Національна академія наук України

Інститут кібернетики імені В.М. Глушкова

Кваліфікаційна наукова праця на правах рукопису

Маланін Владислав Юрійович

УДК 004.89:314.4:614.2

ДИСЕРТАЦІЯ

РОЗРОБЛЕННЯ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ ТРИВАЛОСТІ ЖИТТЯ НА ОСНОВІ ПОЄДНАННЯ ВПЛИВУ ДЕМОГРАФІЧНИХ, СОЦІАЛЬНИХ ТА МЕДИЧНИХ ФАКТОРІВ

122 — «Комп’ютерні науки»

Галузь знань 12 — «Інформаційні технології»

Подається на здобуття наукового ступеня доктора філософії.

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

Маланін В.Ю.

Наукові керівники Тульчинський Вадим Григорович

д.ф.-м.н., с.н.с.

Чайковський Ілля Анатолійович

к.м.н., с.н.с.

Київ — 2025

АНОТАЦІЯ

Маланін В.Ю. Розроблення технології прогнозування тривалості життя на основі поєднання впливу демографічних, соціальних та медичних факторів. — Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки». — Інститут кібернетики імені В.М. Глушкова Національної академії наук України, Київ. — 2025.

У дисертаційній роботі вирішено актуальні науково-прикладні завдання створення ефективної інтегрованої технології прогнозування очікуваної тривалості життя населення України. Проведено аналіз існуючих методів прогнозування очікуваної тривалості життя, визначено їх переваги й недоліки. Обґрунтовано необхідність інтегрованого підходу, що охоплює демографічні, соціальні та медичні фактори, які найбільш повно відображають особливості здоров'я та умов життя українського населення.

В рамках розроблення інтегрованої технології було створено чотири математичні моделі прогнозування очікуваної тривалості життя. Дві моделі (LC19 та LC20) є детерміністичними, які базуються відповідно на офіційних статистичних даних ВООЗ за 2019 рік та Держстату України за 2020 рік. Інші дві моделі (LCR та LCX) створено з використанням сучасних алгоритмів машинного навчання — Random Forest та XGBoost відповідно, навчання яких було проведено на основі даних, отриманих у результаті проходження понад 260 тисяч персоналізованих анкет українськими користувачами. Коефіцієнти впливу факторів були визначені на основі сучасного досвіду прогнозування очікуваної тривалості життя з додатковим включенням факторів, специфічних для українського населення, зокрема в умовах війни.

Запропоновано алгоритм адаптивного оновлення прогнозів очікуваної тривалості життя, який дозволяє динамічно реагувати на зміну індивідуальних показників здоров'я користувачів та соціальних умов.

Вперше реалізовано технологію інтеграції цифрових медичних даних з платформ HealthKit та Google Fit у процес персоналізованого прогнозування очікуваної тривалості життя. Запропонований підхід дозволяє автоматично враховувати зміни медичних показників користувача в мобільних додатках, суттєво підвищуючи точність прогнозів.

Створено веб-застосунок «Калькулятор життя» з унікальним адаптивним інтерфейсом, високою швидкістю роботи, відповідністю міжнародним стандартам безпеки даних, який підтримує дві мови (українську та англійську). Веб-застосунок забезпечує швидке проходження персоналізованих анкет (середній час проходження — близько 4 хвилин 15 секунд), адаптивність результатів прогнозу при зміні відповідей користувачем, а також потенційну можливість інтеграції та аналізу даних електрокардіограм (ЕКГ).

Порівняльний аналіз результатів апробації показав, що всі чотири розроблені моделі, особливо на основі алгоритмів машинного навчання, мають значно кращі показники точності прогнозування порівняно з міжнародними аналогами (MPoRT, GBD, ONS), демонструючи найнижчі показники похибок (RMSE, MAE, MSE, MAPE) та достатньо високий C-index (0,707 для LCX і 0,705 для LCR). Ці моделі підтвердили свою високу стабільність та практичну придатність для персоналізованого прогнозування очікуваної тривалості життя українського населення.

Результати роботи успішно впроваджено у діяльність Центру громадського здоров’я МОЗ України. Оцінка користувацького досвіду продемонструвала високу зручність та ефективність застосунку, підтверджену позитивними відгуками та високим рівнем завершення користувацьких опитувань.

Отримані результати мають вагоме наукове та практичне значення, сприяючи розвитку комп’ютерних технологій у сфері охорони здоров’я, покращенню управління здоров’ям населення та зменшенню ризиків передчасної смертності.

Ключові слова: прогнозування тривалості життя, інтегрована модель, демографічні фактори, соціальні фактори, медичні фактори, машинне навчання, штучний інтелект, веб-застосунок, цифрове здоров'я, медична кібернетика, програмне забезпечення, популяційна норма, інформаційні технології, українська популяція.

ANNOTATION

Malanin V. Development of a technology for predicting life expectancy based on the combined influence of demographic, social and medical factors. — Qualifying scientific work as a manuscript.

Dissertation for the degree of Doctor of Philosophy in the specialty 122 "Computer science". — V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv. — 2025.

The dissertation addresses an important scientific and applied task: developing an effective integrated technology for predicting life expectancy in the Ukrainian population. An analysis of existing life expectancy forecasting methods was conducted, identifying their strengths and weaknesses. The necessity of an integrated approach covering demographic, social, and medical factors, which most comprehensively reflect the specific health conditions and living conditions of the Ukrainian population, was justified.

