведеній нижче ілюстрації (рис. 2.10).

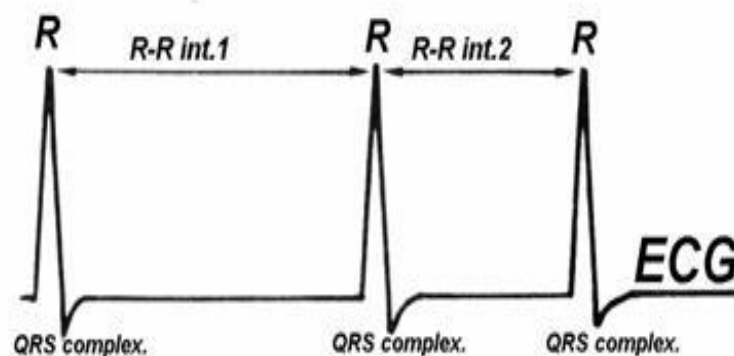


Рисунок 2.10 – RR-інтервали

Таким чином, для розрахунку скатерограмми використовувався набір RR-інтервалів, відфільтрований за типом комплексів.

Інтервал R-R – проміжок часу між сусідніми зубцями R електрокардіограми, рівний тривалості серцевого циклу; використовується при визначенні частоти серцевих скорочень, в діагностиці аритмій. З набору

виключені комплекси неясною етимологією. Потім був обчислений набір відносних значень RR-інтервалів за формулою:

$$RR_r = RR_n / RR_{n+1}$$

Потім була розрахована скатерограма в якій вісь X це представлена інтервалами значень від 0.2 до 1.9 RR_m розбитих на 60 рівних діапазонів, а вісь Y це також інтервали значень від 0.2 до 1.9 RR_{m+1} аналогічно розбиті на 60 рівних діапазонів.

Отримані значення були переведені в логарифмічну шкалу і нормовані на значення 100 для виділення негативними значеннями малого представництва в діапазоні.

Обчислена за даною методикою скатерограма була візуалізована за допомогою вбудованих засобів бібліотеки Seaborn. З її допомогою дуже просто створювати специфічну візуалізацію: такі як теплові карти.

Метод головних компонент

Для попередньої підготовки набору даних використовувався один з основних способів зменшити розмірність даних, втративши найменшу кількість даних, а саме метод головних компонент.

Метод головних компонент - це технологія багатовимірного статистичного аналізу, яка використовується для скорочення розмірності простору ознак з мінімальною втратою корисної інформації. Запропоновано К.Пірсоном в 1901 р, а потім детально розроблений американським економістом і статистиком Г.Хоттелінгом.

З математичної точки зору метод головних компонент являє собою ортогональне лінійне перетворення, яке відображає дані з вихідного простір ознак в новий простір меншої розмірності.

При цьому перша вісь нової системи координат будується таким чином, щоб дисперсія даних уздовж неї була б максимальна. Друга вісь будується ортогонально першій так, щоб дисперсія даних уздовж неї, була б максимальною з решти можливих і т.д. Перша вісь називається першою головною компонентою, друга - другою і т.д.

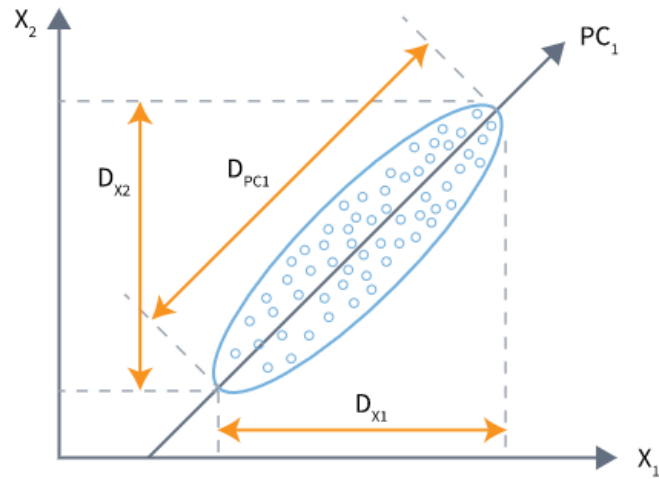


Рисунок 2.11 – Зниження розмірності вихідного 2-мірного простору

На рис. 2.11 показано зниження розмірності вихідного 2-мірного простору (X_1 , X_2) за допомогою методу головних компонент до 1-мірного.

Перша головна компонента PC_1 орієнтована вздовж напрямку найбільшої витягнутості еліпсоїда розсіювання точок об'єктів вихідного набору даних в просторі ознак, тобто з нею пов'язана найбільша дисперсія.

На малюнку, також, неважко помітити, що проекція дисперсії даних на вісь першої головної компоненти D_{PC1} , більше, ніж її проекції на вихідні осі DX_1 і DX_2 , але менше їх суми. Тобто за допомогою першої головної компоненти висловити всю дисперсію даних не вдалося. Тому будують другу, третю і т.д. головні компоненти, поки вони сумарно не відіб'ється всю дисперсію.

Алгоритм методу головних компонент:

- обчислюється матриця стандартизованих ознак Z ;
- обчислюється кореляційна матриця стандартизованих ознак R_Z ;
- обчислюється матриця власних значень;

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

та матриця нормованих власних векторів $U = \begin{pmatrix} \bar{u}_1 & \bar{u}_2 & \dots & \bar{u}_n \end{pmatrix}$.

- обчислюємо матрицю факторних навантажень;

$$W = U\Lambda^{1/2}$$

д) за матрицею власних значень обчислюється частка дисперсії (власне значення це і є дисперсія);

$$v_k = \frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^n \lambda_j} \geq \delta$$

е) обчислюється матриця значень факторів;

$$F = ZW\Lambda^{-1}$$

ж) за початковими Y та одержаною F будемо регресійну лінійну модель.

Склад компоненти визначається за коефіцієнтом інформованості:

$$k_i = \frac{\sum_{j=1}^{n_i} w_{ji}}{\sum_{j=1}^n w_{ji}}$$

Набір пояснюючих ознак вважається задовільним, якщо величина коефіцієнта інформованості становить не менше 0,70.

Таким чином, сенс методу полягає в тому, що з кожної головної компонентою пов'язана певна частка загальної дисперсії вихідного набору даних (її називають навантаженням). У свою чергу, дисперсія, що є мірою мінливості даних, може відображати рівень їх інформативності. Основними обмеженнями методу головних компонент є:

- неможливість смислової інтерпретації компонент, оскільки вони "вбирають" в себе дисперсію від декількох вихідних змінних;
- метод може працювати тільки з безперервними даними.

Метод головних компонент включається до складу більшості аналітичних платформ і широко використовується для зниження розмірності вхідних даних на етапі їх попередньої обробки.

Методика визначення подібних груп

Для того щоб класифікувати отримані скатерограми була застосована методика кластеризації, тобто класифікація без вчителя.

В якості базової методики був застосований найбільш популярний метод кластеризації, а саме метод К-середніх, винайдений в 1950-х роках математиком Гуго Штейнгаузом.

Дія алгоритму таке, що він прагне мінімізувати сумарне квадратичне відхилення точок кластерів від центрів цих кластерів:

$$V = \sum_{i=1}^k \sum_{x \in S_i} (x - \mu_i)^2$$

Де k - число кластерів, S_i - отримані кластери, μ_i - центри мас всіх векторів X з кластера.

За аналогією з методом головних компонент центри кластерів називаються також головними точками, а сам метод називається методом головних точок і включається в загальну теорію головних об'єктів, що забезпечують найкращу апроксимацію даних. Алгоритм є версією EM-алгоритму, що застосовується також для поділу суміші Гауссіан. Він розбиває безліч елементів векторного простору на заздалегідь відоме число кластерів k .

Основна ідея полягає в тому, що на кожній ітерації переобчислюють центр мас для кожного кластера, отриманого на попередньому кроці, потім вектори розбиваються на кластери знову відповідно до того, який з нових центрів виявився ближчим за обраною метриці.

Алгоритм завершується, коли на якийсь ітерації не відбувається зміни внутрікластерної відстані. Це відбувається за кінцеве число ітерацій, так як

кількість можливих розбиттів кінцевої безлічі кінцеве, а на кожному кроці сумарна квадратичне відхилення V зменшується, тому зациклення неможливо.

Методика визначення найліпшої кількості груп

Для визначення найліпшої кількості груп пацієнтів по зонам ризику була використана метрика Silhouette.

Silhouette дає інформативне зображення того, наскільки добре класифікований кожен об'єкт. Значення Silhouette для кожного об'єкту - це міра того, наскільки він схожий або згуртований в власному кластері у порівнянні з іншими кластерами. Silhouette для об'єкту коливається від -1 до +1, де високе значення вказує на те, що об'єкт добре підходить до власного кластеру і погано відповідає сусіднім кластерам. Якщо більшість об'єктів мають високе значення, то конфігурація кластеризації визнається найкращою. Якщо багато точок мають низьке або негативне значення, то в конфігурації кластеризації може бути занадто багато або занадто мало кластерів. У даній роботі Silhouette була обчислена за допомогою Евклідової метрики відстані.

Дефініція метрики Silhouette

Припустимо, дані були класифіковані за допомогою будь-якої техніки, наприклад k-means, у k кластери. Для кожної точки даних $i \in C_i$ (точка даних i в кластері C_i), нехай

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i, j)$$

Де - $a(i)$ - середня відстань між i та всіма іншими точками даних у тому ж кластері,

$d(i, j)$ - відстань між точками даних i, j в кластері C_i .

$a(i)$ можемо інтерпретувати як міру того, наскільки добре точка i віднесена до свого кластеру (чим менше значення, тим краще).

Потім визначаємо середню несхожість точки i на деякий кластер C як середню відстань від i до всіх точок C (де $C \neq C_i$).

Для кожної точки даних де $i \in C_i$, визначимо

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j)$$

найменшу середню відстань i до всіх точок будь-якого іншого кластеру. Кластер з цією найменшою середньою різницею вважається "сусіднім кластером" точці i , оскільки це наступний кластер, що найкраще підходить для цієї точки.

Тепер визначимо значення Silhouette для однієї точки даних i

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \text{ if } |C_i| > 1$$

та

$$s(i) = 0, \text{ if } |C_i| = 1$$

Що також можна записати як:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases}$$

Використання метрики

Для визначення найліпшої кількості груп був розрахований інтегральний показник де 0 відповідає найгіршій конфігурації, а 1 відповідає найліпшій[9].

2.5 Розробка алгоритму розв'язання

Алгоритм розв'язання задачі проходить у такі етапи:

а) з набору записів електричної активності серця виділяються дані по серцевим ритмам, як показником, найбільш відображаючим загальний стан здоров'я людини;

б) будуються набори скатерограмм в логнормальній шкалі для ефективної візуалізації режимів роботи серця;

в) згортка даних проводиться методикою головних компонент та виділяються перші дві головні компоненти для подальшої кластеризації;

г) методом кластеризації K-means дані згуртовуються та розподіляються на 7 кластерів (груп).

Алгоритм K-means:

Етап 1. Початковий розподіл об'єктів по кластерам

а) вибір випадковим чином k точок даних з D як початкова множина представників кластера C ;

б) розподіл об'єктів по кластерам у відповідність з формулою.

Етап 2. Перерозподіл серединних елементів:

а) обчислення центру для кожного кластера;

б) перерозподіл об'єктів по кластерам.

2.6 Результати досліджень ефективності методу

Алгоритм k-means є досить простим та повторюваним алгоритмом кластеризації даних, котрий розподіляє певний набір даних на задане користувачем число груп(кластерів), k . Алгоритм простий у використанні та для реалізації і запуску, порівняно швидкий, легко адаптується і дуже поширений на практиці. Це історично один з найвідоміших алгоритмів інтелектуального аналізу даних.

Потенційна проблема алгоритму - проблема «порожніх» кластерів. При запуску k-means, особливо з великим значенням k і / або коли дані знаходяться у величезному розмірному просторі, може статись так, що в якийсь момент виконання, може існувати представник кластера c_j , такий, що всі крапки x_i в D ближче до деякого іншому представнику кластера, який не є c_j . Коли точки в D будуть присвоєні до найближчого кластеру, j -му кластеру будуть присвоєні нульові точки.

Алгоритм k-means простий ітеративний алгоритм кластеризації даних, що розділяє купу даних на визначену кількість k кластерів. За своєю суттю, алгоритм працює за методом перебору в два етапи: (1) проходить кластеризація усіх точок даних в залежності від відстані між точкою та її найближчим представником кластера та (2) проводиться переоцінка представників кластеру.

Обмеження алгоритму k-means включає чутливість k-means до ініціалізації і визначення значення k.

Незважаючи на всі недоліки, k-means залишається найбільш широко розповсюдженим алгоритмом кластеризації даних на практиці. Алгоритм простий, зрозумілий і досить масштабований і може бути легко модифікований для вирішення різних конкретних завдань, таких як часткове навчання з учителем або потокових даних. Стабільні поліпшення і узагальнення основних алгоритмів забезпечили її актуальність і поступово збільшують його ефективність [21].

В результаті проведених експериментів було визначено найбільшу ефективність саме цього методу, який і був використаний у дослідженні. По результатам обчислення метрики Silhouette(рис. 2.12) було обрано 7 груп як найкраща конфігурація. У середині кожної з яких були локалізовані візуально подібні скатерограми.

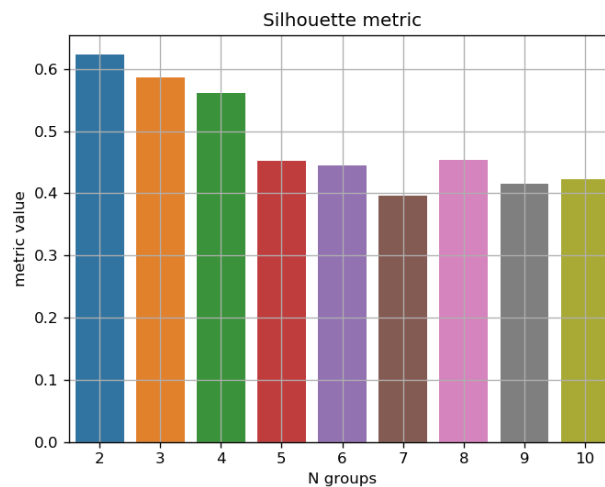


Рисунок 2.12 – Результати дослідження проведені метрикою Silhouette

Висновки до розділу

У даному розділі було описано змістовну та математичну постановку задачі інтелектуального аналізу аномалій ритмів електрокардіограм. Були розглянуті методи вирішення завдання, оцінка якостей кожного, та вибір оптимального для використання в даному дослідженні. Був розроблений та представлений алгоритм розв'язання та показані результати дослідження.

