

МАТЕРІАЛИ ІV МІЖНАРОДНОЇ
СТУДЕНТСЬКОЇ НАУКОВОЇ
КОНФЕРЕНЦІЇ

НАУКА СЬОГОДЕННЯ:
ВІД ДОСЛІДЖЕНЬ ДО
СТРАТЕГІЧНИХ РІШЕНЬ



**М. ІВАНО-ФРАНКІВСЬК,
УКРАЇНА**

**17 ЧЕРВНЯ
2022 РІК**

МАТЕРІАЛИ IV МІЖНАРОДНОЇ
СТУДЕНТСЬКОЇ НАУКОВОЇ
КОНФЕРЕНЦІЇ

**НАУКА СЬОГОДЕННЯ:
ВІД ДОСЛІДЖЕНЬ ДО
СТРАТЕГІЧНИХ РІШЕНЬ**

м. Івано-Франківськ, Україна
17 червня 2022 рік

Вінниця, Україна
«Європейська наукова платформа»
2022

**УДК 001(08)
Н 34**



Голова оргкомітету: Кореньюк І.О.

Верстка: Зрада С.І.

Дизайн: Бондаренко І.В.



Матеріали конференції знаходяться у відкритому доступі на умовах ліцензії CC BY-NC 4.0 International.

Н 34

Наука сьогодення: від досліджень до стратегічних рішень: матеріали IV Міжнародної студентської наукової конференції, м. Івано-Франківськ, 17 червня, 2022 рік / ГО «Молодіжна наукова ліга». — Вінниця: ГО «Європейська наукова платформа», 2022. — 368 с.

ISBN 978-617-8037-71-0

DOI 10.36074/liga-inter-17.06.2022

Викладено матеріали учасників IV Міжнародної мультидисциплінарної студентської наукової конференції «Наука сьогодення: від досліджень до стратегічних рішень», яка відбулася 17 червня 2022 року у місті Івано-Франківськ, Україна.

УДК 001 (08)

© Колектив учасників конференції, 2022

© ГО «Молодіжна наукова ліга», 2022

ISBN 978-617-8037-71-0

© ГО «Європейська наукова платформа», 2022

ЗМІСТ

СЕКЦІЯ 1. ЕКОНОМІЧНА ТЕОРІЯ. МАКРО- ТА РЕГІОНАЛЬНА ЕКОНОМІКА

ВИНИКНЕННЯ ТА ЗАГАЛЬНА ХАРАКТЕРИСТИКА НЕОЛІБЕРАЛІЗМУ
Мороз М.Ю., *Науковий керівник: Курденко О.В.* 14

КРИПТОВАЛЮТИ І СВІТОВА ЕКОНОМІКА
Туріщева М.Б., *Науковий керівник: Дорошенко Н.О.* 17

ПРІОРИТЕТИ ПІСЛЯВОЄННОГО РОЗВИТКУ УКРАЇНИ
Кругла С.Ю., *Науковий керівник: Дугієнко Н.О.* 20

СЕКЦІЯ 2. ФІНАНСИ ТА БАНКІВСЬКА СПРАВА; ОПОДАТКУВАННЯ, ОБЛІК І АУДИТ

БАНКІВСЬКА СИСТЕМА УКРАЇНИ В УМОВАХ ВОЄННОГО СТАНУ
Дубініна С.Д., *Науковий керівник: Дорошенко Н.О.* 22

НАПРЯМИ АКТИВІЗАЦІЇ ВИКОРИСТАННЯ ЛІЗИНГУ ДЛЯ РОЗБУДОВИ
ВІТЧИЗНЯНОГО БІЗНЕСУ
Безвиненко Б.В., *Науковий керівник: Тимошенко І.П.* 24

ПІДХІД ДО МОДЕЛЮВАННЯ ФІНАНСОВИХ РИНКІВ
Колесник А.С., *Науковий керівник: Діденко А.В.* 26

УПРАВЛІНСЬКИЙ ОБЛІК ДОХОДІВ В ГОСПОДАРСЬКІЙ ДІЯЛЬНОСТІ
ПІДПРИЄМСТВ КОНДИТЕРСЬКОЇ ПРОМИСЛОВОСТІ
Шуневич А.В., *Науковий керівник: Гайдучок Т.С.* 28

УПРАВЛІНСЬКИЙ ОБЛІК СОБІВАРТОСТІ НА ПІДПРИЄМСТВАХ ВИРОБНИКІВ
ЯГІД
Науменко І.В., *Науковий керівник: Гайдучок Т.С.* 31

ФІНАНСОВІ РИЗИКИ ПІДПРИЄМСТВА
Вакуленко В.О., *Науковий керівник: Бондаренко Н.М.* 34

ФУНКЦІЇ ФІНАНСІВ. ДИСКУСІЙНІ ПИТАННЯ ЩО ДО ФУНКЦІЙ ФІНАНСІВ
Шинкаренко В.О. 36

ЦИФРОВІ ВАЛЮТИ ЦЕНТРАЛЬНИХ БАНКІВ ЯК НОВА ФОРМА ГРОШЕЙ В
УМОВАХ ЦИФРОВІЗАЦІЇ ЕКОНОМІКИ
Ситник Д.Р., *Науковий керівник: Дорошенко Н.О.* 38

ШЛЯХИ УДОСКОНАЛЕННЯ ОРГАНІЗАЦІЇ ОБЛІКУ ВИРОБНИЧИХ ЗАПАСІВ
Тищенко Т.О., *Науковий керівник: Глухова В.І.* 41