As part of developing the integrated technology, four mathematical models were created to predict life expectancy. Two deterministic models (LC19 and LC20) are based, respectively, on official statistical data from the WHO for 2019 and the State Statistics Service of Ukraine for 2020. The other two models (LCR and LCX) were developed using modern machine learning algorithms, Random Forest and XGBoost, respectively, trained on data obtained from over 260 thousand personalized questionnaires completed by Ukrainian users. Influence coefficients for the factors were determined based on modern experience in life expectancy forecasting, supplemented by specific factors relevant to the Ukrainian population, particularly under wartime conditions.

An adaptive algorithm for updating life expectancy predictions was proposed, enabling dynamic responses to changes in individual health indicators and social conditions.

For the first time, technology was implemented to integrate digital health data from the HealthKit and Google Fit platforms into personalized life expectancy forecasting. This approach automatically accounts for changes in user medical indicators within mobile applications, significantly improving prediction accuracy.

A web application, "Life Calculator," was created featuring a unique adaptive interface, high performance, adherence to international data security standards, and bilingual support (Ukrainian and English). The application ensures rapid completion of personalized questionnaires (average completion time about 4 minutes and 15 seconds), adaptive updating of prediction results as user responses change, and potential integration and analysis of electrocardiogram (ECG) data.

Comparative analysis of test results showed that all four developed models, particularly those based on machine learning algorithms, demonstrated significantly better accuracy compared to international counterparts (MPoRT, GBD, ONS). They showed the lowest error rates (RMSE, MAE, MSE, MAPE) and a sufficiently high C-index (0.707 for LCX and 0.705 for LCR). These models proved their high stability and practical applicability for personalized life expectancy forecasting in the Ukrainian population.

The research results have been successfully implemented at the Public Health Center of the Ministry of Health of Ukraine. User experience assessments demonstrated high convenience and effectiveness of the application, confirmed by positive feedback and a high rate of questionnaire completion.

The obtained results have significant scientific and practical implications, contributing to the advancement of computer technologies in healthcare, improving population health management, and reducing risks of premature mortality.

Keywords: life expectancy, integrated model, demographic factors, social factors, medical factors, machine learning, artificial intelligence, web application, digital health, mathematical model, population norm, medical cybernetics, Ukrainian population.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ 9

ВСТУП 11

1 ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ ПРОГНОЗУВАННЯ ТРИВАЛОСТІ ЖИТТЯ 22

1.1 Аналіз підходів і моделей прогнозування 22

1.1.1 Аналіз традиційних демографічних моделей (Lee-Carter, Gompertz-Makeham) 22

1.1.2 Аналіз сучасних цифрових та інтегрованих моделей (MPoRT, ONS, GBD) 26

1.1.3 Аналіз методів машинного навчання та нейронних мереж для прогнозування 30

1.1.4 Порівняльний аналіз критеріїв оцінювання точності прогнозних моделей 46

1.1.5 Аналіз інтеграції цифрових медичних та персональних даних .. 56

1.2 Концептуальні основи інтегрованої моделі прогнозування 61

1.3 Вибір і формалізація демографічних, соціальних та медичних факторів 64

1.4 Математична модель та алгоритми прогнозування 73

1.4.1 Розробка математичної моделі прогнозування з використанням коефіцієнтів впливу факторів на ймовірність виживання 73

1.4.2 Розробка алгоритму адаптивного прогнозування з урахуванням специфіки українського населення та врахування особливих категорій населення у моделі 79

1.5 Висновки до першого розділу 83

2 РОЗРОБКА ТА ПРАКТИЧНА РЕАЛІЗАЦІЯ ВЕБ-ЗАСТОСУНКУ ДЛЯ ПРОГНОЗУВАННЯ ТРИВАЛОСТІ ЖИТТЯ 84

2.1 Архітектура веб-застосунку «Калькулятор життя» 84

2.1.1 Бізнес-рівень: модель взаємодії користувача 85

2.1.2 Рівень застосунків: фронтенд та бекенд-логіка 87

2.1.3 Технологічний рівень: серверна інфраструктура, захист даних та інформаційна безпека 90

2.1.4 Компресія та обробка зображень інтерфейсу за допомогою ШІ 94

2.2 Інтерфейсні рішення для ефективного збирання та аналізу даних 95

2.3 Інтеграція зовнішніх цифрових медичних і соціальних даних 101

2.4 Приклад застосування розробленого інструменту 104

2.5 Висновки до другого розділу 106

3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ, ВАЛІДАЦІЯ ТА ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ 107

3.1 Методика та аналіз результатів експериментального дослідження 107

3.1.1 Методика проведення експерименту 107

3.1.2 Аналіз точності прогнозів на реальних даних українського населення 112

3.1.3 Аналіз впливу факторів на тривалість життя за результатами експерименту 117

3.2 Оцінювання користувацького досвіду та зручності використання веб-застосунку.....119

3.3 Перспективні напрями розвитку технології прогнозування 121

3.3.1 Напрями вдосконалення розробленої моделі та методів..... 121

3.3.2 Перспективи інтеграції з медичними та державними сервісами 124

3.3.3 Адаптація технології до міжнародних стандартів і використання в інших країнах 126

3.4 Висновки до третього розділу..... 128

ЗАГАЛЬНІ ВИСНОВКИ 130

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ 133

ДОДАТКИ..... 146

Скорочення

ШІ — штучний інтелект

ML — машинне навчання (Machine Learning)

NN — нейронна мережа (Neural Network)

CNN — згорткова нейронна мережа (Convolutional Neural Network)

RNN — рекурентна нейронна мережа (Recurrent Neural Network)

LSTM — мережа довгої короткочасної пам'яті (Long Short-Term Memory)