3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ

3.1 Засоби розробки

Python - інтерпретована мова програмування високого рівня, загального призначення. Створена Гідо ван Россумом та вперше випущена в 1991 році, філософія дизайну Python підкреслює читабельність коду завдяки помітному використанню значного пробілу. Його мовні конструкції та об'єктно-орієнтований підхід мають на меті допомогти програмістам написати чіткий логічний код для малих та масштабних проектів. Він підтримує декілька парадигм програмування, включаючи процедурне, об'єктно-орієнтоване та функціональне програмування. Python часто описується мовою "включені батареї" завдяки своїй всебічній бібліотеці стандартів.

Python – це скриптова мова загального призначення, для якої був розроблений спеціальний інтерфейс, призначений для взаємодії з внутрішніми функціями Blender'a. Більшість функцій не залежать від Python. Єдиний виняток це меню Help яке відкриває зовнішні посилання в веб-браузер.

Python портований та працює майже на всіх нині відомих платформах — від КПК до мейнфреймів.

Серед основних переваг Python можна назвати такі:

- чистота синтаксису (для виділення блоків слід використовувати відступи);
- переносимість програм;
- дистрибутив має велику кількість корисних модулів;
- можливість використання Python в діалоговому режимі (дуже корисне для розв'язання простих задач);
- стандартний дистрибутив має просте, але досить потужне середовище розробки, яке зветься IDLE та яке написане на мові Python;
- зручний для розв'язання математичних проблем.

Пайтон має ефективну структуру даних високого рівня та простий, але водночас із тим, ефективний підхід до об'єктно-орієнтованого програмування. Витончений синтаксис Пайтона, динамічна обробка, роблять його ідеальним для написання скриптів та швидкісної розробки прикладних програм у багатьох різноманітних галузях та на більшості з існуючих платформ.

Також інтерпретатор пайтону може бути розширений функціями та типами даних, розробленими на С або С++. Пайтон також зручний як мова розширення для прикладних програм, які потребують подальшого налагодження.

Як базу даних в роботі була використана SQLite 3 - це автономний, що працює без сервера транзакційний механізм бази даних SQL. Python отримав модуль sqlite3 в версії 2.5, що означає що можна створювати базу даних SQLite в будь-якої справжньої версії Python, без необхідності завантаження додаткових інструментів.

Окремо також слід звернути увагу на фреймворк Django

Django - це програмний «каркас» з багатими можливостями, що підходить для розробки складних сайтів і веб-додатків, написаний на мові програмування Python. «Каркас» - це не що інше, як сукупність модулів, що полегшують розробку. Вони згруповані разом і дозволяють створювати додатки чи веб-сайти з наявного джерела, а не з нуля.

Основна мета Django - полегшити створення складних веб-сайтів, керованих базами даних. Фреймворк наголошує на повторному використанні та «підключеності» компонентів, меншій кількості коду, низькому з'єднанні, швидкому розвитку та принципі не повторювати себе. Python використовується на всьому протязі, навіть для файлів налаштувань та моделей даних. Django також надає додатковий адміністративний інтерфейс для створення, читання, оновлення та видалення, який динамічно генерується за допомогою самоаналізу та налаштовується за допомогою адміністраторських моделей.

Django проектувався для роботи під управлінням Apache (з модулем mod_python) і з використанням PostgreSQL в якості бази даних. В даний час, крім PostgreSQL, Django може працювати з іншими СУБД: MySQL (MariaDB),

SQLite, Microsoft SQL Server, DB2, Firebird, SQL Anywhere і Oracle. Для роботи з базою даних Django використовує власний ORM, в якому модель даних описується класами Python, і по ній генерується схема бази даних.

Архітектура Django схожа на «Модель-Представлення-Контролер» (MVC). Контролер класичної моделі MVC приблизно відповідає рівню, який в Django називається Представлення (View), а презентаційна логіка Уявлення реалізується в Django рівнем Шаблонів (Templates). Через це рівневу архітектуру Django часто називають «Модель-Шаблон-Подання» (MTV).

Спочатку розробка Django велася для забезпечення більш зручної роботи з новинними ресурсами, що досить сильно відбилося на архітектурі: фреймворк надає ряд засобів, які допомагають у швидкій розробці веб-сайтів інформаційного характеру. Наприклад, розробнику не потрібно створювати контролери та сторінки для адміністративної частини сайту, в Django є вбудований додаток для керування вмістом, яке можна включити в будь-який сайт, зроблений на Django, і яке може управляти відразу декількома сайтами на одному сервері. Адміністративне додаток дозволяє створювати, змінювати і видаляти будь-які об'єкти наповнення сайту, протоколюючи всі скоєні дії, і надає інтерфейс для управління користувачами і групами (з пооб'єктного призначенням прав).

Веб-фреймворк Django використовується в таких великих і відомих сайтах, як Instagram, Disqus, Mozilla, The Washington Times, Pinterest, lamoda і ін.

Деякі можливості Django:

- ORM, API доступу до БД з підтримкою транзакцій;
- вбудований інтерфейс адміністратора, з уже наявними перекладами багатьма мовами;
- диспетчер URL на основі регулярних виразів;
- розширювана система шаблонів з тегами і спадкуванням;
- система кешування;
- інтернаціоналізація;

- підключається архітектура додатків, які можна встановлювати на будь-які Django-сайти;
- «Generic views» - шаблони функцій контролерів;
- авторизація та аутентифікація, підключення зовнішніх модулів аутентифікації: LDAP, OpenID та інші;
- система фільтрів («middleware») для побудови додаткових обробників запитів, як наприклад включені в дистрибутив фільтри для кешування, стиснення, нормалізації URL і підтримки анонімних сесій;
- бібліотека для роботи з формами (успадкування, побудова форм по існуючій моделі БД);
- вбудована автоматична документація по тегам шаблонів і моделей даних, доступна через адміністративне додаток [23], [24].

Також важливо приділити увагу найбільш вживаним бібліотекам які були використані – Scikit-learn та Matplotlib.

Scikit-learn (раніше scikits.learn і також відомий як sklearn) - це безкоштовна бібліотека машинного навчання для мови програмування Python. Він містить різні алгоритми класифікації, регресії та кластеризації, включаючи підтримуючі векторні машини, випадкові ліси, збільшення градієнта, k-засоби та DBSCAN, і призначений для взаємодії з числовими та науковими бібліотеками Python NumPy та SciPy.

Matplotlib — це бібліотека побудови графіків Python 2D, яка створює якісні публікації у різноманітних форматах у 2D форматі та інтерактивному середовищі на різних платформах. Matplotlib може використовуватися в сценаріях Python, оболонках Python та IPython, ноутбуках Jupyter, на серверах веб-додатків та чотирьох графічних наборах інструментів користувальницького інтерфейсу.

Matplotlib намагається зробити легкі речі легкими і важкими. Ви можете генерувати графіки, гістограми, спектри потужності, смугові діаграми, діаграми помилок, розсіювачі тощо, лише за допомогою декількох рядків коду.

Для простого побудови модуля pyplot передбачений інтерфейс, подібний до MATLAB, особливо в поєднанні з IPython. Для енергокористувача ви маєте

повний контроль над стилями рядків, властивостями шрифту, властивостями осей тощо, через об'єктно-орієнтований інтерфейс або через набір функцій, знайомих користувачам MATLAB.

3.2 Вимоги до технічної частини

Дана інформаційна система складається саме з обчислюваного сервера та мережі під'єднаних до нього клієнтів.

Програмне забезпечення серверу не має прив'язки до конкретного технічного забезпечення, тому може бути розгорнуто будь де, від хмарного кластера до звичайного фізичного серверу.

Для правильної роботи програмного забезпечення серверу необхідно мати наступний мінімальний склад технічних засобів:

- а) сервер з наступною конфігурацією:
 - мінімальна тактова частота процесору не нижче за 2 ГГц;
 - не менш ніж 4 Гб об'єм оперативної пам'яті;
 - HDD або SSD об'ємом не менше 10 Гб.
- б) додатково має бути встановлене таке програмне забезпечення:
 - операційна система: будь-яка Linux базована ОС.

Апаратне забезпечення клієнтської частини повинне забезпечувати підтримку програмного забезпечення клієнтської частини.

3.3 Вимоги до програмного продукту

Наведемо функціональні вимоги до обох частин програмного забезпечення системи.

Функціональні вимоги до серверної частини програмного забезпечення програмно-інформаційної системи. Система повинна:

- надавати одночасний доступ для багатьох користувачів;
- приймати дані від клієнтського програмного забезпечення у форматі даних *.beat*;

- надавати інформацію про завантажений користувачем запис;
- можливість вивантаження звіту з інформацією про запис.

Функціональні вимоги до клієнтської частини програмного забезпечення програмно-інформаційної системи. Клієнтська частина повинна:

- мати доступ до всесвітньої мережі інтернет.

3.4 Архітектура програмного забезпечення

Програмне забезпечення побудоване за клієнт-серверної архітектури. Це виконано з ціллю централізованого зберігання даних, що надходять з територіально різних місць.

При обміні інформацією, зачинателем зв'язку є клієнт, що відправляє дані на сервер, а також має змогу завантажувати їх з серверу.

3.5 Інструкція користувача

Визначити рівень ризику серцево-судинних захворювань тепер можна за допомогою спеціально спроектованої і навченої на великій клінічній базі нейронної мережі.

Машинний інтелект ретельно перегляне запис, що містить майже сотню тисяч биття серця, записаних за допомогою переносного апарату для добового ЕКГ моніторингу і порівняє з наявною у нього клінічною базою.

Отриману загальну картину роботи серця нейронна мережа співвіднесе з картою, побудованою на вже закладеній клінічній базі та визначить до якої групи рівня ризику серцево-судинних захворювань відноситься переглянута запис.

Для початку роботи необхідно увімкнути живлення приладу та відкрийте ваш браузер і скопіюйте адресу сайту Smart Heart Checker

<http://heatmap.zobin.org:8080/>

На ній користувач зможе побачити головну сторінку інформаційної системи(рис. 3.1.) На ній відображено опис роботи системи та коротку інструкцією роботи з нею.

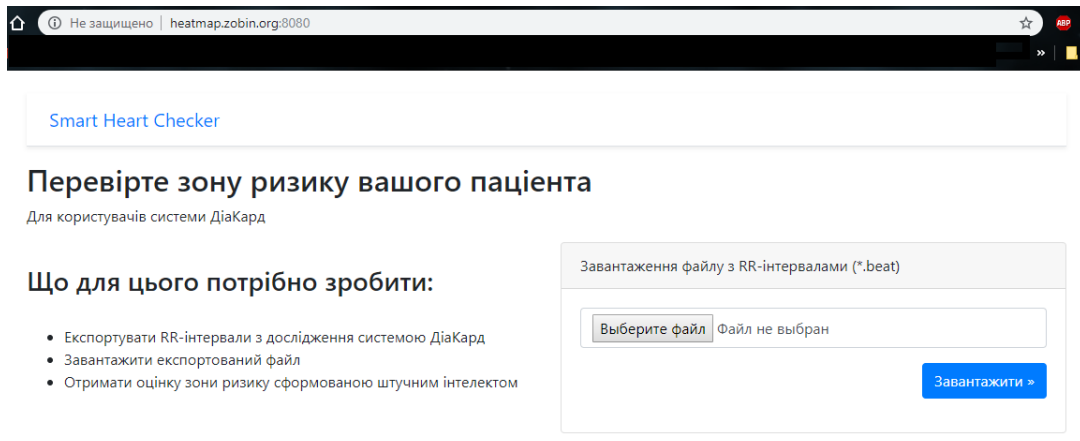


Рисунок 3.1 – Головна сторінка програми

На головній сторінці також є вікно для завантаження даних через яке користувач зможе почати роботу та вибрати файл (натиснути по кнопці "Виберіть файл") зі своєї файлової системи для завантаження у інформаційну систему(рис.3.2)

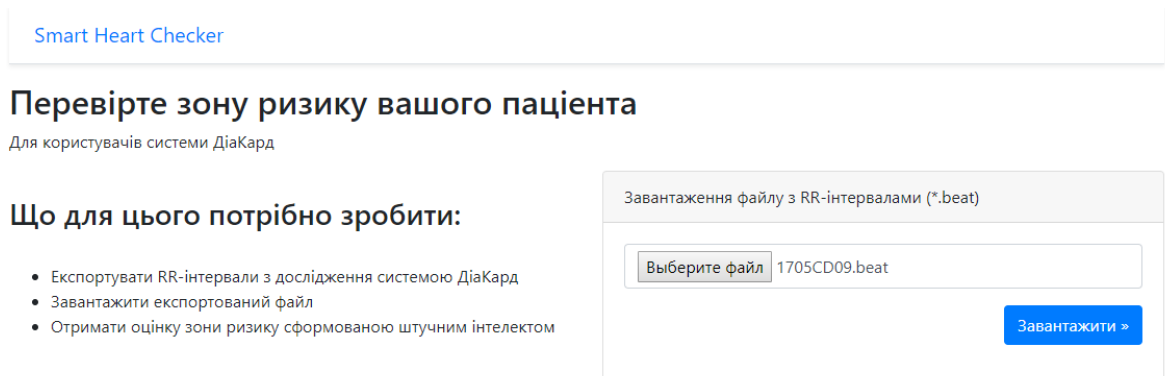


Рисунок 3.2 – Вікно завантаження даних

Після очікування поки система проводить аналіз отриманих даних користувача спрямує на наступну сторінку системи на якій буде відображено декілька зон з важливою інформацією. А саме :

- Мапа активності серця пацієнта
- Анотація до запису ЕКГ
- Карта зон ризику пацієнта
- Завантаження звіту

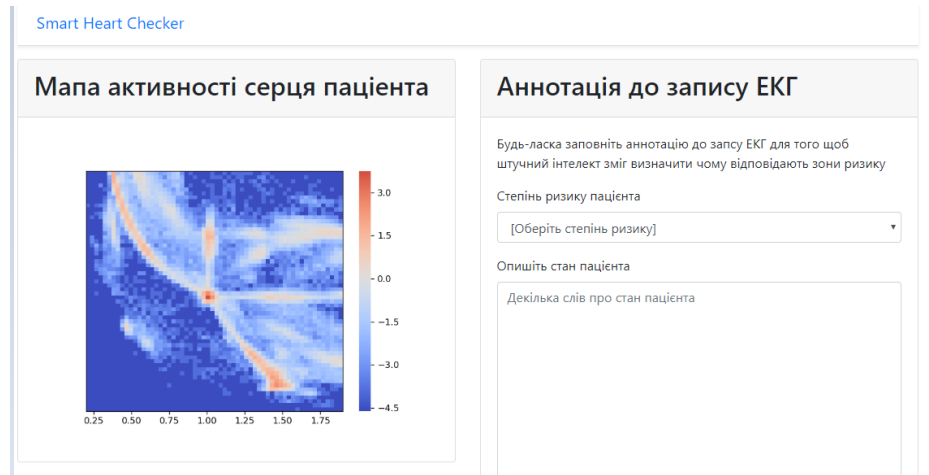


Рисунок 3.3 – Вікно активності

На першій зоні відображено мапу активності серця пацієнта, саме його запис електрокардіограми у вигляді скатерограмми сформованої системою на основі завантажених користувачем даних. (рис. 3.3.)