СЕКЦІЯ 3. МАРКЕТИНГОВА ТА ЛОГІСТИЧНА ДІЯЛЬНІСТЬ

TRANSFER OF VALUES THROUGH THE MERGERS
Ploskina Y., *Supervisor: Mostova A.* 43

ЛЕПБУК ЯК ІННОВАЦІЙНИЙ СПОСІБ ОРГАНІЗАЦІЇ ОСВІТНЬОЇ ДІЯЛЬНОСТІ В ПОЧАТКОВІЙ ШКОЛІ Козловська Х.Р., Науковий керівник: Багай Б.М.....	266
ОРГАНІЗАЦІЯ ІГРОВОЇ ДІЯЛЬНОСТІ МОЛОДШИХ ШКОЛЯРІВ Усач Н.С., Науковий керівник: Лесик А.С.....	270
ОРГАНІЗАЦІЙНА ПОВЕДІНКА СТУДЕНТІВ І ВИКЛАДАЧІВ ПІД ЧАС НАВЧАЛЬНОЇ ДІЯЛЬНОСТІ Черниш В.Р., Науковий керівник: Бугаєвська Ю.В.....	273
САМОМОТИВАЦІЯ СТУДЕНТІВ ДО НАВЧАННЯ Дамчук М.С., Науковий керівник: Бойко Л.М.....	275
САМОСТІЙНА РОБОТА ІНОЗЕМНИХ СТУДЕНТІВ-НЕФІЛОЛОГІВ ПРИ НАВЧАННІ РЕФЕРУВАННЯ НАУКОВИХ ТЕКСТІВ Сагдієва І., Науковий керівник: Копилова О.В.....	277
УРОК МИСТЕЦТВА В ЗАГАЛЬНООСВІТНІЙ ШКОЛІ ЯК ФОРМА ВИХОВАННЯ ЕСТЕТИЧНОЇ КУЛЬТУРИ УЧНІВ Божко Д.К., Науковий керівник: Сидоренко Т.Д.....	279

СЕКЦІЯ 23. ПСИХОЛОГІЯ ТА ПСИХІАТРІЯ

ВІД УСАМІТНЕННЯ ДО САМОТНОСТІ Федоренко С.І., Науковий керівник: Грабар Н.Г.	282
ПОРІВНЯЛЬНИЙ АНАЛІЗ ВИРАЖЕНОСТІ ТОЛЕРАНТНОСТІ ДО НЕВИЗНАЧЕНОСТІ У РІЗНИХ ПОКОЛІНЬ Власенко Т.М., Симоненко А.Ю., Науковий керівник: Чумак О.О.....	285
ТЕМПЕРАМЕНТ ЯК ЧИННИК ФОРМУВАННЯ ЕМОЦІЙНИХ ПЕРЕЖИВАНЬ: АРИСТОТЕЛЬ ТА СУЧАСНА ПСИХОЛОГІЯ Амрахова В.О., Науковий керівник: Фролова Є.В.	289

СЕКЦІЯ 24. МЕДИЧНІ НАУКИ ТА ГРОМАДСЬКЕ ЗДОРОВ'Я

ВПЛИВ НА ФОРМУВАННЯ ОСОБИСТОСТІ СОЦІАЛЬНИХ МЕРЕЖ Луців Л.В., Лемак-Деміда Я.В., Науковий керівник: Рогач І.М.....	291
ДО ПИТАННЯ АКТУАЛЬНОСТІ САНІТАРНОЇ ГРАМОТНОСТІ НАСЕЛЕННЯ УКРАЇНИ Дильдіна А.А., Науковий керівник: Валецький Ю.М.....	293
ДОСЛІДЖЕННЯ НЕВИЗНАЧЕНИХ МЕДИЧНИХ ЗНАНЬ ТА КІЛЬКІСНА ОЦІНКА НЕВИЗНАЧЕНОСТІ Койдан А.А., Науковий керівник: Кузьомін О.Я.....	296
КАРДІОНЕВРОЗ Банетішвілі Д.Д., Науковий керівник: Екштейн К.С.	307
КЕСОННА, ВИСОТНА ХВОРОБИ, ЇХ ПРИЧИНИ, ПРОЯВИ, НЕВІДКЛАДНА ДОПОМОГА, ПРОФІЛАКТИКА. "ВИБУХОВА ДЕКОМПРЕСІЯ" Солошенко Марія С., Науковий керівник: Курило Т.М.....	310

Койдан Анастасія Андріївна, здобувач вищої освіти факультету інформаційних радіотехнологій та технічного захисту інформації
Харківський національний університет радіоелектроніки, Україна

Науковий керівник: Кузьомін Олександр Якович, д-р. техн. наук,
професор, професор кафедри Інформатики
Харківський національний університет радіоелектроніки, Україна

ДОСЛІДЖЕННЯ НЕВИЗНАЧЕНИХ МЕДИЧНИХ ЗНАНЬ ТА КІЛЬКІСНА ОЦІНКА НЕВИЗНАЧЕНОСТІ

Вступ

Спроба використовувати логіку першого порядку для подання знань в таких проблемних областях, як медична діагностика COVID-19, у багатьох випадках закінчується невдачею з трьох основних причин, описаних нижче [1,3].

– **Економія зусиль на дослідження.** Для формування повної множини антецедентів або консеквентна, необхідного для складання правил, які не мають винятків, потрібно занадто багато роботи і часу, крім того застосування таких правил є сама по собі занадто складною справою.