GRU — блок з вентилями (Gated Recurrent Unit)

MLP — багатошаровий персептрон (Multilayer Perceptron)

API — інтерфейс прикладного програмування (Application Programming Interface)

SVD — сингулярний розклад (Singular Value Decomposition)

ARIMA — інтегрована модель авторегресії і ковзного середнього (Autoregressive Integrated Moving Average)

ECG, ЕКГ — електрокардіограма

GBD — глобальний тягар хвороб (Global Burden of Disease)

DALY — роки життя, скориговані на непрацездатність (Disability-Adjusted Life Years)

YLL — втрачені роки життя (Years of Life Lost)

YLD — роки, прожиті з інвалідністю (Years Lived with Disability)

UX/UI — дизайн користувацького досвіду/інтерфейсу (User Experience/User Interface)

ROC-AUC — площа під ROC-кривою (Receiver Operating Characteristic — Area Under Curve)

RMSE — корінь середньоквадратичної похибки (Root Mean Squared Error)

MAE — середня абсолютна похибка (Mean Absolute Error)

MSE — середньоквадратична похибка (Mean Squared Error)

MAPE — середня абсолютна процентна похибка (Mean Absolute Percentage Error)

TP — істинно позитивні (True Positive)

FP — хибнопозитивні (False Positive)

TN — істинно негативні (True Negative)

FN — хибнонегативні (False Negative)

ВСТУП

Актуальність. Дисертація присвячена проблемі персоналізованого прогнозування тривалості життя українського населення на основі інтеграції медичних і соціальних даних, що створює передумови для більш ефективного управління здоров'ям населення, формування індивідуальних профілактичних стратегій, а також підвищення загальної якості життя та довголіття громадян. Висока актуальність та соціальна значущість розробленої технології прогнозування тривалості життя підтверджується значним суспільним резонансом після її публічного релізу.

Основи персоналізованого прогнозування тривалості життя закладені такими вченими, як Р.Д. Лі, Л.Р. Картер, Б. Гомперц, У. Мейкхам, Дж.Р. Уілмот, Х. Бут, Л. Тіккл, Л.А. і Н.С. Гаврилови, Х.Л. Шанг, М. Мурано, С. Левантесі, А. Нігрі, М.С. Фанг, Г.У. Петерс, П.В. Шевченко та іншими. Ними було розроблено концепцію законів смертності та відомі демографічні моделі, починаючи з моделей Лі-Картера і Гомперца-Мейкхама. Подальші дослідження з'ясували певні обмеження традиційних моделей спрямованих на аналіз загальних тенденцій, але не придатних для прогнозування тривалості життя окремих осіб з урахуванням їх індивідуальних характеристик. Тому сьогодні особливий інтерес становлять інтегровані моделі, що поєднують демографічні, медичні й поведінкові фактори та використовують сучасні алгоритми машинного навчання й штучного інтелекту. Серед таких сучасних моделей, що побудовані на основі обчислення великих даних можна відмітити канадську модель Інструмент популяційного ризику смертності (Mortality Population Risk Tool, MPoRT) побудовану у рамках ініціативи Big Life,

американську модель Глобальний тягар хвороб (Global Burden of Disease GBD) Інституту показників та оцінки здоров’я (Institute for Health Metrics and Evaluation, IHME) при Університеті Вашингтону, та британську модель Управління національної статистики (Office for National Statistics, ONS).

Для України додаткову актуальність теми обумовлюють сучасні виклики: війна, окупація частини території, пандемія COVID-19, та викликані ними екологічні та соціальні кризи, які безпосередньо впливають на стан здоров’я, стресові умови життя, та відповідно, змінюють структуру факторів ризику для населення. В цих умовах важливо мати надійні й оперативні технології прогнозування, здатні швидко реагувати на динамічні зміни ситуації, забезпечуючи громадян і державні установи інструментами для прийняття своєчасних і обґрунтованих рішень.

Важливим додатковим аспектом актуальності роботи є її значення для індивідуальних користувачів. Персоналізовані прогнози тривалості життя надають людині потужну мотивацію для ведення більш здорового способу життя. Наявність чітких, науково обґрунтованих прогнозів щодо потенційних наслідків поточних звичок і поведінки стимулює усвідомлене ставлення до власного здоров'я, сприяє прийняттю більш здорових рішень та підвищує готовність до позитивних змін. Таким чином, крім суспільного значення, ця робота має яскраво виражену практичну користь на рівні окремого індивіда, сприяючи формуванню відповідального ставлення до власного життя та здоров'я.

Запуск веб-застосунку «Калькулятор життя» в рамках наукового співробітництва Інституту кібернетики імені В.М. Глушкова Національної академії наук України з Державною установою «Центр громадського здоров’я МОЗ України» викликав широке зацікавлення у суспільстві та привернув увагу провідних українських медіа, зокрема новинних порталів і соціальних мереж. Такий активний інтерес свідчить про важливість та своєчасність тематики дослідження, а також наголошує на практичній корисності та попиті на інструменти персоналізованого прогнозування у сфері охорони громадського здоров’я.

Дисертаційна робота виконувалась у межах наукової співпраці Інституту кібернетики імені В.М. Глушкова Національної академії наук України та Державною установою «Центр громадського здоров'я МОЗ України», відповідно до пріоритетних напрямів розвитку цифровізації охорони здоров'я України, що визначені Кабінетом Міністрів України, а також з міжнародними проєктами з впровадження цифрових медичних технологій. Результати досліджень було впроваджено у межах розробки та реалізації державної цифрової ініціативи Центру громадського здоров'я МОЗ України — веб-застосунку «Калькулятор життя», що є частиною загальнодержавних заходів щодо підвищення тривалості та якості життя громадян України.