На другій зоні відображено вікно в якому користувач зможе анотувати даний запис електрокардіограми та визначити степінь ризику захворювання пацієнта. Ці данні будуть збережені на сервері для подальшого їх використання. (рис. 3.3.)

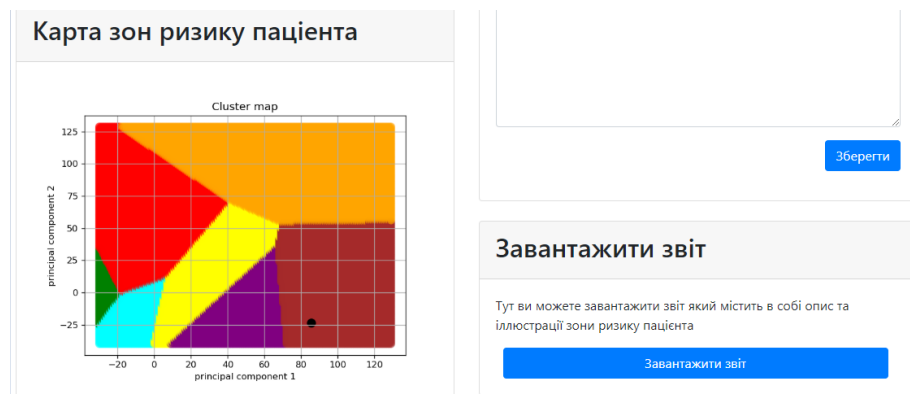


Рисунок 3.4 – Вікно завантаження даних

На третій зоні відображено карту зон ризику пацієнта на якій чорною точкою позначено місцезнаходження даного запису електрокардіограми на карті розподіленою за кольором на 7 різних кластерів відповідно до типології запису електрокардіограми визначеної системою. (рис. 3.4.)

На четвертій зоні відображено вікно завантаження звіту, через яке користувач зможе отримати та зберегти усю інформацію щодо запиту яка присутня на даний момент на сервері. (рис. 3.4.)

Рівні ризику серцево-судинних захворювань, що визначаються нейронною мережею:

- Дуже низький (область з зеленим кольором). Добовий запис показує відмінну роботу серця пацієнта, всім би так;
- Низький (область з блакитним кольором). Робота серця майже ідеальна, але щоб було без "майже" можливо варто трохи подбати про нього;
- Скоріше за низький (область з рожевим кольором). Хороша робота серця ось-ось опиниться під загрозою, подумайте про проходження лікування;
- Середній (область з бурим кольором). Правильна робота серця під загрозою, потрібне лікування;
- Скоріше за великий (область з світло-коричневим кольором). Виявлено помірний ризик смерті від серцево-судинного захворювання;
- Великий (область з жовтим кольором). Виявлено ризик смерті від серцево-судинного захворювання;
- Дуже великий (область з червоним кольором). Виявлено високий ризик смерті від серцево-судинного захворювання.

На рисунку 3.5 зображений схематичний розбір карти зон ризику пацієнта

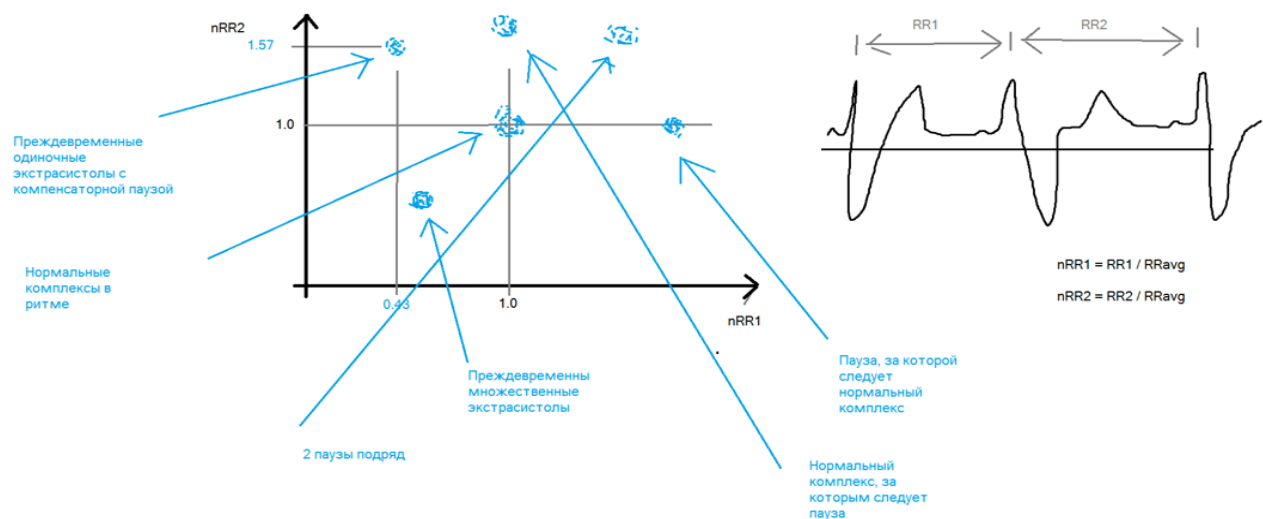
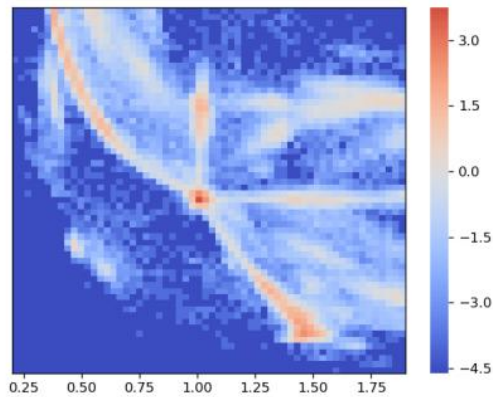


Рисунок 3.5 – Розбір карти зон ризику пацієнта

Звіт степені ризику пацієнта

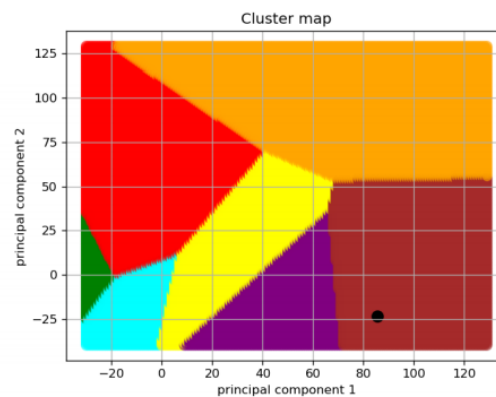
Мапа активності серця пацієнта



Карта зон ризику пацієнта

Рисунок 3.6 – Звіт про запис

Карта зон ризику пацієнта



Анотація до запису ЕКГ

Степінь ризику: Велика

Стан пацієнта:

test test test

Рисунок 3.7 – Звіт про запис

На рисунках 3.6, 3.7 Відображено звіт про запис електрокардіограми у вигляді pdf-файлу, який користувач зможе зберегти у свою файлову систему.

3.6 Випробовування програмного продукту

Результати дослідження та ефективності методу представлені на схемі у таб. 3.1 у якій відображена схема перетворення вхідних даних від їх отримання до кінцевого результату інтелектуального аналізу.

Таблиця 3.1 – Приклад проведеного дослідження

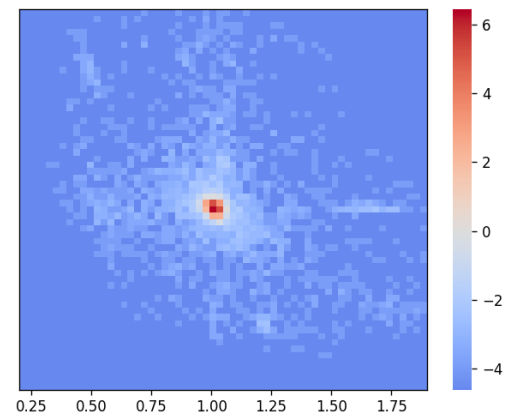
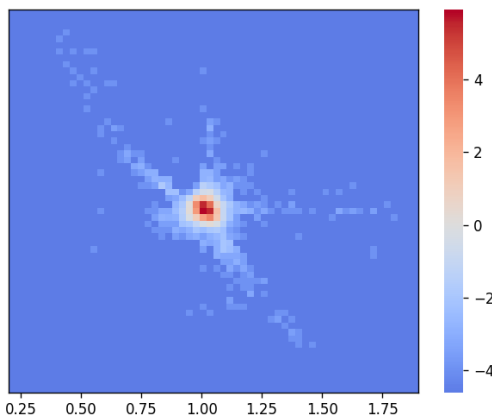
Пацієнт 1

| N | Комп | Время | R-R мс | QRS | QR мс | RS мс | скор. | амп.Т | сиг/шум |
|----|------|-------|--------|-----|-------|-------|-------|-------|---------|
| 1 | A | 87 | 680 | 8 | ---- | 1006 | 527 | -1025 | 173 |
| 2 | N | 207 | 938 | 86 | 1855 | 2354 | 1465 | --- | 52 |
| 3 | N | 299 | 719 | 94 | 1826 | 2324 | 1504 | 88 | 50 |
| 4 | N | 392 | 727 | 94 | 1836 | 2314 | 1436 | 117 | 52 |
| 5 | N | 484 | 719 | 94 | 1709 | 2188 | 1445 | 59 | 52 |
| 6 | N | 578 | 734 | 86 | 1748 | 2256 | 1484 | 137 | 53 |
| 7 | N | 670 | 719 | 86 | 1797 | 2236 | 1406 | 146 | 53 |
| 8 | N | 763 | 727 | 94 | 1895 | 2354 | 1523 | 117 | 52 |
| 9 | N | 856 | 727 | 86 | 1846 | 2324 | 1436 | 156 | 51 |
| 10 | N | 949 | 727 | 86 | 1777 | 2295 | 1406 | 127 | 50 |
| 11 | N | 1041 | 719 | 94 | 1748 | 2227 | 1484 | 127 | 52 |
| 12 | N | 1134 | 727 | 94 | 1768 | 2217 | 1406 | 137 | 54 |
| 13 | N | 1226 | 719 | 86 | 1650 | 2139 | 1367 | -98 | 49 |
| 14 | N | 1319 | 727 | 94 | 1895 | 2324 | 1484 | 107 | 50 |
| 15 | N | 1412 | 727 | 94 | 1816 | 2275 | 1426 | 166 | 50 |
| 16 | N | 1505 | 727 | 86 | 1748 | 2217 | 1406 | 146 | 53 |
| 17 | N | 1598 | 727 | 102 | 1807 | 2334 | 1494 | 78 | 51 |
| 18 | N | 1691 | 727 | 117 | 1875 | 2451 | 1533 | --- | 58 |
| 19 | N | 1784 | 727 | 86 | 1836 | 2451 | 1504 | 98 | 66 |
| 20 | N | 1877 | 727 | 86 | 1943 | 2383 | 1543 | 127 | 50 |

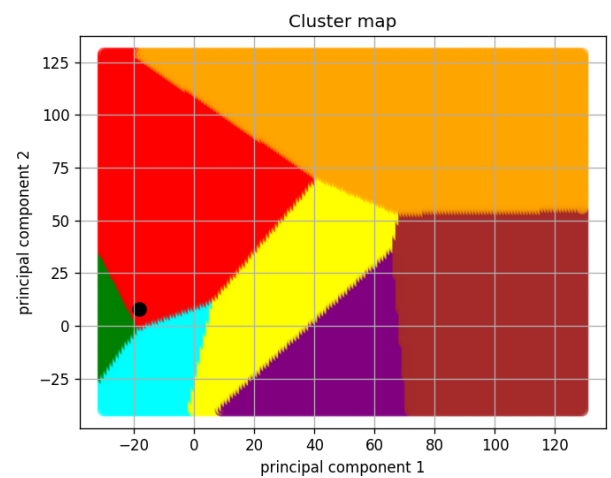
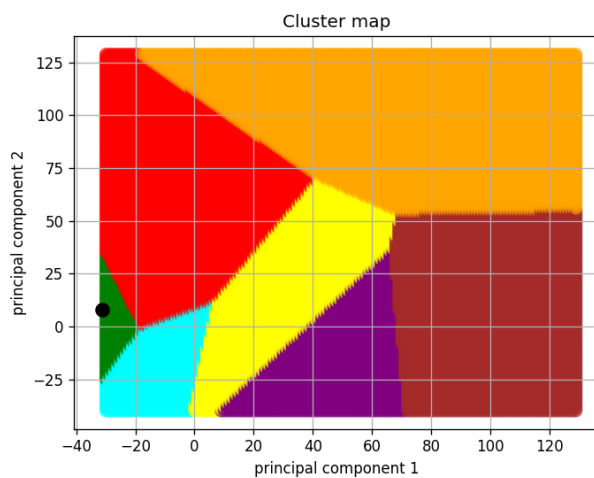
Пацієнт 2

| N | Комп | Время | R-R мс | QRS | QR мс | RS мс | скор. | амп.Т | сиг/шум |
|----|------|-------|--------|-----|-------|-------|-------|-------|---------|
| 1 | A | 87 | 680 | 8 | ---- | 1006 | 527 | -1025 | 173 |
| 2 | N | 207 | 938 | 86 | 1855 | 2354 | 1465 | --- | 52 |
| 3 | N | 299 | 719 | 94 | 1826 | 2324 | 1504 | 88 | 50 |
| 4 | N | 392 | 727 | 94 | 1836 | 2314 | 1436 | 117 | 52 |
| 5 | N | 484 | 719 | 94 | 1709 | 2188 | 1445 | 59 | 52 |
| 6 | N | 578 | 734 | 86 | 1748 | 2256 | 1484 | 137 | 53 |
| 7 | N | 670 | 719 | 86 | 1797 | 2236 | 1406 | 146 | 53 |
| 8 | N | 763 | 727 | 94 | 1895 | 2354 | 1523 | 117 | 52 |
| 9 | N | 856 | 727 | 86 | 1846 | 2324 | 1436 | 156 | 51 |
| 10 | N | 949 | 727 | 86 | 1777 | 2295 | 1406 | 127 | 50 |
| 11 | N | 1041 | 719 | 94 | 1748 | 2227 | 1484 | 127 | 52 |
| 12 | N | 1134 | 727 | 94 | 1768 | 2217 | 1406 | 137 | 54 |
| 13 | N | 1226 | 719 | 86 | 1650 | 2139 | 1367 | -98 | 49 |
| 14 | N | 1319 | 727 | 94 | 1895 | 2324 | 1484 | 107 | 50 |
| 15 | N | 1412 | 727 | 94 | 1816 | 2275 | 1426 | 166 | 50 |
| 16 | N | 1505 | 727 | 86 | 1748 | 2217 | 1406 | 146 | 53 |
| 17 | N | 1598 | 727 | 102 | 1807 | 2334 | 1494 | 78 | 51 |
| 18 | N | 1691 | 727 | 117 | 1875 | 2451 | 1533 | --- | 58 |
| 19 | N | 1784 | 727 | 86 | 1836 | 2451 | 1504 | 98 | 66 |
| 20 | N | 1877 | 727 | 86 | 1943 | 2383 | 1543 | 127 | 50 |

Рисунки 3.8, 3.9 – Зразки перших 20-ти рядків даних ЕКГ пацієнта отриманих з системи ДіаКарт



Рисунки 3.10, 3.11 – Зразки теплової карти скатерограми RR-інтервалів для логнормалізованих даних



Рисунки 3.12, 3.13 – Зразки положення точки стану пацієнта на карті зон ризику у просторі перших двох головних компонент

Висновки до розділу

В даному розділі були розглянуті засоби які були використані для розробки системи. На основі цих засобів були сформульовані вимоги до технічного та програмного забезпечення. Була описана система загалом, її компоненти та архітектура. Була розроблена детальна інструкція використання системи для користувача. Показано та описано які дії може виконувати користувач.