– **Відсутність теоретичних знань про COVID-19.** Медична наука не має під час діагностування повної теоретичної обґрунтованості для даної проблемної ситуації, яку отримали медичні фахівці під час навчання. Необхідно постійно підвищувати фаховий рівень, мати інтерес до нових технологій і результатів досліджень. Тут може використовуватись різноманітні сучасні можливості, зокрема відомості, які у великому обсязі знаходяться в Інтернет.

– **Відсутність практичних знань.** Навіть якщо відомі всі правила логічного виводу, може залишатися невизначеність щодо діагнозу даного конкретного пацієнта, оскільки всі необхідні обстеження були або не могли бути виконані, або мали мале практичне підтвердження в особистій лікувальній практиці лікаря, або не вистачило часу на отримання достовірного діагнозу чи коштів у пацієнта та інше. Для скорочення часу на діагностування і підвищення якості діагнозу.

Отже, при аналізі проблем діагностичною невизначеності необхідно враховувати таку невизначеність, яка пов'язана з вимірюваннями параметрів стосовно стану організму, які поповнюють бази даних (БД) так і знань (БЗ) на всіх етапах діагностування [1,5].

Відзначимо, що знання дозволяють саме краще та точніше сформулювати висловлювання, які стосуються медичної проблеми діагностування COVID-19 тільки з певним **ступенем впевненості** (degree of belief).

У цьому дослідженні інструментальним засобом для обліку і аналізу ступенів впевненості буде **теорія ймовірностей**, яка зв'язана з баєсовим представленням даних і знань про багатофакторний стан проблеми [1].

1. Аналіз предметної області

Використання ймовірності до аналізу даних і знань надають до методу оцінки сумарного обліку невизначеності, що виникає з причин економії зусиль і відсутності знань. Ми не можемо знати з усією впевненістю, що турбує даного конкретного пацієнта, які у нього ознаки і симптоми хвороби найбільш значимі для виявлення правильного діагнозу. Ступінь істинності, на відміну від ступенів впевненості, є

предметом нечіткої логіки [2].

Проблема оцінки і врахування невизначеності медичних даних і знань щодо медичних аналізів і різноманітних досліджень має дуже важливе значення. Системний аналіз і прийняття рішень стосовно діагностування стану організму людини є актуальною та досить складною проблемою [1,3].

У дослідженні розглядаються можливості оцінити та врахувати невизначеність даних на прикладі розповсюдженого засобу контролю температури тіла, при використанні комп'ютерної томографії та об'єднання використання іммуносигнатури щодо аналізу даних та зображень [7] під час пандемії COVID-19.

В медичній сфері поняття невизначеності тісно пов'язане з проблемою повноти знання про середовище функціонування організму людини, зокрема обмеженістю інформації, яка необхідна для прийняття діагностичних рішень. У даному випадку обмеженість інформаційного ресурсу викликана невідповідністю його основним якісним характеристикам, зокрема таких:

- **повнота** – відображення всіх медичних процесів та явищ, які є суттєвими та важливими для прийняття раціонального діагностичного рішення;
- **зрозумілість** – доступність її усвідомлення тим користувачам (лікарям), для яких вона призначена;
- **достовірність** – об'єктивність та відповідність реальному стану частин та цілком організму людини;
- **актуальність** – відповідність поточному моменту часу.

Виходячи з цього, невизначеністю є недостатність інформації про умови, в яких буде здійснюватися діагностування чи моніторинг стану пацієнта під час лікування, яка обумовлює складність визначення кінцевих результатів діагностування на всіх рівнях медичної діяльності.

Крім зазначеного, зауважимо, що для вирішення поставленої проблеми будимо використовувати мультиагентне представлення медичної системи, яка буде використана для діагностування та лікування [1].

2. Постановка задачі дослідження навколишнього середовища та стану пацієнта

Необхідно розробити адаптивний метод проектування спеціалізованого медичного комплексу для діагностування захворювань COVID-19 з мультиагентним представленням компонентів комплексу (ММКД) [1,5].

Формальне представлення структури ММКД $\{\text{Str}\}$ має такий вигляд:

$$\begin{aligned} \{\text{Str}\} &= \left\{ \text{Str}_i, \Pi(\text{Sit}_i), \text{Alg}''(\Pi), K_{\text{Str}_i} \right\}, \\ \text{Str}_{\text{opt}}(K_{\text{Str}}) &\in \{\text{Str}\}, \\ \mathfrak{R}(\text{Str}_{\text{opt}}) &= \left(\mathfrak{R}(\text{Str}_i), i = \overline{1, m} \right) \end{aligned}$$

$de K_{\text{Str}} = f(\mathfrak{Z}, \mathfrak{R})$ – агентні компоненти структури ММКД, які визначаються на основі багато профільних досліджень стану організму пацієнта (опитування пацієнта, клінічних аналізів, знімків після призначених лікарем досліджень стану з елементів організму, наприклад, знімків комп'ютерної томографії (КТ), результатів магніторезонансного дослідження (МРТ) та інше), крім того треба враховувати оцінку витрат для залучення необхідних ресурсів \mathfrak{R} для

діагностування; $\{\text{Str}\}$ - множина структурних рішень ММКД; Str_i - елемент цієї множини; $\Pi(\text{Str}_i)$ - простір агентних структур, який розташовано у базі прецедентів, що відповідає вдалим раціональним, аналогічним медичним рішенням відносно діагностування чи стратегії лікування; Str_{opt} - оптимальна структура ММКД (у разі мінімально допустимих ризиків і витрат часу на дослідження щодо діагностування стану пацієнта); Alg'' - алгоритми діагностування.