Зв'язок роботи з науковими програмами, планами, темами. Наукова спрямованість дисертації відповідає пріоритетам науково-технічної політики України, що визначені постановою Кабінету Міністрів України від 30 квітня 2024 р. № 476, а саме «Інформаційні технології в медицині», «Системи штучного інтелекту», «Інтелектуальні інтерактивні інформаційно-аналітичні системи. Інтегровані системи баз даних та знань. Національні інформаційні ресурси». Робота виконана як складова наукової співпраці Інституту кібернетики імені В.М. Глушкова НАН України з Державною установою «Центр громадського здоров'я МОЗ України» в рамках наступних науково-дослідних тем Інституту кібернетики імені В.М. Глушкова НАН України:

- ВК.145.37.20 «Розвиток грід-інфраструктури ресурсного центру СКІТ і хмарних сервісів для вирішення складних міждисциплінарних задач НАН України» (№ держреєстрації 0120U103329, 2020).
- ВФ.146.20 «Розробити методи взаємної адаптації архітектури обчислювального комплексу та властивостей моделей і програм» (№ держреєстрації 0124U001601, 2020-2025).
- ВП.145.34 «Розробити і дослідити методи розподілених обчислень для задач комп'ютерного моделювання і машинного навчання» (№ держреєстрації 0122U001515, 2022-2024).

- ВП.145.46 «Розробити моделі та методи машинного навчання у прикладних задачах інформатизації» (№ держреєстрації 0125U001129, 2025).

Мета роботи — оцінити очікувану тривалість життя населення України на основі розроблення інтегрованої технології персоналізованого прогнозування тривалості життя, що враховує демографічні, соціальні та медичні фактори з використанням сучасних алгоритмів машинного навчання.

Для досягнення поставленої мети визначені такі **наукові завдання**:

1. Проаналізувати сучасні методи та підходи до прогнозування очікуваної тривалості життя.
2. Визначити та формалізувати основні демографічні, соціальні та медичні фактори для побудови інтегрованої моделі прогнозування очікуваної тривалості життя.
3. Розробити модель персоналізованого прогнозування тривалості життя на основі інтеграції різних типів даних та сучасних алгоритмів машинного навчання.
4. Створити алгоритми адаптивного прогнозування тривалості життя з урахуванням специфіки українського населення, зокрема в умовах війни.
5. Реалізувати розроблену модель у вигляді практичного веб-застосунку та провести експериментальне дослідження на реальних даних українського населення.
6. Оцінити ефективність, точність та зручність запропонованої технології, а також визначити перспективні напрями її розвитку.

Об’єкт дослідження — прогнозування очікуваної тривалості життя українського населення. **Предметом дослідження** є методи, алгоритми й інтегровані технології прогнозування тривалості життя українського населення на основі комплексного аналізу медичних, соціальних та демографічних факторів.

Методи дослідження. У роботі використано наступні методи:

- Системний аналіз – для аналізу та узагальнення існуючих моделей і методів прогнозування тривалості життя, що дозволило визначити напрями вдосконалення прогнозних технологій (завдання 1).
- Статистичні методи (кореляційний аналіз, регресійний аналіз) – для виявлення значущості та формалізації впливу демографічних, соціальних та медичних факторів на тривалість життя (завдання 2).
- Методи машинного навчання (Random Forest, XGBoost) – для створення інтегрованої математичної моделі прогнозування тривалості життя на основі великої кількості персоналізованих даних (завдання 3).
- Методи адаптивного моделювання – для реалізації алгоритму оновлення персонального прогнозу тривалості життя при зміні вхідних даних, що важливо в умовах динамічних змін (завдання 4).
- UX/UI-дизайн, емпіричний аналіз користувацького досвіду – для створення зручного й ефективного інтерфейсу веб-застосунку «Калькулятор життя» (завдання 5).
- Методи експериментального дослідження та порівняльного аналізу – для оцінювання ефективності, точності та практичної цінності розробленої технології прогнозування, порівняння отриманих результатів з міжнародними аналогами та визначення перспективних напрямів її подальшого розвитку (завдання 6).

Наукова новизна отриманих результатів полягає у тому, що було вперше розроблено інтегровану модель персоналізованого прогнозування тривалості життя для української популяції на основі комплексного аналізу індивідуальних демографічних, соціальних та медичних факторів із застосуванням алгоритмів машинного навчання. Розроблена модель відрізняється врахуванням групової специфіки українського населення (регіонів, військових та цивільних осіб, які пережили бойові дії, ВПО тощо) та адаптивного оновлення прогнозів за умови отримання нових даних.

Практичне значення. Розроблено і впроваджено веб-застосунок «Калькулятор життя», який використовується Центром громадського здоров’я МОЗ України як інструмент персоналізованого прогнозування очікуваної тривалості життя населення та ефективного управління ризиками здоров’я громадян. Експериментальна апробація веб-застосунку, проведена на великій вибірці користувачів (понад 200 тис. осіб), підтвердила його високу точність та ефективність у реальних умовах.

Завдяки запропонованому алгоритму адаптивного оновлення прогнозів тривалості життя, користувачі можуть отримувати актуальні персоналізовані рекомендації щодо способу життя, що сприяє зменшенню ризиків передчасної смертності, підвищенню ефективності заходів профілактики захворювань та покращенню загальної якості життя населення.