А також були проведені випробовування системи, та наведені результати експериментів.

4 РОЗРОБКА СТАРТАП-ПРОЕКТУ

4.1 Опис ідеї проекту

Ідея стартап-проекту полягає у виході на ринок ПЗ для інтелектуального аналізу аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань. Ціллю проекту є заключення партнерського договору з однією або кількома провідними виробниками таких систем для отримання фінансування подальших розробок, та впровадження розробленої системи в якості інтегрованого модуля окремої лінійки обладнання для серійної нарізки.

Таблиця 4.1 – Опис ідеї стартап-проекту

| <i>Зміст ідеї</i> | <i>Напрямки застосування</i> | <i>Вигоди для користувача</i> |
|--|--|--|
| Аналіз аномалій ритмів електрокардіограм | Впровадження систем у провідні клініки та підпис контрактів на її використання | Багатократний приріст швидкості у роботі лікарів, автоматизація процесу аналізу ритмів електрокардіограм, допомога у постановці діагнозу пацієнтам |

Таблиця 4.2 – Сильні, слабкі та нейтральні сторони проекту

| Ідея | (потенційні) товари/концепції конкурентів | | Слабка сторона (W) | Нейтральна сторона (N) | Сильна сторона (S) |
|--|---|-------------------------|--------------------------|------------------------------|--------------------------|
| | <i>Мій проект</i> | <i>HRV software</i> | | | |
| <i>Обробка серцевого ритму</i> | Наявний | Наявний | | + | |

Продовження таблиці 4.2

| Ідея | (потенційні) товари/концепції конкурентів | | Слабка сторона (W) | Нейтральн а сторона (N) | Сильна сторона (S) |
|--|---|-----------|--------------------------|-------------------------------|--------------------------|
| <i>Перетворення запису у скатерограми для подальшого аналізу</i> | Наявний | Відсутній | | | + |
| <i>Проведення кластерного аналізу</i> | Наявний | Відсутній | | | + |
| <i>Можливість анування записів</i> | Наявний | Відсутній | | | + |
| <i>Завантаження звіту про запис</i> | Наявний | Відсутній | | | + |

4.2 Технологічний аудит ідеї проекту

Для визначення здійсненності задуму проекту було проведено його ретельний аналіз. Результати подано у таблиці 4.3.

Таблиця 4.3 – Технологічна здійсненність ідеї проекту

| Ідея проекту | Технології реалізації | Наявність технології | Доступність технології |
|--------------------------------|---|-------------------------|---------------------------|
| <i>Обробка серцевого ритму</i> | інформаційна система розроблена на мові програмування Python3.7 | Наявна | Доступна |

Продовження таблиці 4.3

| Ідея проекту | Технології реалізації | Наявність технології | Доступність технології |
|--|---|--|-------------------------------|
| <i>Перетворення запису у скатерограми для подальшого аналізу</i> | Система на Python3.7 | Наявна | Доступна |
| | Бібліотека для Python - Seaborn | Наявна | Доступна |
| <i>Проведення кластерного аналізу</i> | Метод кластерного аналізу – K-means | Наявна | Доступна |
| <i>Можливість анування записів</i> | інформаційна система розроблена на мові програмування Python3.7 | Додатково реалізується всередині системи | Доступна |
| <i>Завантаження звіту про запис</i> | Додаток до мови Python – xhtml2pdf | Наявна | Доступна |

4.3 Аналіз ринкових можливостей запуску стартап-проекту

Під час впровадження стартап-проекту на ринок, у першу чергу, постають питання з ринкових можливостей проекту, та загроз, що можуть зашкодити підприємству.

Проведемо аналіз на наявність попиту та його обсяг. Для цього оцінимо головних гравців ринку та зведемо дані у таблиці 4.4.

Таблиця 4.4 – Попередня характеристика потенційного ринку стартап-проекту

| Показник стану ринку | Характеристика |
|-----------------------------------|-----------------------|
| <i>Кількість головних гравців</i> | 1. HRV software |
| <i>Загальний обсяг продаж</i> | Невідома |
| <i>Динаміка ринку</i> | Зростає |

Продовження таблиці 4.4

| Показник стану ринку | Характеристика |
|--|--|
| <i>Наявність обмежень для входу</i> | Написання програмного забезпечення, та збільшення можливостей серверної частини для покращення роботи з великою кількістю користувачів |
| <i>Специфічні вимоги до стандартизації та сертифікації</i> | Вимоги до роботи інформаційної системи у медичних закладах |
| <i>Середня норма рентабельності в галузі</i> | Невідома |

Попередня оцінка показала що ринок є досить привабливим для входження, хоча існують і деякі обмеження на нішу, яку розроблюваний продукт здатен зайняти у ньому. Розглянемо потенційні групи клієнтів, їх характеристики та вимоги до товару.

Таблиця 4.5 – Характеристика потенційних клієнтів стартап-проекту

| Потреба | Цільова аудиторія | Відмінності у поведінці цільових груп | Вимоги споживачів до товару / послуги |
|--|--------------------------|--|--|
| <i>Швидкому аналізу даних ритмів електрокардіограм</i> | Пацієнти | Відсутні | Автоматизація процесів |
| | | | Відсутність натовпу у кініках |
| | | | Швидкість отримання результатів |

Продовження таблиці 4.5

| Потреба | Цільова аудиторія | Відмінності у поведінці цільових груп | Вимоги споживачів до товару / послуги |
|--|--------------------------|--|--|
| <i>Консультативна допомога інформаційної системи</i> | Лікарі | Відсутні | Допомога у постановці діагнозу, та складання передбачення щодо захворювання пацієнта |

Після визначення цільових груп споживачів послуг та їх вимог, логічним є аналіз ринкового середовища, а саме визнання факторів загроз та можливостей.

Відомою тактикою боротьби з загрозами є передбачення їх можливих джерел та планування роботи з уникання або зведення наслідків до прийнятних втрат. З відомих на поточний час фактів, виокремимо існуючі та майбутні загрози та наведемо шляхи їх подолання у таблиці 4.6.

Таблиця 4.6 – Фактори загроз

| Фактор | Зміст загрози | Реакція компанії |
|--------------------------------|--|---|
| <i>Кадри</i> | Необхідне навчання кадрів для використання ПЗ | Пошук та навчання людей |
| <i>Відмова у співпраці</i> | Медична компанія відмовляється від співробітництва | Отримання коментарів щодо причини відмови, пропонування компромісу, введення змін |
| <i>Недостатнє фінансування</i> | Недостатнє фінансування компаній | Аналіз та оптимізація витрат, маркетингові заходи з пошуку клієнтів |

Визначимо фактори можливостей, які у змозі позитивно вплинути на стан компаній. Згруповані дані наведемо у таблиці 4.7.

Таблиця 4.7 – Фактори можливостей

| Фактор | Зміст можливості | Реакція компанії |
|--|---|--|
| <i>Науково-технічний</i> | Вдосконалення інформаційної системи | Впровадження в роботу |
| <i>Можливість встановлення монополії</i> | Можливість захопити ринок послуг | Робота над вдосконаленням системи |
| <i>Робота з закордонними замовниками</i> | Встановлення контактів з іноземними замовниками | Локалізація, інтернаціоналізація та сертифікація системи |
| <i>Економічний</i> | Підтримка інновацій у виробництві | Підвищення / Пониження ціни на послугу |

Проаналізувавши ринкове середовище, можна зробити висновок щодо найбільшої загрози підприємству – недостатнє фінансування, тому проект повинен буде зосередити усі зусилля на пошук нових клієнтів та підтримку вже існуючих зв'язків.

Надалі проведемо аналіз пропозицій на ринку та визначимо загальні риси конкуренції на ринку (таблиця 4.8).

Таблиця 4.8 – Ступеневий аналіз конкуренції на ринку

| Особливості конкурентного середовища | В чому проявляється дана характеристика | Вплив на діяльність підприємства |
|---|--|--|
| <i>Монополія</i> | В галузі домінує одна фірма | Надання конкурентоспроможних послуг |
| <i>Міжнародний рівень конкурентної боротьби</i> | Компанії конкуренти з інших країн | Створити основу ПП таким чином, щоб можна було легко переробити даний ПП для використання у галузях інших країн. |
| <i>Внутрішньогалузева ознака</i> | Продукт використовується тільки в одній галузі | Постійне вдосконалення продукту |

Продовження таблиці 4.8

| Особливості конкурентного середовища | В чому проявляється дана характеристика | Вплив на діяльність підприємства |
|---|---|--|
| <i>Товарно-видова конкуренція</i> | Конкуренція між видами ПП, їх особливостями. | Вдосконалити ПП, враховуючи недоліки конкурентів |
| <i>Нецінові переваги</i> | Вдосконалення технології створення ПП, щоб собівартість була нижчою | Удосконалення моделі. Використання більш дешевих технологій для розробки, ніж використовують конкуренти, але тільки якщо ці технології відповідають необхідним вимогам якості. |
| <i>Марочна інтенсивність</i> | Бренд присутній, але його роль незначна | Створення власної марки, Реклама, участь у конференціях, семінарах. |

Після аналізу конкуренції, проведемо аналіз умов конкуренції у галузі та наведемо результати у таблиці 4.9. З проведеного аналізу конкуренції на ринку нами було виявлено, що існує можливість виходу на ринок. Як стратегію на початку роботи можна обрати напрямок роботи на встановлення зв'язків з існуючими замовниками послуг.

Таблиця 4.9 – Аналіз конкуренції в галузі за М. Портером

| | Конкуренти | | Постачальники | Клієнти | Товари замітники |
|-------------------------|---------------------|-------------------|----------------------|--------------------------|---|
| | Прямі | Потенційні | | | |
| <i>Складові аналізу</i> | <i>HRV software</i> | Невідомі | - | Контроль якості продукту | Наявність більш широкого функціоналу, зручнішого інтерфейсу та авторитет (перевірена якість), наявність необхідних сертифікацій |

Продовження таблиці 4.9

| | Конкуренти | | Постачальники | Клієнти | Товари замітники |
|-----------------|---|---------------|----------------------|---|---|
| <i>Висновки</i> | Досить інтенсивна конкурентна боротьба з вже закріпленими на ринку гравцями | Дані відсутні | - | Клієнти диктують умови роботи на ринку: зручний інтерфейс, надійний, швидкий, точний та достовірний ПП для побудови моделей і прогнозів | Необхідно випускати ПЗ не гірше, ніж у конкурентів та розширяти функціонал. |

З аналізу конкуренції можна визначити тип конкуренції – монополістична. Це відбулося тому, що вони були першими на ринку.

Сформуємо перелік факторів конкурентоспроможності (таблиця 4.10). Поки проект не є реалізованим, ми можемо навести лише приблизні можливі фактори.

Таблиця 4.10 – Обґрунтування факторів конкурентоспроможності

| Фактор конкурентоспроможності | Обґрунтування |
|--|--|
| <i>Точність</i> | Використання сучасних технологій для отримання найбільш точних даних |
| <i>Ціна</i> | Можливість давати більш точні результати аналізу, та вихід на медичні ринки. |
| <i>Дані, що зчитуються</i> | Можливість безпечно змінювати дані |
| <i>Орієнтованість на кінцевого споживача</i> | Продукт орієнтований на взаємодію з клієнтом |

На основі вже визначених факторів конкурентоспроможності, проведемо аналіз сильних та слабких сторін стартап-проекту та наведемо результати у таблиці 4.11.

Таблиця 4.11 – Порівняльний аналіз сильних та слабких сторін

| Фактор конкурентоспроможності | Бали 1-20 | Рейтинг товарів конкурентів у порівнянні | | | | | | |
|--|-----------|--|----|----|---|----|----|----|
| | | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| <i>Точність</i> | 15 | | | + | | | | |
| <i>Ціна</i> | 19 | | + | | | | | |
| <i>Дані, що зчитуються</i> | 20 | + | | | | | | |
| <i>Орієнтованість на кінцевого споживача</i> | 12 | | | + | | | | |

Фіналізуємо ринковий аналіз можливостей впровадження проекту складанням матриці аналізу сильних та слабких сторін підприємства і його можливостей та загроз у матриці SWOT-аналізу (таблиця 4.12).

Таблиця 4.12 – SWOT – аналіз стартап-проекту

| Сильні сторони | Слабкі сторони |
|---|--|
| <ul style="list-style-type: none"> – Ціна – Зчитування варіабельності серцевого ритму окрім ЕКГ | <ul style="list-style-type: none"> – Нерозуміння потреб ринку |
| Можливості | Загрози |
| <ul style="list-style-type: none"> – Конкуренція – Поява нових методів аналізу електрокардіограм – Поява нових типів даних | <ul style="list-style-type: none"> – Кадри – Відмова у співпраці – Невідповідність медичним стандартам – Недостатнє фінансування |

Після проведеного SWOT-аналізу можна зробити висновок, що проект буде конкурентоспроможним за рахунок своїх сильних сторін та можливостей.

Основаючись на SWOT-аналізі, логічним є розробка поведінки виведення стартап-проекту на ринок. Для цього аналізується ймовірність отримання ресурсів за певні строки (таблиця 4.13).

Таблиця 4.13 – Альтернативи ринкового впровадження стартап-проекту

| Альтернатива | Ймовірність отримання ресурсів | Строки реалізації |
|---|---------------------------------------|--------------------------|
| <i>Безкоштовне розповсюдження створеного ПП</i> | Дуже висока | 9-12 місяців |
| <i>Створення ПП з подальшим розповсюдженням за певну оплату</i> | Висока | 18-24 місяці |

Після аналізу було обрано альтернативу №1

4.4 Розроблення ринкової стратегії проекту

Для розробки ринкової стратегії, визначимо стратегію охоплення ринку через опис гру потенційних споживачів (таблиця 4.14).