Обчислення невизначеності можливо виконати за наступними методами:

1. Використати середнє арифметичне значення, наприклад, з усіх вимірів освітлення в даній точці:

$$E = \frac{1}{n} \sum_{i=1}^n E_i \quad (1)$$

де E_i – виміри в даній точці.

2. Для джерел невизначеності випадкового характеру обчислюємо невизначеність за типом А (Рис.1):

$$u_A(E) = \sqrt{\frac{\sum_{i=1}^n (E_i - E)^2}{n(n-1)}}. \quad (2)$$

3. Для джерел невизначеності систематичного характеру (приладова похибка) обчислюємо невизначеність за типом Б (Рис.1):

$$u_B(E) = \frac{\Delta E}{\sqrt{3}}. \quad (3)$$

4. Застосувати сумарну стандартну невизначеність:

$$u_C(E) = \sqrt{u_A^2(E) + u_B^2(E)}. \quad (4)$$

5. Використати довірчу ймовірність (ймовірність охоплення) $P = 0.95$ (рекомендується). Наприклад, задаємо коефіцієнт охоплення $k = 2$ маємо розширену невизначеність вимірювань:

$$u = k u_C. \quad (5)$$

Методи оцінки щодо невизначеності не стверджують відсутність справжнього значення, але стверджують відсутність точного значення знання про справжню цінність (Рис.1).

Загальний ланцюжок випробувань у лабораторній медицині включає кілька можливих джерел невизначеності від клінічного рішення про замовлення тесту через біологічні варіації, преданалітичну, аналітичну та постаналітичну фази до значення результату тесту у поточні клінічні рішення.

Метод помилок зосереджуються насамперед на аналітичній фазі (властивості вимірювальної системи), тоді як методи невизначеності фокусуються на підрахунку всіх джерел невизначеностей, включаючи біологічні зміни, доаналітичні варіації, аналітичну варіацію та постаналітичну варіацію як допоміжний засіб у діагностиці

та моніторингу лікування.

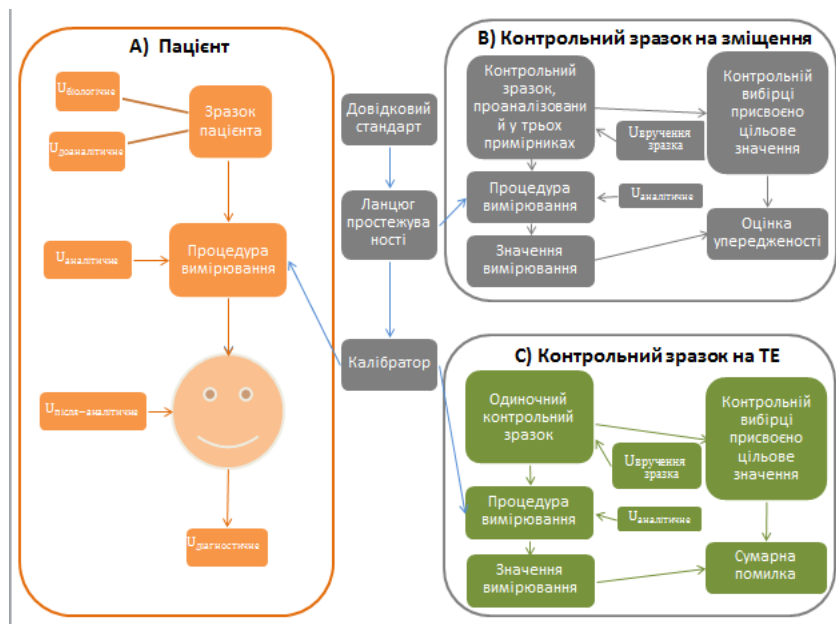


Рис. 1. Концептуальні відмінності між діагностичною невизначеністю та підходами до похибки вимірювання при оцінці (А) діагностичної невизначеності, (В) упередженість та (С) ТЕ в лабораторній медицині

Підходи до діагностичної невизначеності (А) забезпечують оцінку діагностичної невизначеності, яка врахує сукупну невизначеність (Рис.2).

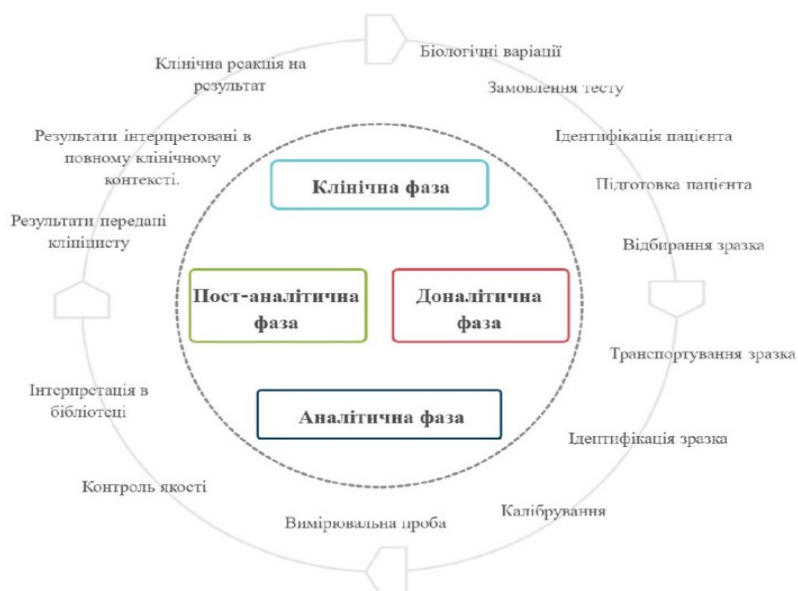


Рис. 2. Сукупна невизначеність

Ефекти усіх причин невизначеності у загальному ланцюжку випробувань при використанні лабораторних результатів для діагностики пацієнтів, включаючи біологічні варіації та аналітичні, аналітичні та постаналітичні варіації. Як LPU, наданий GUMом, так і просте додавання дисперсій відповідно до теореми Піфагора можна використовувати для оцінки діагностичної невизначеності.