Впровадження технології автоматичної інтеграції медичних даних з цифрових платформ в мобільних додатках забезпечує додаткову точність персоналізованих прогнозів тривалості життя, дозволяючи максимально точно відстежувати зміни здоров’я користувачів у режимі реального часу. У роботі запропоновано нову концепцію прогнозування тривалості життя з застосуванням принципу динамічної корекції коефіцієнтів впливу факторів у процесі накопичення нових статистичних даних.

Застосування розробленої технології отримало значний суспільний резонанс, підтверджений численними згадками у медіа, новинних виданнях та соціальних мережах. Висока зацікавленість суспільства свідчить про значний потенціал для масштабного впровадження технології у державні медичні сервіси та інтеграції в національні програми охорони громадського здоров'я, особливо в умовах війни.

Теоретичне значення. Отримані результати розширюють існуючі уявлення щодо взаємозв’язку та взаємного впливу медичних, соціальних і демографічних факторів на тривалість життя людини. Розроблена інтегрована модель і технологія прогнозування є суттєвим внеском у розвиток сучасних методів системного аналізу та персоналізованого моделювання. Запропоновані алгоритмічні рішення щодо

динамічного оновлення прогнозів і автоматичної інтеграції цифрових даних є теоретичною основою для подальших досліджень у сфері цифровізації медицини та громадського здоров'я.

Особистий внесок здобувача. Особистий внесок здобувача полягає у формулюванні та постановці наукового завдання, аналізі сучасних підходів до прогнозування тривалості життя, розробленні й реалізації інтегрованої математичної моделі та алгоритмів персоналізованого прогнозування тривалості життя для українського населення, а також створенні веб-застосунку «Калькулятор життя». Автор особисто здійснив експериментальну перевірку моделі, аналіз та інтерпретацію отриманих результатів, а також підготував наукові публікації, в яких відображено результати дослідження. Усі основні наукові положення та результати, винесені на захист, отримані автором самостійно.

Апробація результатів дисертації. Результати дослідження були високо оцінені Державною установою «Центр громадського здоров'я МОЗ України». У подяці (Додаток Г) особливо відзначено актуальність розробленого калькулятора очікуваної тривалості життя, високий рівень роботи, а також важливість внеску в покращення громадського здоров'я. Результати роботи були освітлені провідними ЗМІ України (Суспільне, Українська Правда та інші) та на сайті НАН України в окремому повідомленні «Що таке очікувана тривалість життя і як її рахують». Положення дисертації доповідалися та обговорювалися на міжнародних наукових конференціях, наукових семінарах та експертних панелях, зокрема на:

1. IX International Congress on Information and Communication Technology (ICICT 2024), London, 19.02.2024. Доповідь: «Combined ECG Analysis Approach Based on Detection of Deviation from the Personal or Population Norm Using Wearable Devices» та участь у дискусії.
2. Forbes Technology Council Expert Panel «Emerging Wearable Health Technologies Transforming Remote Care», 11.06.2024. Доповідь в експертній панелі.
3. Науково-практичний семінар «Штучний інтелект. Наука. Бізнес». ДУ «Інститут ринку і економіко-екологічних досліджень НАН України», Одеса,

22.10.2024. Доповідь «Комп’ютерний зір та штучний інтелект: трансформація різних галузей через інновації» та участь у дискусії.

Публікації. За матеріалами дисертаційного дослідження опубліковано 13 наукових праць. При перерахунку згідно постанові Кабінету Міністрів України від 12 січня 2022 р. № 44 (Із змінами, внесеними згідно з постановою Кабінету Міністрів України № 507 від 03.05.2024) кількість публікацій дорівнює 5. З них:

Стаття у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України за спеціальністю 122

1. Маланін В. Ю., Чайковський І. А. Розробка математичної моделі для персоналізованого оцінювання очікуваної тривалості життя в Україні // Кібернетика та комп'ютерні технології. 2025. 2. 5—6. DOI: 10.34229/2707-451X.25.2.4.

Особистий внесок здобувача: Запропонував концепцію, розробив методологію та здійснив формальний аналіз математичної моделі персоналізованого оцінювання очікуваної тривалості життя в Україні. Виконав розробку програмного забезпечення, а також підготував оригінальний текст статті.

3 статті у міжнародних наукових журналах та виданнях, що індексуються у наукометричних базах Scopus та Web of Science:

2. Malanin V., Chaikovsky I. Combined ECG Analysis Approach Based on Detection of Deviation from the Personal or Population Norm Using Wearable Devices // Lecture Notes in Networks and Systems. Singapore: Springer, 2024. Vol. 1000, 193—203 (2024). DOI: 10.1007/978-981-97-3289-0_16.

Особистий внесок здобувача: Розробив двоетапний персоналізований підхід до аналізу ЕКГ з використанням носимих пристроїв, включаючи створення моделей персональних норм і архітектури для постійного моніторингу та виявлення відхилень. Внесок полягає в запропонуванні методології, обробці та аналізі даних ЕКГ і вирішенні технічних викликів застосування цієї технології, а також написання тексту статті.

3. Bocharov M., Stasiuk V., Osyodlo V., Ryzhenko T., Malanin V., Chumachenko D., Chaikovsky I. Assessment of the physiological cost of activities for Ukrainian defense forces officers using a miniature ECG device // Frontiers in Cardiovascular Medicine. 2023. № 10. DOI: 10.3389/fcvm.2023.1239128.

Особистий внесок здобувача: Брав участь на етапах розроблення багаторівневої архітектури для обробки ЕКГ-сигналів, зокрема з питань інтеграції даних із носимих пристроїв, аналітики в реальному часі та алгоритмів оцінювання психоемоційного стану військовослужбовців.