Таблиця 4.14 – Вибір цільових груп потенційних споживачів

| Опис профілю цільової групи | Готовність споживачів сприйняти послугу | Орієнтовний попит в межах цільової групи | Інтенсивність конкуренції в сегменті | Простота входу у сегмент |
|---|--|---|---|---------------------------------|
| <i>Компанії (українські Висока Високий Сильна Склад но 89 та міжнародні) діяльність яких пов'язана з сферою страхової медицини.</i> | Висока | Високий | Сильна | Складно |

Продовження таблиці 4.14

| Опис профілю цільової групи | Готовність споживачів сприйняти послугу | Орієнтовний попит в межах цільової групи | Інтенсивність конкуренції в сегменті | Простота входу у сегмент |
|--|--|---|---|---------------------------------|
| <i>Приватні підприємства міського та міжнародного рівня, діяльність яких пов'язана з медичними дослідженнями</i> | Висока | Високий | Сильна | Складно |
| <i>Приватні клініки</i> | Помірна | Помірний | Помірна | Середня |
| <i>Заклади медичного туризму</i> | Помір на | Слабкий | Слабка | Просто |
| <i>Держ. клініки</i> | Слабка | Слабкий | Слабка | Просто |

За результатами аналізу потенційних груп споживачів було обрано цільові групи – 1,2,3, для яких буде запропоновано даний товар, та визначено стратегію охоплення ринку – стратегію диференційованого маркетингу (компанія працює з декількома сегментами). З метою роботи в обраному ринковому сегменті, сформуємо базову стратегію розвитку (таблиця 4.15).

Таблиця 4.15 – Визначення базової стратегії розвитку

| Обрана альтернатива розвитку проекту | Стратегія охоплення ринку | Ключові конкурентоспроможні позиції відповідно до обраної альтернативи | Базова стратегія розвитку |
|---|---|---|----------------------------------|
| Стратегія спеціалізації | Налагодження зв'язків з клієнтами, індивідуальна модифікація ПЗ під потреби | Висока якість та точність, ухил на довготривалі стосунки | Стратегія диференціації |

Базовою стратегією оберемо стратегію диференціації – орієнтування на потреби користувача. Альтернативною до неї (у разі провалу) буде обрано стратегію спеціалізації – налаштування під окремий цільовий сегмент.

Далі оберемо стратегію конкурентної поведінки на ринку (таблиця 4.16).

Таблиця 4.16 – Визначення базової стратегії конкурентної поведінки

| Чи є проект «першопрохідцем» на ринку? | Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів? | Чи буде компанія копіювати основні характеристики товару конкурента, і які? | Стратегія конкурентної поведінки |
|--|--|---|-------------------------------------|
| Проект не є «першопрохідцем» | Пошук нових клієнтів та перехоплення існуючих | Копіювання лише спільного функціоналу та його розширення | Стратегія заняття конкурентної ніші |

Обраною є стратегія заняття конкурентної ніші з орієнтацією на декілька сегментів. Головною діяльністю є підтримку клієнтів та розширення функціоналу.

Означимо стратегію позиціонування продукту (таблиця 4.17).

Підсумуємо обрані стратегії поведінки:

- **базова стратегія розвитку** – *диференціація*,
- **альтернативна стратегія розвитку** – *спеціалізація*;
- **конкурентна поведінка** – *заняття конкурентної ніші*;
- **обрані сегменти ринку** – *державні і приватні підприємства*.

Таблиця 4.17 – Визначення стратегії позиціонування

| Вимоги до товару цільової аудиторії | Базова стратегія розвитку | Ключові конкурентоспромож ні позиції власного стартап-проекту | Вибір асоціацій, які мають сформувати комплексну позицію власного проект |
|--|---------------------------------|--|---|
| Легкість розуміння, зручний інтерфейс, надійний, швидкий, точний та достовірний ПП для побудови моделей і прогнозів | Стратегія диференціації | Позиція на основі порівняння фірми з товарами конкурентів; Відмінні особливості споживача | Економія часу; Зручність застосування; Практичність та точність результату |

Результатом виконання підрозділу стала узгоджена система рішень щодо ринкової поведінки стартап-компанії, яка визначає напрями роботи стартап компанії на ринку.

4.5 Розроблення маркетингової програми стартап-проекту

Сформовано маркетингову концепцію товару, який отримає споживач. Для цього у таблиці 4.18 підсумовано результати попереднього аналізу конкурентоспроможності товару. Концепція товару – письмовий опис фізичних та інших характеристик товару, які сприймаються споживачем, і набору вигод, які він обіцяє певній групі споживачів.

Таблиця 4.18 – Визначення ключових переваг концепцій потенційного товару (послуги)

| Потреба | Вигода, яку пропонує товар (послуга) | Ключові переваги перед конкурентами |
|--|---|---|
| <i>Швидкість отримання результату;</i> | Швидке зняття кардіограми | Відсутність необхідності звертатися до сторонньої особи/компанії для зняття електрокардіограми. Дані користувача, якими оперує ПП, не передаються третім особам, чого вимагає політика безпеки багатьох компаній. |
| <i>Зручність застосування</i> | Не потрібно мати глибоких знань, для того щоб проводити зняття, та отримувати аналіз електрокардіограми | ПП має вбудовану інструкцію по користуванню |
| <i>Практичність та точність результату</i> | Користувач отримує точні (з малою похибкою розбіжності) результати. | Користувач на виході роботи ПП отримує модель та прогноз, котрі відповідають необхідним показникам достовірності та точності. Отриманий прогноз можна для інтерпретації стану здоров'я. |

Розроблено тривірневу маркетингову модель товару: уточнюється ідея продукту та/або послуги, його фізичні складові, особливості процесу його надання (таблиця 4.19).

1-й рівень При формуванні задуму товару вирішується питання щодо того, засобом вирішення якої потреби і / або проблеми буде даний товар, яка його основна вигода. Дане питання безпосередньо пов'язаний з формуванням технічного завдання в процесі розробки конструкторської документації на виріб.

2-й рівень Цей рівень являє рішення того, як буде реалізований товар в реальному/ включає в себе якість, властивості, дизайн, упаковку, ціну.

3-й рівень Товар з підкріпленням (супроводом) – додаткові послуги та переваги для споживача, що створюються на основі товару за задумом і товару в реальному виконанні (гарантії якості, доставка, умови оплати та ін).

Таблиця 4.19 – Опис трьох рівнів моделі товару

| Рівні товарів | Сутність та складові | | |
|--|--|-------------|-----------------------|
| I. Товар за задумом | Зручність та швидкість отримання практичного результату щодо побудови моделі та прогнозування процесів | | |
| II. Товар у Реальному виконанні | Властивості/характеристики | <i>М/Нм</i> | <i>Вр/Тх /Тл/Е/Ор</i> |
| | <i>Якість</i> | Нм | Тл |
| | <i>Точність</i> | Нм | Тл |
| | <i>Встановлення у медичний заклад</i> | М | Тх |
| | <i>Ціна</i> | Нм | Е |
| | Якість: достовірність побудови аналізу здоров'я серця у домашніх умовах | | |
| Пакування: відсутнє | | | |
| Марка: Smart Heart Checker | | | |
| III. Товар із підкріпленням | Після продажу: персональна підтримка в обслуговуванні за додаткову платню. | | |
| За рахунок чого потенційний товар буде захищено від копіювання: захист інтелектуальної власності, патент на винахід. | | | |

Після формування маркетингової моделі товару слід відмітити, що проект буде захищено від копіювання. Наступним кроком є визначення цінових меж, якими необхідно керуватись при встановленні ціни на потенційний товар (остаточне визначення ціни відбувається під час фінансово-економічного аналізу проекту), яке передбачає аналіз ціни на товари-аналоги або товари

субститути, а також аналіз рівня доходів цільової групи споживачів (таблиця 4.20). Аналіз проведено експертним методом

Таблиця 4.20 – Визначення меж встановлення ціни

| Рівень цін на товари-замінники | Рівень цін на товари-аналоги | Рівень доходів цільової групи споживачів | Верхня та нижня межі встановлення ціни на товар/послугу |
|---------------------------------------|-------------------------------------|---|--|
| Відсутні товари замінники | 1000\$ | Високий рівень доходів | Базова покупка 100\$ |

Наступним кроком є визначення оптимальної системи збуту, в межах якого було прийняте рішення (таблиця 4.21)

Таблиця 4.21 – Формування системи збуту

| Специфіка закупівельної поведінки цільових клієнтів | Функції збуту, які має виконувати постачальник товару | Глибина каналу збуту | Оптимальна система збуту |
|--|---|---------------------------------------|---|
| Цільові клієнти – компанії, які бажають впровадити у своїй роботі сучасні засоби, які допоможуть отримати вигоду та покращити дохідність. Вони цікавляться сучасними розробками та інноваційними рішеннями | Встановлення контактів із споживачами і підтримання їх. Формування попиту і стимулювання збуту. Дослідницька робота зі збору маркетингової інформації. Доробка товару, виходячи з потреб конкретного покупця. | Один (від виробника одразу споживачу) | Прямий канал збуту до споживача, мінімізувати збутові витрати розвиток маркетингового спілкування із споживачем |

Останньою складовою маркетингової програми є розроблення концепції

маркетингових комунікацій, що спирається на попередньо обрану основу для позиціонування, визначену специфіку поведінки клієнтів (таблиця 4.22).

Таблиця 4.22 – Концепція маркетингових комунікацій

| Специфіка поведінки цільових клієнтів | Канали комунікації | Ключові позиції для позиціонування | Концепція рекламного звернення |
|--|---|---|---|
| <i>Пошук шляхів мінімізації витрат</i> | Інтернет ресурси; Спеціалізовані виставки; | Канал першого рівня | Зменшення витрат на обслуговування маршрутів та підвищення лояльності пасажирів |

Результатом підрозділу стала ринкова (маркетингова) програма, що включає в себе концепції товару, збуту, просування та попередній аналіз можливостей ціноутворення, спирається на цінності та потреби потенційних клієнтів, конкурентні переваги ідеї, стан та динаміку ринкового середовища, в межах якого впроваджено проект, та відповідну обрану альтернативу ринкової поведінки.

Висновок стартап-проекту

В даному розділі було проведено аналіз розробленого методу, як частину стартап проекту. Можна зазначити, що проект є досить цікавим з точки зору комерціалізації, так як ринок на який виходить проект все іще формується. На ринку наявна певна кількість конкурентів, але завдяки грамотній стратегії виходу, можливо зайняти правильну нішу. Можна сказати, що подальший розвиток проекту є доцільним, оскільки він знайде свою цільову аудиторію.

ВИСНОВКИ

Серцево-судинні захворювання – це глобальна соціальна проблема, для вирішення якої потрібна не тільки підтримка держави, зосередження всіх зусиль органів виконавчої влади, а й залучення сучасних технологій для розробки нових методів у лікуванні та запобіганню можливих хвороб.

Одним із варіантів є розробка, яка описана в даній магістерській дисертації - «Інтелектуальний аналіз аномалій ритмів електрокардіограм за допомогою спеціалізованих баз знань», який являє систему поєднання методів аналізу даних у вигляді часових рядів.

Основним завданням даної системи - є допомога в виявленні серцево-судинних хвороб на їх ранніх етапах, щоб надати швидке і ефективне лікування для пацієнта, а також забезпечення глобального контролю хронічних захворювань, що безперечно підвищити рівень охорони здоров'я та мінімізувати витрати на нього. В процесі виконання представленої магістерської дисертації було реалізовано такі завдання:

- проведений аналіз існуючих методів та підходів до кластеризації часових рядів;
- обрано оптимальний алгоритм кластеризації даних у вигляді часових рядів
- спроектована власна система виявлення аномалій у ритмах ЕКГ
- проведений аналіз методів та підходів згортання даних
- розроблене власне програмне забезпечення аналізу ритмів ЕКГ та спеціалізованої бази даних до неї
- проведені експерименти на основі тестових та реальних даних.

Для виконання вище зазначених завдань було обрано неієрархічний метод кластеризації k-means, що об'єднує спостереження в k-кластери, в яких кожне спостереження належить до кластера з найближчим середнім з центру цього кластера.

У дисертаційній роботі представлені результати, які відповідно до поставленої мети є рішеннями актуальної задачі. Проведені експериментальні дослідження довели, що запропоновані системи можуть бути успішно використані для вирішення прикладних задач.

Отримані теоретичні результати можуть бути використані для інтелектуального аналізу даних і обробки медико-біологічної, інформації. Середовище проектування: Python, платформи, на яких проводилися дослідження: Microsoft Windows, Linux

Також в магістерській науковій роботі було проведено розгляд власного виробу у вигляді стартап-проекту, та перспектива торгової реалізації продукту, рівень потреби на продукт та його рентабельність. Виконано оцінювання ринкових перспектив і можливостей комерціалізації у вигляді розроблення концепції стартап-проекту в умовах висококонкурентної ринкової економіки глобалізаційних процесів. На ринку наявна монополістична конкуренція, існує декілька фірм конкурентів, але їх товар трохи відрізняється, тому вихід на ринок не буде легким і вимагає розумної стратегії підходу. Для введення ринкової реалізації проекту важливо обрати альтернативу, яка віщує розробку програмного продукту з подальшим поширенням за певну ціну. Можна виділити, що подальший розвиток проекту є раціональним, оскільки він виокремить свою цільову аудиторію. За результатами проведеного дослідження, було визначено можливі шляхи для покращення роботи алгоритму.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бокс, Д., Дженкинс, Г., & Левшин, А. Л. (1974). Анализ временных рядов: Прогноз и управление. Вып. 2. Мир.
2. Kay, S. M. (1993). Fundamentals of statistical signal processing. Prentice Hall PTR.
3. Hamilton, J. (1994). D. (1994), Time Series Analysis.
4. Айвазян, С. А., Енюков, И. С., & Мешалкин, Л. Д.
5. (1983). Прикладная статистика: Основы моделирования и первичная обработка данных. Финансы и статистика.
6. Айвазян, С. А., Енюков, И. С., & Мешалкин, Л. Д. (1985).
7. Прикладная статистика: Исследование зависимостей: Справ.изд. М.:Финансы и статистика, 487.
8. Григорович В.В. Інтелектуальний аналіз аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань / В.В. Григорович, В.І. Баклан, А.К. Очеретяний // Збірник статей LIV Міжнародна конференція "Розвиток науки в ХХІ столітті", 15 листопада 2019р. - Харків. Ч1. - С. 65-74.
9. Айвазян, С. А. (1989). Прикладная статистика: Классификация и снижение размерности: Справочное издание (Vol. 3). Финансы и статистика.
10. Haykin, S. (1994). Neural networks: a comprehensive foundation. Prentice Hall PTR.
11. Abonyi, J., & Feil, B. (2007). Cluster analysis for data mining and system identification. Springer Science & Business Media.
12. Borgelt, C. (2006). Prototype-based classification and clustering.
13. Klawonn, F., Kruse, R., & Timm, H. (1997). Fuzzy shell cluster analysis. In Learning, networks and statistics (pp. 105-119). Springer, Vienna.
14. Gath, I., & Geva, A. B. (1989). Unsupervised optimal fuzzy clustering. IEEE Transactions on Pattern Analysis & Machine Intelligence, (7), 773-780.