Методи LPU мають свою основну силу в здатності боротися з численними та складними причинами невизначеності, але їх основною слабкістю є їх теоретична та практична складність та відсутність практичного впровадження в лабораторній медицині.

Підходи до похибки вимірювання (В, С) надати оцінки невизначеності методів вимірювання. Якщо основною метою програми зовнішнього контролю якості є те, щоб окремо визначити упередженість та неточність, підхід (В) є кращим. Зовнішня якість має контрольні програми, спрямовані на оцінку ТЕ, використовують одномісні вимірювання (С). Їх підтримують вичерпні теоретичні моделі, прийняті контролюючими органами та які мають широке практичне використання в лабораторній медицині. Біологічні, доаналітичні та постаналітичні варіації не є актуальними щодо моніторингу змінних у контрольних зразках, але ці варіації актуальні при роботі із зразками аналізів пацієнтів (Рис. 3).

Похибка вимірювання або просто похибка, є властивістю одного вимірювання - значення кількості мінус еталонне значення величини. Значення еталонної величини служить для змінної як сурогатне справжнє значення «в системі (або моделі)». Справжнє значення або контрольне значення вимірюваної величини в конкретному зразку показнику аналізу пацієнта не може бути відомим. Отже, справжнє значення результату одного пацієнта не може бути відомим. Ми не маємо впевненості у результаті, який повинен мати ймовірнісні термінів на основі частотних статистичних даних (довірчих інтервалів) для моделей помилок та баєсової статистики (щільності імовірності) для моделей невизначеності.

Серед міфів про методи вимірювань невизначеності є те, що вони вимагають щоб упередження було усунуте до того, як можуть бути зроблені розрахунки для оцінки невизначеності. Підходи до обчислення невизначеності стверджують, що упередження слід усунути при їх виявленні. Однак, якщо упередженості неможливо усунути, коли усунення упередженості має збільшити ризик, то загальна невизначеність вимірювання може бути оброблена як будь-який інший тип В невизначеності.



Рис. 3. Термінологія, що використовується у VIM3 для опису компонентів похибки та невизначеності вимірювання

Підходи до вимірювань невизначеності у лабораторній медицині стикаються з

декількома викликами:

1. Рівень знань з математики та передової статистики в лабораторній медицині, як правило, занадто низький, щоб повністю зрозуміти і застосувати закон поширення невизначеності (LPU), включаючи рівняння вимірювання, матриці коваріації, часткові похідної, розкладання Тейлора та байєсівські статистичні дані, які необхідні для повної реалізації підходів невизначеності.

2. Методи помилок / частоти використовуються у всіх лабораторіях і, як правило, добре зрозуміли. Їх використання при впровадженні Міжнародної організації зі стандартів - Стандартів (ISO) загалом добре приймаються органами з акредитації. Існує обмеження стимулу щодо для лабораторій, щоб залишати методи помилок на користь невизначеності методи з ЛПУ (Рис. 1). Термінологія, яка використовується у VIM3 для опису компонентів похибки та невизначеності вимірювань. Похибка вимірювання або просто похибка, є властивістю одного вимірювання:

- значення кількості мінус еталонне значення величини. Значення еталонної величини служить сурогатне справжнє значення "в системі (або моделі). справжнє значення або контрольне значення вимірюваної величини в конкретному зразку пацієнта не може бути відомим. Отже, справжнє значення результату аналізу одного пацієнта не може бути відомим. Впевненість у результаті повинна виражатися в ймовірнісних термінах на основі частотних статистичних даних (довірчих інтервалів) для моделей помилок та баєсової статистики (щільність ймовірності) для моделей невизначеності;

- рівень знань щодо біологічних, доаналітичних та постаналітичних змін у лабораторній медицині все ще значно відстає від знань про причини аналітичних варіацій, і це зменшує надію на те, що правильне використання невизначеності та застосування різноманітних методів покращать клінічне використання методів вимірювання.

Незважаючи на перешкоди, методи невизначеності були належним чином застосовані в лабораторній медицині.

3. Використання методу Монте-Карло

Моделювання з використанням методу Монте-Карло легше зрозуміти, ніж традиційні методи LPU. Це вимагає інформації про розподіл ймовірностей усіх факторів, що впливають на невизначеність.

Дисперсія в дослідженні, що використовується для оцінки біологічних змінних, звикла у моделі, що використовується щодо внесоку біологічної варіації та оцінці діагностичної невизначеності.

Преаналітична варіація оцінюється на основі квадратичної функції ймовірності, властивості якої оцінювання як невизначеність типу В. Аналітична варіація оцінюється як дисперсія повторюваності; всі результати вимірювань в лабораторній організації при вимірюванні однакових зразків для певної вимірюваної величини з використанням усіх систем вимірювання у різних умовах.

Постана алітична варіація оцінюється на основі квадратичної функції ймовірності, тобто властивості оцінюються як невизначеності типу В. Для аналізу діагностичної невизначеності одночасно з усіх функцій ймовірності беруть щонайменше 100 000 повторних зразків для оцінки розподілу ймовірностей.