4. Chaikovsky I., Bocharov M., Stasiuk V., Ryzhenko T., Frolov I., Parovska S., Sharypanov A., Malanin V., Osyodlo V. Determining the Physiological Cost of a Military Serviceman’s Activity in the Field Using Innovative Miniature Devices and Advance ECG Analysis: Different Scenarios of Use // Lecture Notes in Networks and Systems. Singapore: Springer, 2024. Vol. 1013. 187—202 (2024). DOI: 10.1007/978-981-97-3559-4_15.

Особистий внесок здобувача: Брав участь у розробці методології багаторівневого аналізу ЕКГ-сигналів для оцінки фізіологічного навантаження військовослужбовців, зокрема консультував щодо інтеграції даних з носимих пристроїв та алгоритмів оцінки психоемоційного стану.

Інші публікації

5. Malanin V. Adaptive Health Questionnaires: Methods, Implementation and User Impact // International Journal for Multidisciplinary Research (IJFMR). 2025. Vol. 7, Iss. 3. P. 1–20. DOI: ijfmr.2025.v07i03.45768.

Особистий внесок здобувача: Розробив концепцію та методологію адаптивних медичних анкет, створив алгоритми для їх динамічного формування залежно від відповідей користувача, провів експериментальне дослідження впливу таких анкет на залученість користувачів і якість отримуваних даних.

6. Malanin V. Implementing Offline-First Web Apps for Remote Healthcare Monitoring // International Journal of Research Publication and Reviews. 2025. Vol. 6, No. 5. 12351—12361. DOI: 10.5281/zenodo.15486778.

Особистий внесок здобувача: Самостійно розробив концепцію, архітектуру та реалізував приклад офлайн-орієнтованого веб-застосунку для віддаленого медичного моніторингу з використанням технологій Progressive Web Apps (PWA), IndexedDB, Service Workers та методів кешування даних.

7. Malanin V. Designing Accessible Web Interfaces for Aging and Vulnerable Populations // International Journal of Modern Research in Engineering and Technology. 2025. Vol. 8, No. 5. 8449—8460. DOI: 10.15680/IJMRSET.2025.0805159.

Особистий внесок здобувача: Запропонував концепцію, провів аналіз потреб користувачів, здійснив огляд та систематизацію рекомендацій щодо дизайну веб-інтерфейсів для літніх і вразливих груп населення, а також сформулював практичні рекомендації для розробників та регуляторних органів з побудови інклюзивних цифрових середовищ.

8. Malanin V. Survey fatigue in digital health applications: causes, consequences, and prevention strategies // International Journal of Engineering Technology Research & Management (IJETRM). 2025. Vol. 9, No. 5. 280—289. DOI: 10.5281/zenodo.15464581.

Особистий внесок здобувача: Запропонував концепцію дослідження причин та наслідків втоми від опитувань у цифрових медичних додатках, а також сформулював практичні рекомендації щодо її попередження з використанням адаптивних алгоритмів, персоналізації та принципів дизайну, орієнтованого на користувача.

9. Malanin V. Anonymization Techniques for Large-Scale Health Databases: A Critical Review // International Journal of Innovative Research in Science, Engineering and Technology. 2025. Vol. 14, No. 5. 11719—11728. DOI: 10.15680/IJIRSET.2025.1405015.

Особистий внесок здобувача: Запропонував концепцію дослідження, провів систематичний огляд сучасних технік анонімізації великих медичних баз даних (зокрема k-анонімність, l-різноманітність, диференційна приватність), та

сформулював рекомендації щодо їх практичного застосування з урахуванням вимог GDPR та HIPAA.

10. https://github.com/vladmalanin/lifecalculator/ – Бібліотека з відкритими кодами розроблених моделей прогнозування очікуваної тривалості життя // Маланін В.Ю. – 2025.

Патентні заявки на підходи та методи, які застосовуються в роботі:

11. Adaptive Content Delivery Systems and Methods. Патентна заявка США № 18/464,191; 2024; заявники S.K. Zamoshchik, V. Malanin.

12. Systems for Reducing Media Data Size while Maintaining Perceptual Quality. Патентна заявка США № 18/463,549; 2024; заявники S.K. Zamoshchik, V. Malanin.

13. Заявка на корисну модель № u 2022 04866 від 19.12.2022. Розгляд заявки тимчасово призупинено у зв'язку з необхідністю додаткового аналізу відповідності об'єкта заявки вимогам законодавства України щодо патентування методів діагностики та лікування.

Структура дисертації. Дисертаційна робота складається зі вступу, трьох основних розділів, висновків, списку використаних джерел та додатків.

У вступі обґрунтовано актуальність теми дослідження, сформульовано мету, завдання, об'єкт і предмет дослідження, визначено наукову новизну, теоретичне та практичне значення роботи.

У першому розділі здійснено аналіз сучасних моделей прогнозування тривалості життя, систематизовано методичні підходи, визначено теоретичні та методологічні основи розробленої моделі.

Другий розділ присвячено розробці та практичній реалізації веб-застосунку для прогнозування тривалості життя, описано архітектуру, технологічні рішення, алгоритми інтеграції й обробки даних.

У третьому розділі представлено результати експериментального дослідження, проведено аналіз ефективності та точності розроблених технологій, а також окреслено перспективи подальшого розвитку запропонованих методів і моделей.

У висновках узагальнено основні результати дисертаційного дослідження, сформульовано рекомендації щодо їх практичного використання.