16. Gustafson, D. E., & Kessel, W. C. (1979, January). Fuzzy clustering with a fuzzy covariance matrix. In 1978 IEEE conference on decision and control including the 17th symposium on adaptive processes (pp. 761-766). IEEE.143
17. Lindberg, D.A.B., Sharp, G.C., Kingsland, L.C., III, et al., Computer based rheumatology consultant. *Proceedings of MEDINFO-80*, North-Holland, Amsterdam, 1980, pp. 1311–1315.
18. Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project -(edited by Bruce G. Buchanan and Edward H. Shortliffe; ebook version)
19. Hathaway, R. J., & Bezdek, J. C. (2001). Fuzzy c-means clustering of incomplete data. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 31(5), 735-744.
20. Miyamoto, S., Ichihashi, H., Honda, K., & Ichihashi, H. (2008). Algorithms for fuzzy clustering (pp. 1394-1399). Heidelberg: Springer.
21. Li, J., Song, S., Zhang, Y., & Zhou, Z. (2016). Robust k-median and k-means clustering algorithms for incomplete data. *Mathematical Problems in Engineering*, 2016.
22. Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern recognition*, 38(11), 1857-1874.
23. У. Чан, П. Біссекс, Д. форс Django. Розробка веб-додатків на Python - www.symbol.ru/alphabet/666336.html = Python Web Development with Django / пер. з англ. А. Кисельов. - СПб. : Символ-Плюс, 2009. - 456 с. - (High Tech). - ISBN 978-5-93286-167-7
24. Web creator. Складні IT-проекти та Автоматизація бізнесу / Django – Фреймворк на Python[Електронний ресурс]. – Режим доступу : <https://web-creator.ru/articles/django#>
25. Бодянский, Е. В., & Руденко, О. Г. (2004). Искусственные нейронные сети: архитектуры, обучение, применения. Харьков: Телетех, 369.
26. Poljak, B. T., & Tsytkin, J. Z. (1980). Robust identification. *Automatica*, 16(1), 53-63.

27. Huber, P. J. (2011). Robust statistics (pp. 1248-1251). Springer Berlin Heidelberg.
28. Rey, W. J. (2006). Robust statistical methods (Vol. 690). Springer.
29. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (2011). Robust statistics: the approach based on influence functions (Vol. 196). John Wiley & Sons.
30. Gorshkov, Y., Kokshenev, I., Bodyanskiy, Y., Kolodyazhniy, V., & Shylo, O. (2006, September). Robust recursive fuzzy clustering-based segmentation of biological time series. In 2006 International Symposium on Evolving Fuzzy Systems (pp. 101-105). IEEE.
31. Ljung, L. (2002). System Identification: Theory for the User Pers. Peking: Tsinghua University Press and Prentice.
32. Goodwin, G. C., Ramadge, P. J., & Caines, P. E. (1981). A globally convergent adaptive predictor. *Automatica*, 17(1), 135-140.144
33. Li, S. Z. (2009). Markov random field modeling in image analysis. Springer Science & Business Media.
34. Zhang, Z. (1997). Parameter estimation techniques: A tutorial with application to conic fitting. *Image and vision Computing*, 15(1), 59-76.
35. Lee, C. C., Chiang, Y. C., Shih, C. Y., & Tsai, C. L. (2009). Noisy time series prediction using M-estimator based robust radial basis function neural networks with growing and pruning techniques. *Expert Systems with Applications*, 36(3), 4717-4724.
36. Davé, R. N., & Krishnapuram, R. (1997). Robust clustering methods: a unified view. *IEEE Transactions on fuzzy systems*, 5(2), 270-293.
37. Bodyanskiy, Y., Kolodyazhniy, V., & Stephan, A. (2001, October). An adaptive learning algorithm for a neuro-fuzzy network. In International Conference on Computational Intelligence (pp. 68-75). Springer, Berlin, Heidelberg.
38. Otto, P., Bodyanskiy, Y., & Kolodyazhniy, V. (2003). A new learning algorithm for a forecasting neuro-fuzzy network. *Integrated Computer-Aided Engineering*, 10(4), 399-409.

39. Du, W., Inoue, K., & Urahama, K. (2005, August). Robust kernel fuzzy clustering. In International Conference on Fuzzy Systems and Knowledge Discovery (pp. 454-461). Springer, Berlin, Heidelberg.
40. Cochocki, A., & Unbehauen, R. (1993). Neural networks for optimization and signal processing. John Wiley & Sons, Inc.
41. Graupe, D. (2016). Deep learning neural networks: Design and case studies. World Scientific Publishing Company.
42. Григорович В.В. Інтелектуальний аналіз аномалій ритмів електрокардіограм з використанням спеціалізованих баз знань / В.В. Григорович, І.В. Баклан, А.К. Очеретяний // Матеріали III всеукраїнської науково-практичної конференції молодих вчених та студентів «Інформаційні системи та технології управління» (ІСТУ-2019) – м. Київ.: НТУУ «КПІ ім. Ігоря Сікорського», 20-22 листопада 2019 р. – С. 42-44.
43. Цыпкин, Я. З. (1984). Основы информационной теории идентификации. Наука. Гл. ред. физ.-мат. лит.
44. Aggarwal, C. C., & Reddy, C. K. (2014). Data clustering. Algorithms and Application, Boca Raton: CRC Press.
45. Aggarwal, C. C. (2015). Data mining: the textbook. Springer.145
46. Möller-Levet, C. S., Klawonn, F., Cho, K. H., & Wolkenhauer, O. (2003, August). Fuzzy clustering of short time-series and unevenly distributed sampling points. In International Symposium on Intelligent Data Analysis (pp. 330-340). Springer, Berlin, Heidelberg.
47. Cruz, L. P., Vieira, S. M., & Vinga, S. (2015, September). Fuzzy clustering for incomplete short time series data. In Portuguese Conference on Artificial Intelligence (pp. 353-359). Springer, Cham.
48. Evers, F. T., Höppner, F., Klawonn, F., Kruse, R., & Runkler, T.
49. (1999). Fuzzy cluster analysis: methods for classification, data analysis and image recognition. John Wiley & Sons.

50. Bezdek, J. C., Keller, J., Krisnapuram, R., & Pal, N. (1999). *Fuzzy models and algorithms for pattern recognition and image processing* (Vol. 4). Springer Science & Business Media.
51. Bifet, A. (2010, July). Adaptive stream mining: Pattern learning and mining from evolving data streams. In *Proceedings of the 2010 conference on adaptive stream mining: Pattern learning and mining from evolving data streams* (pp. 1-212). Ios Press.
52. Tsoukalas, L. H., & Uhrig, R. E. (1996). *Fuzzy and neural approaches in engineering*. John Wiley & Sons, Inc.
53. Kohonen, T. (2012). *Self-organization and associative memory*(Vol. 8). Springer Science & Business Media.
54. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.
55. Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1, No. 10). New York: Springer series in statistics.
56. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.146
57. Havens, T. C., Bezdek, J. C., & Palaniswami, M. (2012). Incremental kernel fuzzy c-means. In *Computational Intelligence*(pp. 3-18). Springer, Berlin, Heidelberg.
58. Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.
59. Arrow, K. J., & Hurwicz, L. (1958). Gradient method for concave programming, I: Local results. *Studies in Linear and Nonlinear Programming*. Stanford University Press, Stanford, CA, 31, 322-338.
60. Chung, F. L., & Lee, T. (1994). Fuzzy competitive learning. *Neural Networks*, 7(3), 539-551.
61. Zadeh, L. A. (2015). Fuzzy logic—a personal perspective. *Fuzzy sets and systems*, 281, 4-20.
62. Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.

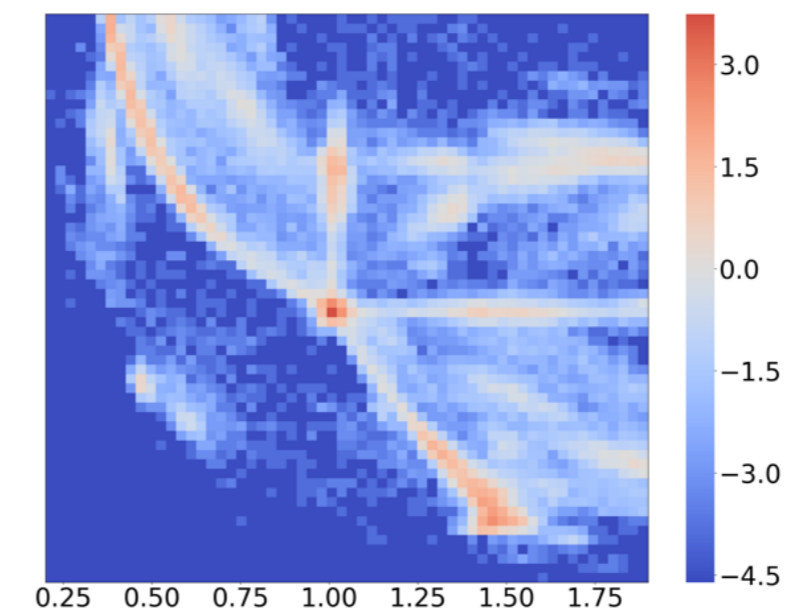
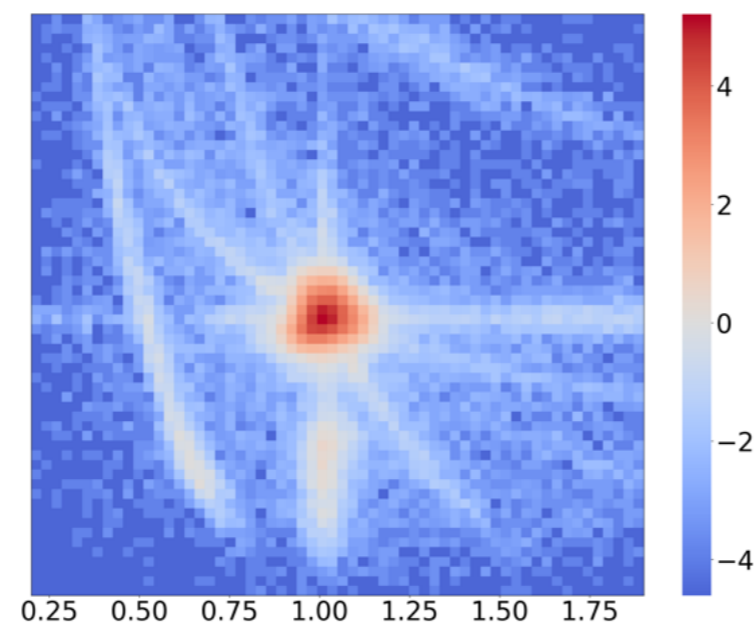
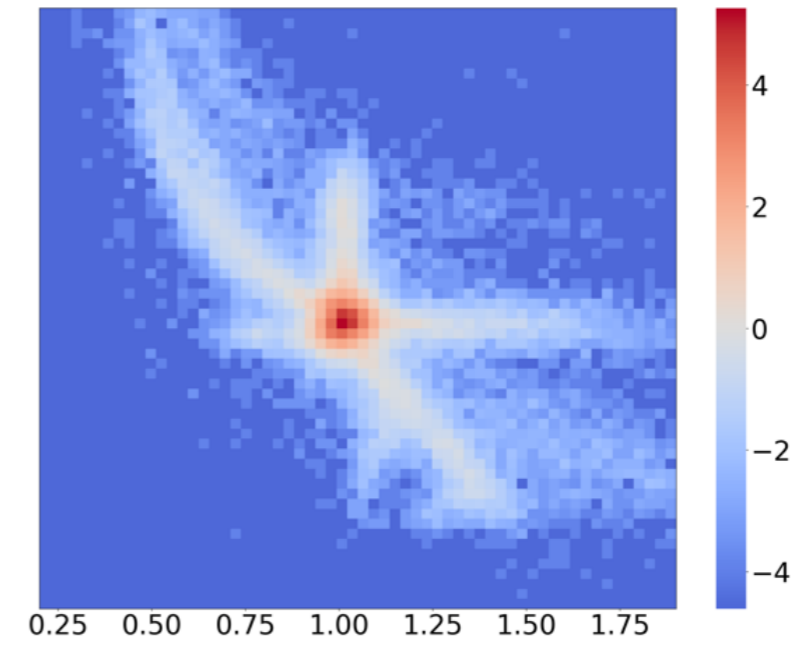
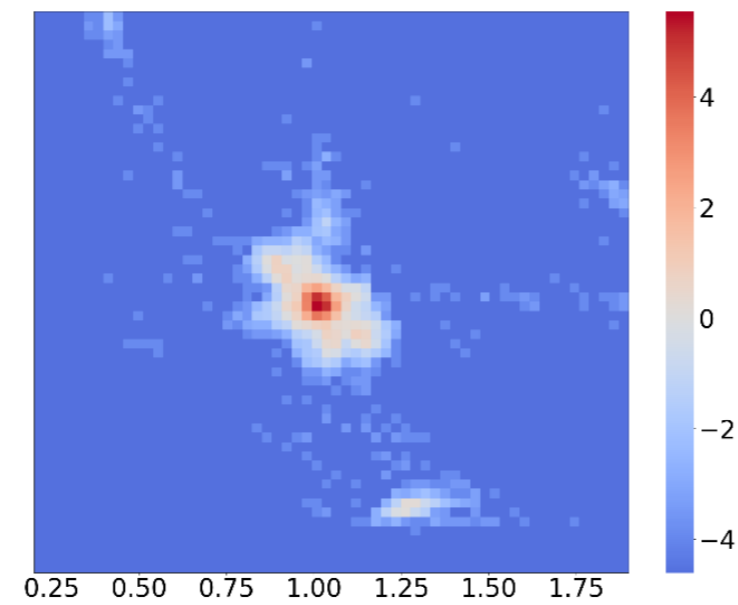
63. Chung, F., & Rhee, H. (2007). Uncertain fuzzy clustering: Insights and recommendations. *IEEE Computational Intelligence Magazine*, 2(1), 44-56.
64. Bezdek, J. C. (2013). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media.
65. Mendel, J. M. (2007). Type-2 fuzzy sets and systems: an overview. *IEEE computational intelligence magazine*, 2(1), 20-29.
66. Zarandi, M. H. F., Zarinbal, M., & Türksen, I. B. (2009, July). Type-II Fuzzy Possibilistic C-Mean Clustering. In *IFSA/EUSFLAT Conf.*(pp. 30-35).
67. Bodyanskiy, Y., Kolodyazhniy, V., & Stephan, A. (2002, September). Recursive fuzzy clustering algorithms. In *Proc. 10th East West Fuzzy Colloquium* (pp. 276-283).
68. Klawonn, F., & Kruse, R. (1997). Constructing a fuzzy controller from data. *Fuzzy sets and systems*, 85(2), 177-193.
69. Prieto, A., Prieto, B., Ortigosa, E. M., Ros, E., Pelayo, F., Ortega, J., & Rojas, I. (2016). Neural networks: An overview of early research, current frameworks and new challenges. *Neurocomputing*, 214, 242-268.147
70. Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, 2(1), 1.
71. Schilling, R. J., Carroll, J. J., & Al-Ajlouni, A. F. (2001). Approximation of nonlinear systems with radial basis function neural networks. *IEEE Transactions on neural networks*, 12(1), 1-15.
72. Nelles, O. (2013). *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Science & Business Media.
73. Specht, D. F. (1991). A general regression neural network. *IEEE transactions on neural networks*, 2(6), 568-576.
74. Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford university press.

75. Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3), 1065-1076.
76. Gan, G., Ma, C., & Wu, J. (2007). *Data clustering: theory, algorithms, and applications* (Vol. 20). Siam.
77. Du, K. L., & Swamy, M. N. (2013). *Neural networks and statistical learning*. Springer Science & Business Media.
78. Rosenblatt, F. (1962). *Principles of Neurodynamics*, Washington, D. DC: Spartan Books.
79. Bodyanskiy, Y., Pliss, I., & Vynokurova, O. (2007). A learning algorithm for forecasting adaptive wavelet-neuro-fuzzy network. In *Fifth International Conference INFORMATION RESEARCH AND APPLICATIONS* (p. 211).
80. Rutkowski, L. (2008). *Computational intelligence: methods and techniques*. Springer Science & Business Media.
81. Everitt, B.S., Landau, S. and Leese, M. (2001), *Cluster Analysis*, Fourth edition, Arnold.
82. Manly, B.F.J. (2005), *Multivariate Statistical Methods: A Case*, Third Edition, Chapman and Hall.
83. Rencher, A.C. (2002), *Methods of Multivariate Analysis*, Second edition, Wiley.

ДОДАТОК А

Графічний матеріал

Плакат 1. Скатерограми центроїди для кожної з груп



Демонстраційний плакат до магістерської дисертації
на тему «Інтелектуальний аналіз аномалій ритмів електрокаріограм з використанням
спеціалізованих баз знань»

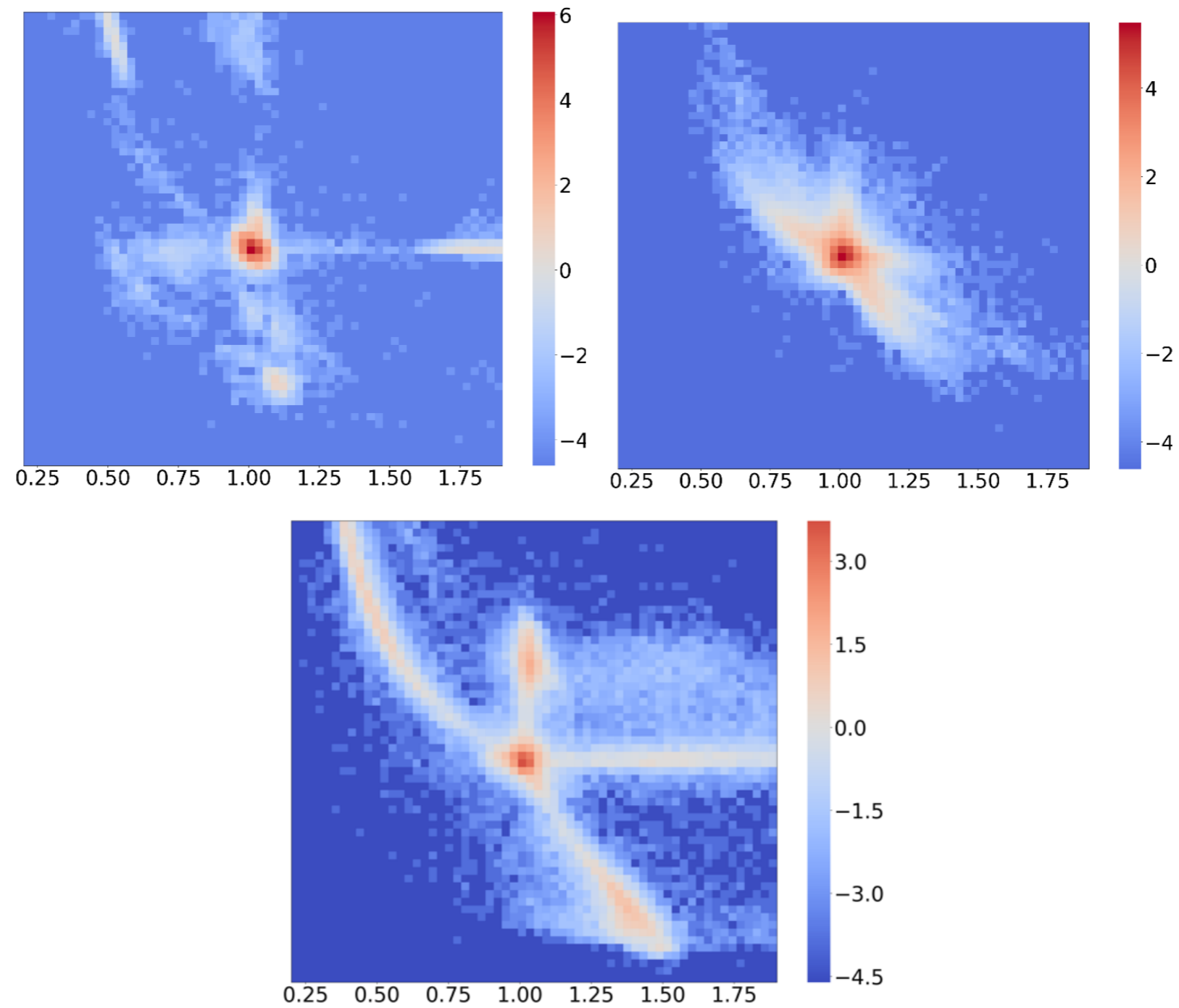
Виконав студент гр. ІС-82мп

Григорович В.В.

Керівник

Баклан І.В.

Плакат 2. Скатерограми центроїди для кожної з груп(продовження)



Демонстраційний плакат до магістерської дисертації
на тему «Інтелектуальний аналіз аномалій ритмів електрокаріограм з використанням
спеціалізованих баз знань»

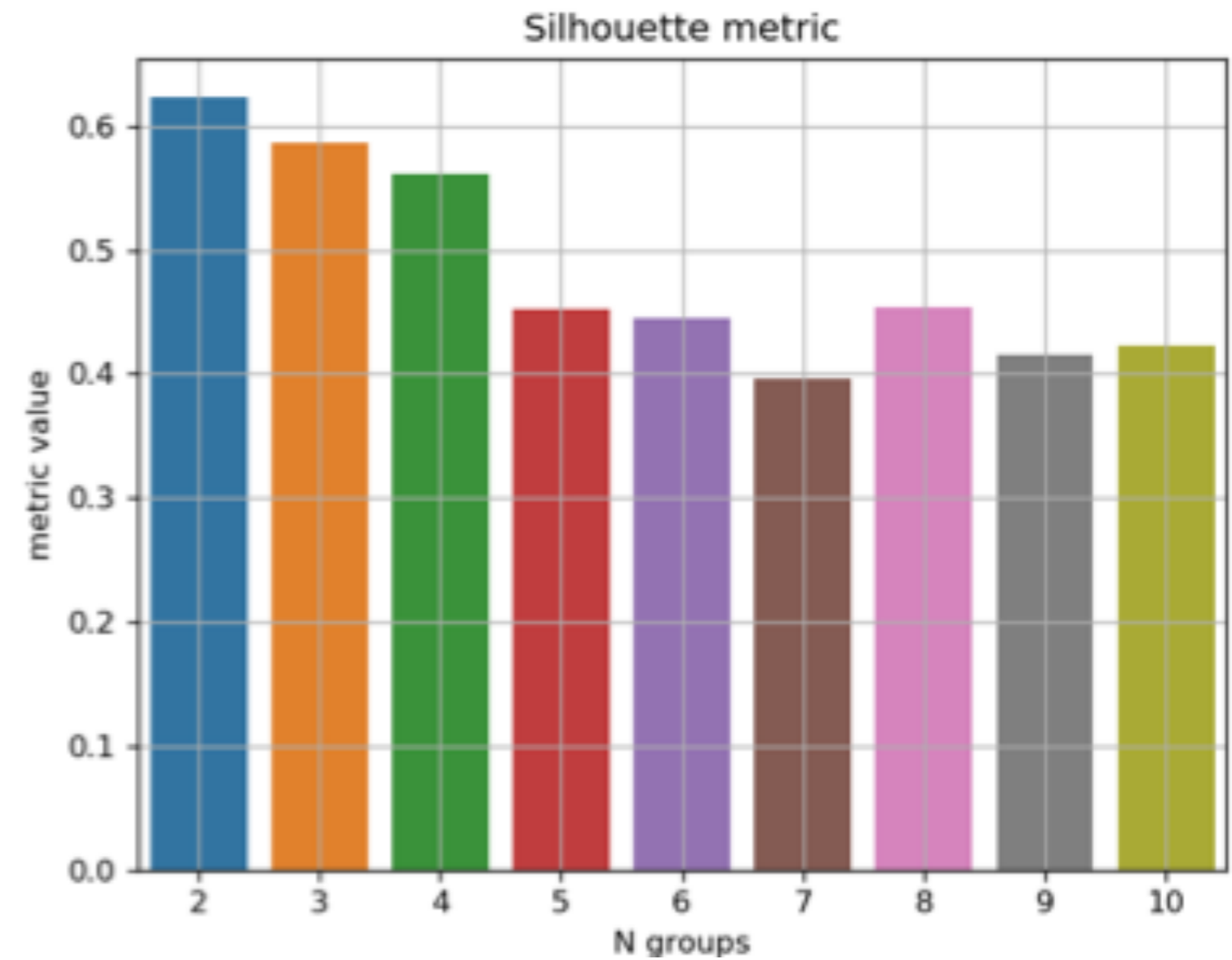
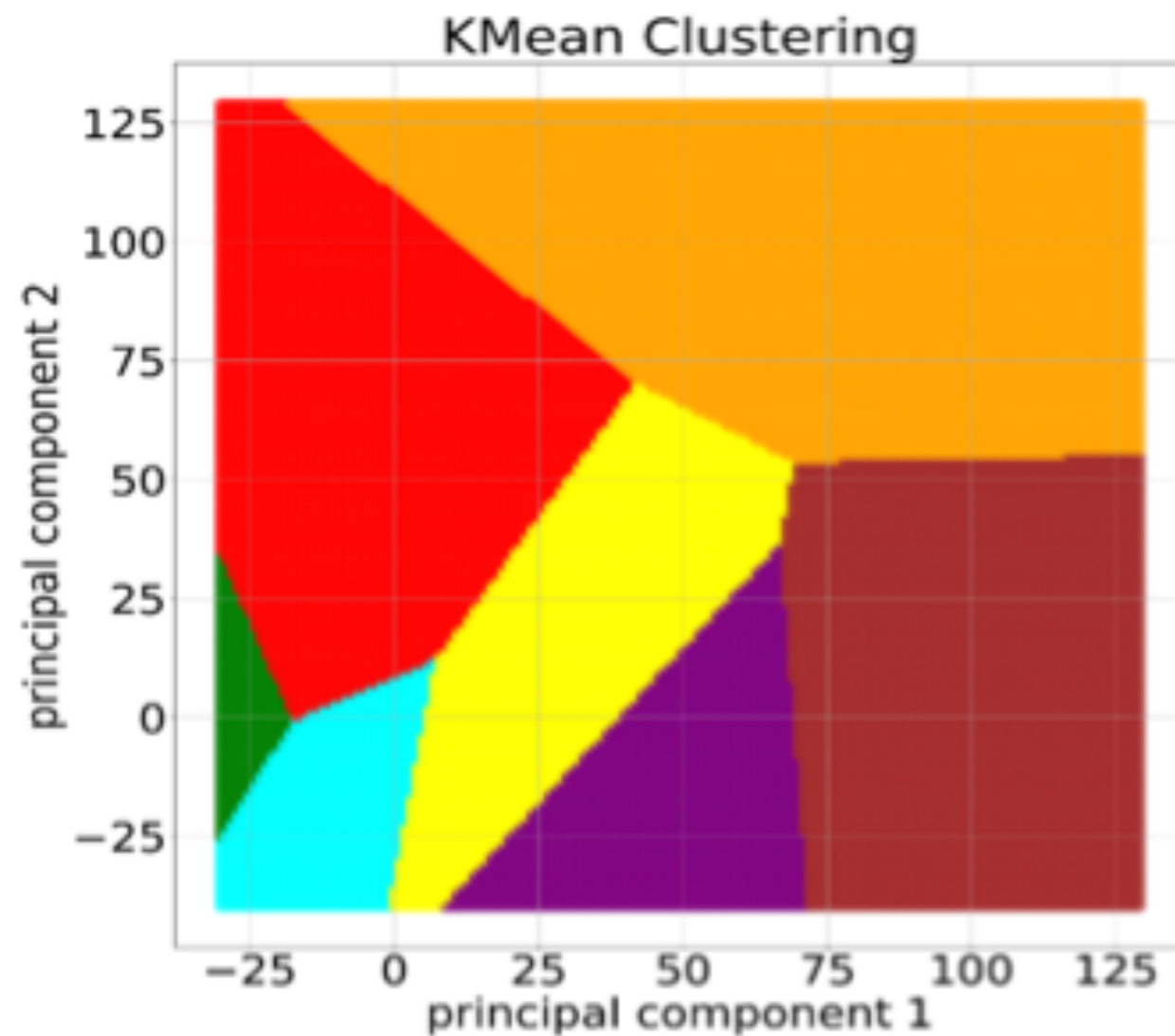
Виконав студент гр. ІС-82мп

Григорович В.В.

Керівник

Баклан І.В.

Плакат 3. Карта зон ризику захворювання у просторі перших двох компонентів та метрика розбиття на кластери



Демонстраційний плакат до магістерської дисертації
на тему «Інтелектуальний аналіз аномалій ритмів електрокаріограм з використанням
спеціалізованих баз знань»

Виконав студент гр. ІС-82мп

Григорович В.В.