4. Методи відновлення даних

Під «неповністю» або «некомплектністю» мається на увазі те, що деякі дані з

тих чи інших причин пропущені або відсутні у вихідному масиві даних. Традиційними причинами, що призводять до появи пропусків, є:

- неможливість їх отримання або обробки;
- спотворення або приховування інформації;
- всілякі поломки технічного обладнання;
- природні явища;
- економічні причини та інше

У підсумку на вхід програм аналізу зібраних даних надходять неповні відомості. З проблемою обробки пропусків у масивах даних доводиться стикатися при проведенні різноманітних технічних, соціологічних, економічних, астрономічних, біологічних, статистичних та ін. досліджень.

Західні дослідники такі дані з пропущеними значеннями називають "missing data" або "incomplete data" [7]. У [7] наведено результати аналізу методів, що застосовується у напрямку врахування невизначеності у даному напрямку (Рис.4).

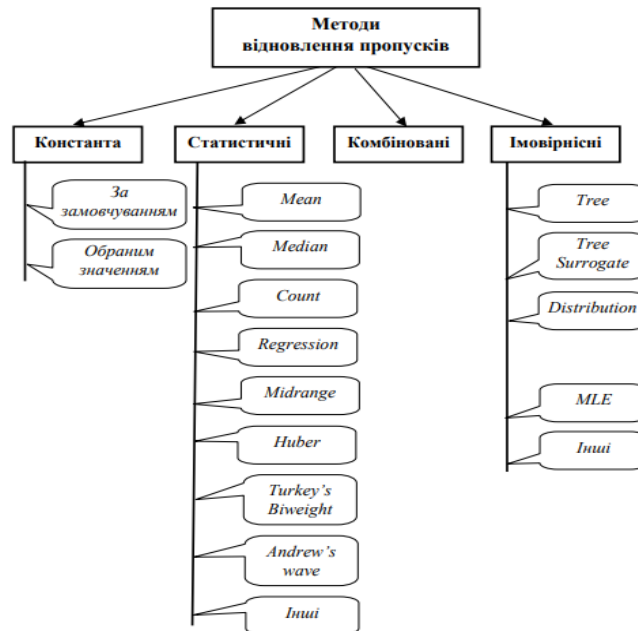


Рис. 4. Класифікація методів заповнення пропущених даних

5. Приклад реалізації алгоритму доповнення даних

Щодо аналізу пропущених даних в статистичному пакеті SPSS передбачена модель Missing Values [7,8], яка містить дві процедури:

1. Процедура «Аналіз пропущених значень» виконує три основні функції:

- Описує структуру пропущених даних. Де розташовані пропущені значення? Наскільки широку область вони охоплюють? Чи є тенденція до пропуску значень в декількох спостереженнях у парі змінних? Чи приймають дані крайні значення? Носять чи пропуски випадковий характер?

- Оцінює середні, середньоквадратичні відхилення, ковариації і кореляції для різних методів обробки пропущених значень: за списками, попарно, регресія або ОМП (максимізація очікувань). Попарний метод виводить також частоти повних пар спостережень.

- Виробляє включення (імпутацію) на місце пропущених значень оціночних

значень, використовуючи метод регресії або ОМП (максимізація очікувань).

2. Процедура «Множинної імпутація». Мета «множинної імпутації» - згенерувати можливі значення для пропущених значень, створивши кілька нібито повних наборів даних. Аналітичні процедури, що працюють з наборами даних багаторазового включення, генерують результати для кожного нібито повного набору даних плюс об'єднаний результат, який оцінює, якими були б результати, якби початковий набір даних був без пропущених значень. Ці об'єднані результати зазвичай точніше, ніж отримані за методами одного включення.

6. Реалізація програмного забезпечення

Розроблене програмне забезпечення у [7] (Додаток А) поMissingness засноване на описаних вище методах. Програма дозволяє створювати нові дані або ж відкривати і редагувати дані з файлу формату .txt з поділом знаком табуляції. Для роботи з даними необхідно:

1) Виберіть в меню:

Аналіз> Відновлення пропущених значень

2) У вікні «Вибір методів» відзначте методи одноразових включень, на основі яких буде будується агрегована оцінка за правилом Рубіна на стадії багаторазового включення. Додатково можна вивести на екран матрицю коефіцієнтів кореляції Пірсона і матрицю ковариаций.

7. Опис алгоритму програми

Після того як користувач вніс у програму дані і вибрав методи одноразових включень, програма на початку роботи знайде матрицю коваріації, матрицю коефіцієнтів кореляції за формулою Пірсона і 95% довірчі інтервали по кожному колонку. Матриця коефіцієнтів кореляції використовується для знаходження повнокомплектною вибірки з найбільшим коефіцієнтом кореляції з неполнокомплектною вибіркою, над якою проводиться аналіз, в іншому випадку (якщо немає повнокомплектних вибірок) програма знайде неполнокомплектную вибірку і заповнить її середнім за вибіркою для подальшого аналізу.

Даний алгоритм, в рамках програми, використовується в усіх процедурах одноразового включення за винятком заповнення середнього по вибірці. Матриця ковариации потрібна для початкових розрахунків в методах, які використовують рівняння регресії. Довірчі інтервали необхідні для подальшого аналізу включень.

При використанні методу програма працює з копією даних для збереження цілісності оригіналу і далі обробляє дані по їх опису.

При цьому дані з включеннями з одноразових методів і з багаторазового методу будуть автоматично збережені в файлі SaveFF.txt в корені програми.