Структура і об’єм роботи. Дисертація викладена на 167 сторінках друкованого тексту, включає 9 рисунків та 10 таблиць, складається зі вступу, 3 розділів, загальних висновків, списку використаних джерел із 102 найменувань та 4 додатків. Загальний обсяг основної частини дисертації становить 122 сторінок.

РОЗДІЛ 1 ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ ПРОГНОЗУВАННЯ ТРИВАЛОСТІ ЖИТТЯ

1.1 Аналіз підходів і моделей прогнозування

1.1.1 Аналіз традиційних демографічних моделей (Lee-Carter, Gompertz-Makeham)

Одним з найважливіших аспектів прогнозування тривалості життя є використання класичних демографічних моделей, які забезпечують розуміння загальних тенденцій та закономірностей смертності населення [1, 2]. Серед таких традиційних моделей найбільшого поширення набули моделі Lee-Carter та Gompertz-Makeham [3].

Модель Lee-Carter була запропонована Рональдом Лі та Лоуренсом Картером у 1992 році [4] і набула значної популярності завдяки простоті, гнучкості та ефективності у довгостроковому прогнозуванні смертності. В основі цієї моделі лежить ідея факторного аналізу часових рядів коефіцієнтів смертності, що дозволяє оцінювати зміни смертності за різними віковими групами та прогнозувати її майбутні значення.

Математично модель Lee-Carter записується наступним чином:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t},\ \ \varepsilon_{x,t} \sim N(0, \sigma^2) \quad (1)$$

де $m_{x,t}$ — коефіцієнт смертності у віці x в момент часу t, a_x — середній рівень логарифма коефіцієнта смертності для віку x, b_x — вікова компонента, що характеризує чутливість смертності у віці x до змін загального рівня смертності, k_t — загальний індекс смертності, що залежить лише від часу. При цьому припускається, що похибка $\varepsilon_{x,t}$ має нормальний розподіл із нульовим середнім і постійною дисперсією σ^2, що формально записується як $\varepsilon_{x,t} \sim N(0, \sigma^2)$. Це означає, що модельні відхилення реальних коефіцієнтів смертності від прогнозованих мають випадковий характер, є незалежними і однаково розподіленими для всіх вікових груп та часових періодів.

Параметри моделі зазвичай оцінюють методом сингулярного розкладу (Singular Value Decomposition — SVD), який полягає в розкладі матриці логарифмів коефіцієнтів смертності на множники:

$$M_{X \times T} = A_{X \times 1} \cdot 1'_{1 \times T} + B_{X \times 1} \cdot K'_{1 \times T} + E_{X \times T} \quad (2)$$

де $M_{X \times T}$ — матриця з елементами $\ln(m_{x,t})$, $A_{X \times 1}$ — вектор середніх значень смертності за віком, $B_{X \times 1}$ — вектор, що показує чутливість смертності за віком до часових змін, $K'_{1 \times T}$ — вектор часових коефіцієнтів, що характеризують загальний рівень смертності у різні моменти часу, $E_{X \times T}$ — матриця похибок, $1'_{1 \times T}$ — одиничний вектор-рядок.

Після розкладу SVD отримують оцінки параметрів, за якими здійснюють прогноз k_t методом часових рядів, найчастіше за допомогою моделі часових рядів типу ARIMA (Autoregressive Integrated Moving Average — інтегрована модель авторегресії і ковзного середнього):

$$k_t = k_{t-1} + c + u_t,\ \ u_t \sim N(0, \sigma^2) \quad (3)$$

де k_t — значення часового індексу смертності в поточний момент часу t, k_{t-1} — значення часового індексу смертності в попередній момент часу $t-1$, c — константа (дрейф), яка описує середню швидкість зміни смертності у часі, u_t — випадкова похибка (залишок моделі), яка має нормальний розподіл із нульовим середнім та дисперсією σ^2. Це означає, що випадкові зміни індексу смертності не мають систематичного ухилу і характеризуються стабільною варіабельністю. Таким чином, ця формула стверджує, що значення часового індексу смертності k_t в кожний наступний період визначається значенням попереднього періоду з додаванням константи та випадкової складової. Подібний підхід дозволяє врахувати тенденції зміни смертності та її випадкові коливання у часі, що є ключовим для достовірності прогнозу.

Модель Gompertz-Makeham є класичною моделлю, яка широко використовується для аналізу смертності людини. Її запропонував Бенджамін Ґомпертц у 1825 році [5], а пізніше вона була розширена Вільямом Мейкхемом у 1860 році [6]. Модель описує смертність як функцію віку та складається з двох компонентів: вікової (експоненційне зростання смертності з віком) та незалежної від віку константи смертності.

Математично модель Gompertz-Makeham визначається формулою:

$$\mu(x) = A + R \cdot e^{\alpha x} \tag{4}$$

де $\mu(x)$ — інтенсивність смертності у віці x, A — константа, що визначає базову інтенсивність смертності, незалежну від віку, R — коефіцієнт, що визначає початковий рівень смертності, яка залежить від віку, α — параметр, що визначає темп зростання смертності з віком.

Функція виживання $S(x)$, що визначає ймовірність дожиття до віку x, обчислюється за формулою:

$$S(x) = e^{\left\{-\int_{\{0\}}^{\{x\}}\mu(t),dt\right\}} \tag{5}$$

Для оцінювання параметрів моделі Gompertz-Makeham використовують метод максимальної правдоподібності (MLE), що дозволяє максимально ефективно підібрати параметри моделі до реальних даних про смертність.