Керівник

Баклан І.В.

Плакат 4. Методика розрахунку

Методика розрахунку:



Демонстраційний плакат до магістерської дисертації
на тему «Інтелектуальний аналіз аномалій ритмів електрокаріограм з використанням спеціалізованих баз знань»

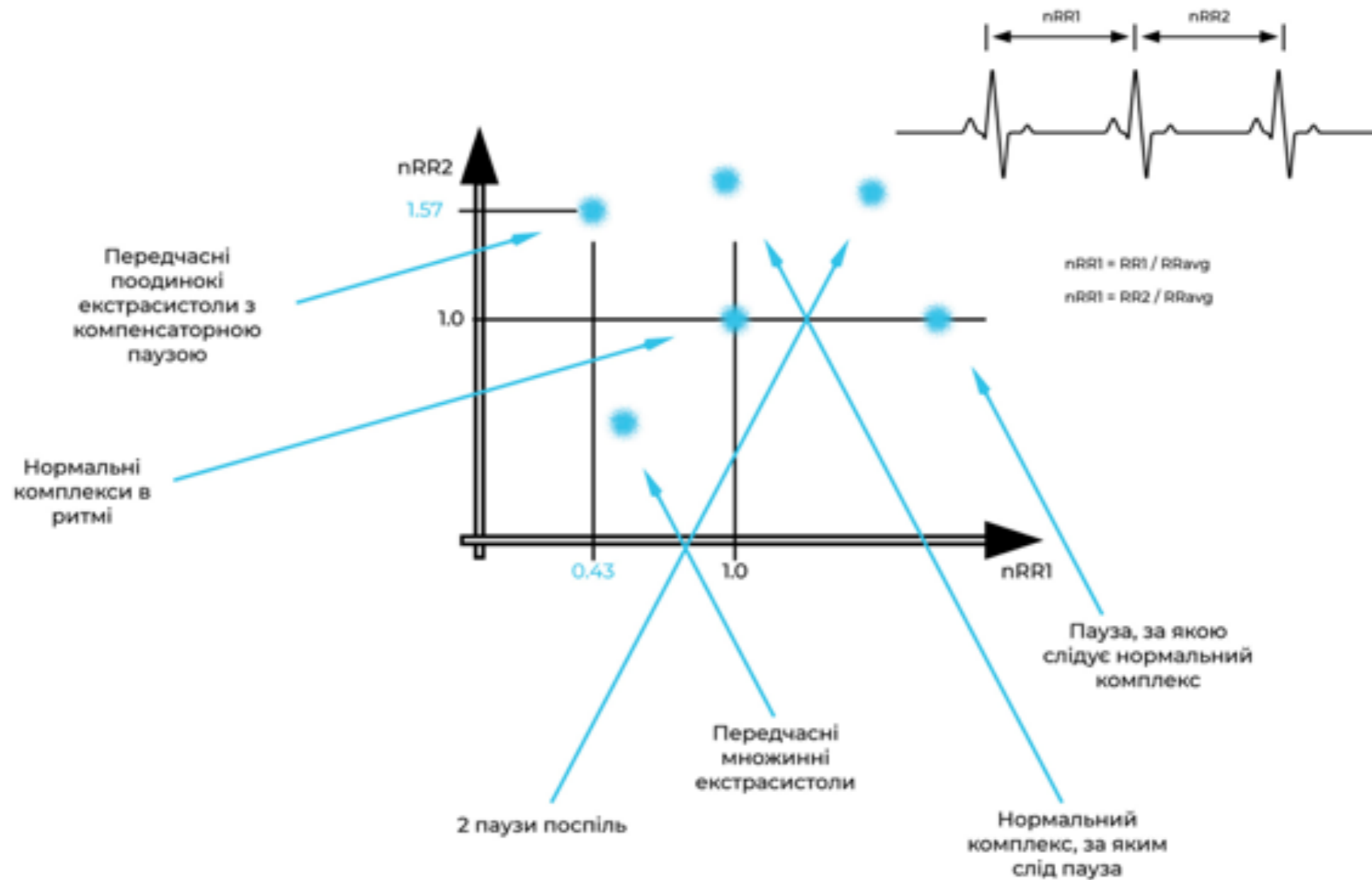
Виконав студент гр. ІС-82мп

Григорович В.В.

Керівник

Баклан І.В.

Плакат 5. Роз'яснення карти зон ризику пацієнта



Демонстраційний плакат до магістерської дисертації
 на тему «Інтелектуальний аналіз аномалій ритмів електрокаріограм з використанням спеціалізованих баз знань»

Виконав студент гр. ІС-82мп

Григорович В.В.

Керівник

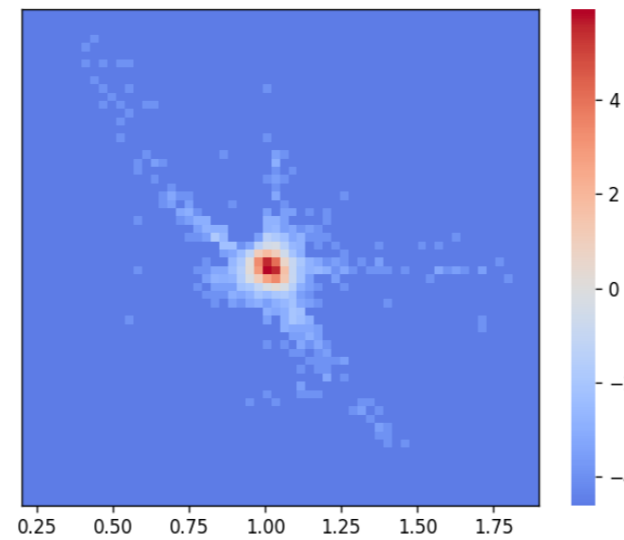
Баклан І.В.

Плакат 6. Схема перетворення вхідних даних до отримання графіків результатів

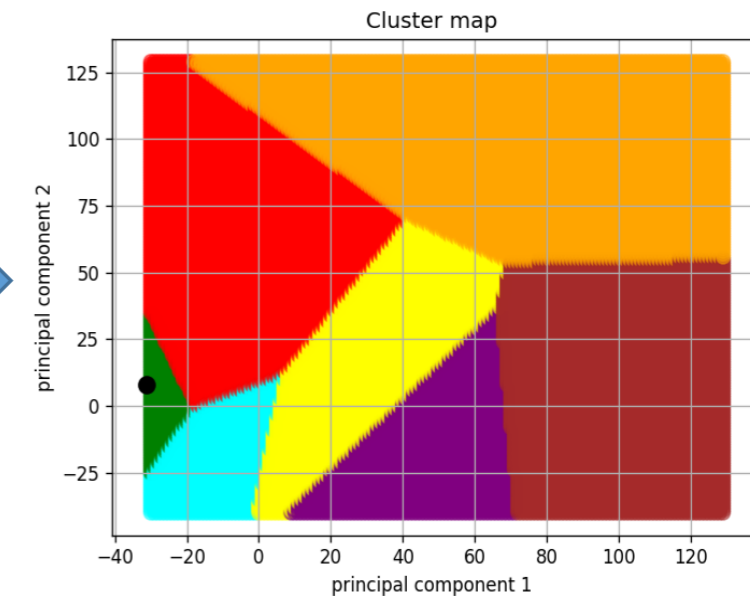
Вхідні дані

| N | Комп | Время | R-R мс | QRS | QR мс | RS мс | скор. | амп.Т | сиг/шум |
|----|------|-------|--------|-----|-------|-------|-------|-------|---------|
| 1 | A | 87 | 680 | 8 | --- | 1006 | 527 | -1025 | 173 |
| 2 | N | 207 | 938 | 86 | 1855 | 2354 | 1465 | --- | 52 |
| 3 | N | 299 | 719 | 94 | 1826 | 2324 | 1504 | 88 | 50 |
| 4 | N | 392 | 727 | 94 | 1836 | 2314 | 1436 | 117 | 52 |
| 5 | N | 484 | 719 | 94 | 1709 | 2188 | 1445 | 59 | 52 |
| 6 | N | 578 | 734 | 86 | 1748 | 2256 | 1484 | 137 | 53 |
| 7 | N | 670 | 719 | 86 | 1797 | 2236 | 1406 | 146 | 53 |
| 8 | N | 763 | 727 | 94 | 1895 | 2354 | 1523 | 117 | 52 |
| 9 | N | 856 | 727 | 86 | 1846 | 2324 | 1436 | 156 | 51 |
| 10 | N | 949 | 727 | 86 | 1777 | 2295 | 1406 | 127 | 50 |
| 11 | N | 1041 | 719 | 94 | 1748 | 2227 | 1484 | 127 | 52 |
| 12 | N | 1134 | 727 | 94 | 1768 | 2217 | 1406 | 137 | 54 |
| 13 | N | 1226 | 719 | 86 | 1650 | 2139 | 1367 | -98 | 49 |
| 14 | N | 1319 | 727 | 94 | 1895 | 2324 | 1484 | 107 | 50 |
| 15 | N | 1412 | 727 | 94 | 1816 | 2275 | 1426 | 166 | 50 |
| 16 | N | 1505 | 727 | 86 | 1748 | 2217 | 1406 | 146 | 53 |
| 17 | N | 1598 | 727 | 102 | 1807 | 2334 | 1494 | 78 | 51 |
| 18 | N | 1691 | 727 | 117 | 1875 | 2451 | 1533 | --- | 58 |
| 19 | N | 1784 | 727 | 86 | 1836 | 2451 | 1504 | 98 | 66 |
| 20 | N | 1877 | 727 | 86 | 1943 | 2383 | 1543 | 127 | 50 |

Скатерограма

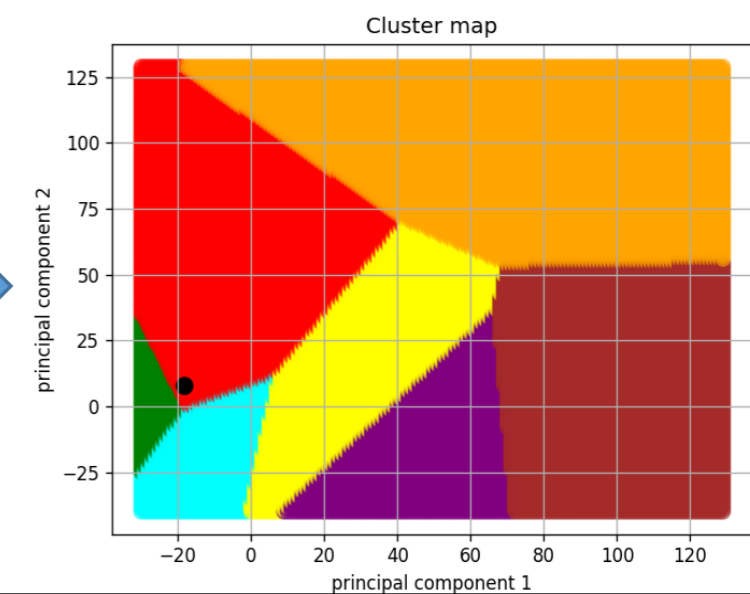
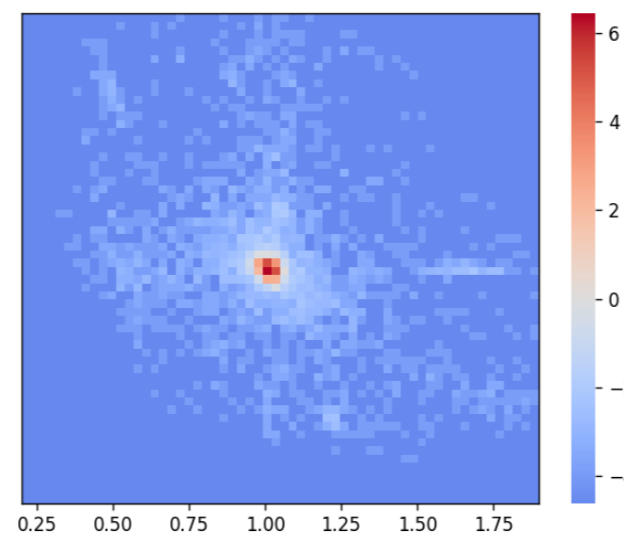


Кластеризована карта



Приклад 1

| N | Комп | Время | R-R мс | QRS | QR мс | RS мс | скор. | амп.Т | сиг/шум |
|----|------|-------|--------|-----|-------|-------|-------|-------|---------|
| 1 | A | 87 | 680 | 8 | --- | 1006 | 527 | -1025 | 173 |
| 2 | N | 207 | 938 | 86 | 1855 | 2354 | 1465 | --- | 52 |
| 3 | N | 299 | 719 | 94 | 1826 | 2324 | 1504 | 88 | 50 |
| 4 | N | 392 | 727 | 94 | 1836 | 2314 | 1436 | 117 | 52 |
| 5 | N | 484 | 719 | 94 | 1709 | 2188 | 1445 | 59 | 52 |
| 6 | N | 578 | 734 | 86 | 1748 | 2256 | 1484 | 137 | 53 |
| 7 | N | 670 | 719 | 86 | 1797 | 2236 | 1406 | 146 | 53 |
| 8 | N | 763 | 727 | 94 | 1895 | 2354 | 1523 | 117 | 52 |
| 9 | N | 856 | 727 | 86 | 1846 | 2324 | 1436 | 156 | 51 |
| 10 | N | 949 | 727 | 86 | 1777 | 2295 | 1406 | 127 | 50 |
| 11 | N | 1041 | 719 | 94 | 1748 | 2227 | 1484 | 127 | 52 |
| 12 | N | 1134 | 727 | 94 | 1768 | 2217 | 1406 | 137 | 54 |
| 13 | N | 1226 | 719 | 86 | 1650 | 2139 | 1367 | -98 | 49 |
| 14 | N | 1319 | 727 | 94 | 1895 | 2324 | 1484 | 107 | 50 |
| 15 | N | 1412 | 727 | 94 | 1816 | 2275 | 1426 | 166 | 50 |
| 16 | N | 1505 | 727 | 86 | 1748 | 2217 | 1406 | 146 | 53 |
| 17 | N | 1598 | 727 | 102 | 1807 | 2334 | 1494 | 78 | 51 |
| 18 | N | 1691 | 727 | 117 | 1875 | 2451 | 1533 | --- | 58 |
| 19 | N | 1784 | 727 | 86 | 1836 | 2451 | 1504 | 98 | 66 |
| 20 | N | 1877 | 727 | 86 | 1943 | 2383 | 1543 | 127 | 50 |



Приклад 2

Демонстраційний плакат до магістерської дисертації
на тему «Інтелектуальний аналіз аномалій ритмів електрокаріограм з використанням спеціалізованих баз знань»

Виконав студент гр. ІС-82мп

Григорович В.В.

Керівник

Баклан І.В.