Медичні дані більш схильні до невизначеності через наявність шуму в даних. Тому для точного діагнозу дуже важливо мати чисті медичні дані без зайвого шуму. Для вирішення цієї проблеми необхідно знати джерела шуму в медичних даних. На підставі медичних даних, отриманих лікарем, призначається діагноз захворювання і план лікування. Отже, в сфері охорони здоров'я зростає невизначеність, а знання для вирішення цих проблем обмежені. У нас мало знань про оптимальні методи лікування, оскільки в медичній науці є багато джерел невизначеності. Наші результати показують, що є кілька проблем, які необхідно вирішити при обробці невизначеності в медичних необроблених даних і нових моделях. Отже, в сфері охорони здоров'я зростає невизначеність, а знання для вирішення цих проблем обмежені. У нас мало знань про оптимальні методи лікування, оскільки в медичній

науці є багато джерел невизначеності. Наші результати показують, що є кілька проблем, які необхідно вирішити при обробці невизначеності в медичних необроблених даних і нових моделях. У цій роботі ми підсумовували різні методи, використовувані для вирішення цієї проблеми. У даний час застосування нових методів глибокого навчання для вирішення таких невизначеностей значно розширилося. і план лікування призначається. Отже, в сфері охорони здоров'я зростає невизначеність, а знання для вирішення цих проблем обмежені. У нас мало знань про оптимальні методи лікування, оскільки в медичній науці є багато джерел невизначеності. У цій роботі ми підсумовували різні методи, використовувані для вирішення цієї проблеми. В даний час застосування нових методів глибокого навчання для вирішення таких невизначеностей значно розширилося. Наші результати показують, що є кілька проблем, які необхідно вирішити при обробці невизначеності в медичних необроблених даних і нових моделях.

В процесі створення програмного продукту була розглянута реалізація функціоналу пакету SPSS, що працює з пропусками в даних. Був реалізований програмний продукт, який застосовує методи відновлення пропущених даних.

а) Програмний продукт використовує популярні методи відновлення даних для заповнення пропусків в некомплектних даних.

б) Взаємодія з користувачем програмного продукту виробляється через графічний інтерфейс (UI).

А також, було проведено обчислювальний експеримент, в результаті якого зроблений порівняльний аналіз ефективності застосування багаторазової вставки за правилом Рубіна і одноразових методів заповнення пропусків. Після оцінки ефективності підходів можна зробити висновок, що Multiple imputation дійсно дає хороші результати відновлення. За реальними чисельним значенням вони можуть відрізнятися, але тим не менше вони будуть входити в 95% довірчий інтервал, а метод Jack Knife на кроці аналізу в якійсь мірі здатний поліпшити результат, але його має сенс використовувати при некомплектності від 20% і вище, оскільки при менших втратах ефекту не видно.

8. Обчислювальний експеримент

У роботі [7] було проведено дослідження основі медичних повнокомплектних даних щодо дослідження з 15 спостережень і 10 ознак обчислювальний експеримент зі згенерували повністю випадковими пропусками (MCAR) як зображено на (Рис.5), що становлять 2.8% від усіх даних. Масиви даних оброблялися описаними ранне методами відновлення пропущених значень: підстановка середнього за вибіркою, підстановка медіани, метод Hot Deck, модель максимальної правдоподібності (EM алгоритм), парна регресія, стохастична регресія і метод multiple imputation Рубіна з використанням Jack Knife.

Скористаємося пакетом SPSS щоб знайти довірчий інтервал для кожного з параметрів з пропусками.

Таблиця 1

Довірчі інтервали при 6 пропусках

		95% довірчий інтервал	
		Нижня межа	Верхня межа
Середнє	VAR2	5.15	6.025
	VAR3	43.975	56.55
	VAR6	18.2	21.775

Таблиця 2

Відхилення при 15 пропусках

Відхилення від реального значення (%)	Метод							
	Підстановка середнього за вибіркою	Підстановка медіани	Hot Deck	EM алгоритм	Парна регресія	Стохастична регресія	Multiple imputation	Jack Knife
0-20	12	12	6	11	7	7	11	13
20-50	2	3	9	5	8	4	4	2
50-100	1	0	0	0	0	4	0	0
100+	0	0	0	0	0	0	0	0

Отримані результати підтвердили можливості використання запропонованого методу додавання пропущених значень клінічних даних.

11. Висновок

Наведений вище аналіз [7] дає змогу продовжити у цьому напрямку дослідження. Отже, використовуючи машинне навчання [1] можливо підвищити якість прогнозування та прийняття рішень за рахунок розвитку **нового методу відновлення пропущеного значень у наборі даних при використанні іммуносігнату** [7] для прогнозування розвитку COVID-19.

Основної метою застосування технології іммуносігнатур є підтримка прийняття рішень в постановці діагнозу, що в термінології аналізу даних вирішується як задача класифікації. У зв'язку з цим для оцінки якості відновлення пропущених значень щодо нашої мети дослідження для умов COVID-19 використовує результати класифікації за даними після процедури відновлення. Цей метод аналізу набору даних вже був застосований раніше у дослідженні [9], де встановлено, що алгоритм випадкового лісу показує високі результати класифікації, у зв'язку з чим його використання в якості класифікатора є цілком доцільним.

З огляду на те, коли набір даних має збалансовану вибірку, доцільно змінити метрику класифікації accuracy. Треба щоб набір даних для умов COVID-19 мав велику кількість ознак, що негативно позначається на часі проведення обчислень. У зв'язку з цим для економії часу для великого набору даних треба вибрати за допомогою t-критерія Стьюдента відібрати тільки 120 найбільш інформативних ознак [7]. Для проведення порівняння ефективності розглянутих методів необхідно сформулювати вибірку з параметрами, для умов COVID-19 маємо змінні в широких межах. У цьому разі доцільно реалізувати усунення невизначеності можливо шляхом додавання пропущених значень штучно. При багатокроковому аналізі можлива поява пропущених значень на будь-якій стадії COVID-19, майже без будь-якої закономірності. зважаючи на це неможливо підібрати метод створення пропущених значень, що імітують всі можливі сценарії. У зв'язку з цим у даному дослідженні можливо рекомендувати використання методу повністю випадкового створення пропущених значень (missing completely at random) [7]. У підсумку цього дослідження рекомендується при виборі методів необхідно виконати наступні кроки:

1. Створено кілька наборів даних з різною кількістю пропущених значень.

2. Відновити пропущені значення у створених наборах даних по черзі із кожним з розглянутих методів.

3. Провести класифікацію за даними кожного з відновленого набору.

4. Повторити кроки 1-3 тридцять разів.

5. Обчислити середні значення точності класифікації, які були проведені з відновленими наборами, для кожного з розглянутих методів для різної кількості пропущених значень.

6. Оцінити отримані результати.

Усю роботу для прогнозування в умовах COVID-19 рекомендується виконати з використанням мови програмування R і програмних бібліотек з публічного сховища CRAN (Табл. 3).

Таблиця 3

Використовувані пакети і їх параметри

Методи	Пакет / параметри налаштування
Просте відновлення	caret / medianImpute
Метод k-зважених найближчих сусідів	wNNSEL
Лінійна регресія	ice / norm.predict
Випадковий ліс	missForest / 100 дерев максимум ітерацій - 10
Лінійний градієнтний бустінг	xgboost / максимум ітерацій - 10
Метод найшвидшого бустінг дерев	максимум раундів - 100, eta = 0.3

Отже у підсумку проведеного дослідження було вибрано метод, який найбільше підходить до умов COVID-19, тому що у цьому разі для прогнозування можливо врахувати ймовірну невизначеність предметної області.

Список використаних джерел:

1. Mobile expert system for diagnostic human state in emergency situations./ Kuzomin, O., Dudka, O., Vasylenko, O., Shylo, R., Lyashenko, V.// International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(4), pp. 6484-6489, 334
2. Литтл Р.Дж.А., Статистический анализ данных с пропусками. Финансы и статистика / Литтл Р.Дж.А., Рубин Д.Б – Москва, 1991, 336 стр.
3. V.S. Petrovichev, A.V. Melekhov, M.A. Saifullin, I. G. Nikitin / Computed tomography for coronavirus infection: differential diagnosis based on clinical examples // Archives of Internal Medicine. Original articles. No. 5.2020. S. 357-371.
4. Using of ontologies for building databases and knowledge bases for consequences management of emergency/Kuzomin, O., Dudka, O., Vasylenko, O., Radchenko, V., Lyashenko, V.//International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(4), pp. 5040-5045.
5. Розробка багатоагентних структур для вирішення проблем медичної системи діагностування / О.Я. Кузьомін, О.О. Василенко, Свістунов І.О. // Радіоелектроніка та інформатика. 2020. №2. С. 47 – 54.
6. Koshechkin A.A., Andryushchenko V.S., Zamyatin A.V. A new method to missing value imputation for immunosignature data. Sovremennye tehnologii v medicine 2019; 11(2): 19–24, <https://doi.org/10.17691/stm2019.11.2>.
7. Andryushchenko VS, Uglov AS, Zamyatin AVStatistical classification of immunosignatures under significantreduction of the feature space dimensions for early diagnosisof diseases. Sovremennye tehnologii v medicine 2018; 10 (3):14-20, <https://doi.org/10.17691/stm 2018.10.3.2>.

НАУКОВЕ ВИДАННЯ

МАТЕРІАЛИ IV МІЖНАРОДНОЇ
СТУДЕНТСЬКОЇ НАУКОВОЇ КОНФЕРЕНЦІЇ

**«НАУКА СЬОГОДЕННЯ:
ВІД ДОСЛІДЖЕНЬ ДО
СТРАТЕГІЧНИХ РІШЕНЬ»**

17 червня 2022 рік • м. Івано-Франківськ, Україна

Українською та англійською мовами

*Всі матеріали пройшли перевірку на плагіат та експертизу за формальними ознаками
(форматування, стиль мови, оформлення цитувань та списку використаних джерел).
За точність викладеного матеріалу відповідальність несуть автори та їх наукові керівники.
Організаційний комітет не завжди поділяє позицію авторів.*

Підписано до друку 17.06.2022.

Папір офсетний. Цифровий друк. Формат 60×84/16.

Гарнітура Times New Roman, Poiret One та Arial.

Умовно-друк. арк. 21,39. Замовлення № 321.

Тираж: 50 екземплярів. Віддруковано з готового оригінал-макету.

Контактна інформація організаційного комітету:

Громадська організація «Молодіжна наукова ліга»
21037, Україна, м. Вінниця, вул. Зодчих, 40, офіс 103
Телефони: +38 098 1948380; +38 098 1956755
E-mail: info@liga.science

Видавець: ГО «Європейська наукова платформа».
21037, Україна, м. Вінниця, вул. Зодчих, 18, офіс 81. E-mail: info@ukrlogos.in.ua
Свідоцтво суб'єкта видавничої справи: ДК № 7172 від 21.10.2020.