Модель Lee-Carter має перевагу у довгострокових прогнозах, оскільки враховує часову динаміку загального рівня смертності, однак вона не завжди добре відображає вікову специфіку смертності для дуже похилого віку [7]. Натомість модель Gompertz-Makeham є більш точною для моделювання смертності у похилому віці завдяки експоненційному зростанню з віком, але менш адаптивна для довгострокового прогнозування через фіксованість параметрів [8].

Незважаючи на поширеність та визнання традиційних демографічних моделей, таких як Lee-Carter та Gompertz-Makeham, вони мають низку суттєвих обмежень, які роблять їх непридатними для використання у даному дослідженні. Модель Lee-Carter, хоча і демонструє високу ефективність при прогнозуванні загальних тенденцій смертності населення, має принципові недоліки у персоналізованому прогнозуванні тривалості життя для окремих осіб. Зокрема, вона базується на агрегованих історичних демографічних даних і не здатна враховувати індивідуальні медичні, соціальні та поведінкові фактори, що є критичними для точності персоналізованих прогнозів у цьому дослідженні.

У свою чергу, модель Gompertz-Makeham ефективна переважно для опису смертності у похилому віці, де смертність експоненційно зростає. Однак вона є статичною моделлю, що не враховує динамічні зміни факторів ризику у часі та індивідуальні відмінності в стані здоров'я, соціальних умовах життя й поведінці людей. Також, модель Gompertz-Makeham не адаптується до змін зовнішніх умов (наприклад, кризових явищ або змін у системі охорони здоров’я), що обмежує її практичну корисність у ситуації сучасної України з її динамічними викликами (війна, пандемія COVID-19, соціальні та екологічні кризи).

Таким чином, через ці обмеження моделей Lee-Carter та Gompertz-Makeham, вони були свідомо виключені з подальших тестувань і порівняльних аналізів у даній

роботі, де основною вимогою є висока персоналізація прогнозу з врахуванням динамічних медичних, соціальних і поведінкових факторів. Зазначені моделі є цікавими насамперед у контексті історичного розвитку методології прогнозування.

1.1.2 Аналіз сучасних цифрових та інтегрованих моделей (MPoRT, ONS, GBD)

Традиційні демографічні моделі прогнозування, хоча й залишаються фундаментом у розумінні загальних закономірностей смертності, мають суттєві обмеження, пов'язані з кількістю та різноманітністю враховуваних факторів [9-12]. Це зумовлює необхідність застосування сучасних прогностичних моделей, що базуються на аналізі великих даних («Big Data») та використовують інтегровані підходи до оцінки тривалості життя та смертності [13]. Нижче розглянемо найбільш визнані сучасні моделі, які спеціально орієнтовані на оцінку ймовірної тривалості життя.

Проєкт Big Life є комплексною канадською платформою прогнозування здоров'я та тривалості життя, що була створена з метою оцінки ризиків смертності та хронічних захворювань на рівні населення та окремих осіб. Основою Big Life є модель MPoRT (Mortality Population Risk Tool), яка дозволяє здійснювати персоналізовані прогнози, базуючись на соціально-економічних, поведінкових та медичних факторах [14].

Математично MPoRT базується на пропорційній моделі ризиків Кокса (Cox proportional hazards):

$$h(t) = h_{0(t)} e^{(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)} \quad (6)$$

де $h(t)$ — ризик (інтенсивність) смерті у момент часу t, $h_{0(t)}$ — базова інтенсивність смерті, що не залежить від факторів, x — значення факторів ризику (наприклад, куріння, фізична активність, хронічні захворювання, дієта, дохід,

освіта), β — коефіцієнти моделі, отримані з великих масивів даних за допомогою методів максимальної правдоподібності.

MPoRT дозволяє розраховувати втрату очікуваної тривалості життя (life expectancy lost):

$$LE_{lost} = LE_{optimal} - LE_{current} \quad (7)$$

де $LE_{optimal}$ — прогнозована тривалість життя за оптимальних умов, а $LE_{current}$ — очікувана тривалість життя за поточних умов життя та здоров'я особи.

Big Life забезпечує доступність прогнозів для громадськості, дозволяючи людям оцінювати вплив своїх звичок та стану здоров'я на тривалість життя через інтерактивні онлайн-калькулятори.

Модель Global Burden of Disease (GBD, українською «Глобальний тягар хвороб») — це масштабний міжнародний науковий проєкт, розроблений та реалізований Інститутом показників та оцінки здоров’я (Institute for Health Metrics and Evaluation, IHME) при Університеті Вашингтона [15, 16]. Основна мета проєкту — оцінка глобального тягаря хвороб та травм, а також факторів ризику для здоров'я населення у різних країнах і регіонах світу. GBD включає:

- аналіз тенденцій смертності та захворюваності;
- прогнозування очікуваної тривалості життя на рівні країн та регіонів;
- оцінку втрати років життя через хвороби та фактори ризику (DALY — Disability-Adjusted Life Years).

В основі моделі лежать регресії Кокса (6) та узагальнені регресійні моделі (Generalized Linear Models, GLMs), які дозволяють враховувати вплив численних медичних, соціально-економічних і поведінкових факторів на очікувану тривалість життя населення.

Основним показником моделі GBD є DALY — роки життя, скориговані на непрацездатність, який можна виразити наступним чином:

$$DALY = YLL + YLD \quad (8)$$

де YLL (Years of Life Lost) — втрачені роки життя через передчасну смерть, YLD (Years Lived with Disability) — роки, прожиті з інвалідністю або хворобою.

$$YLL = N \times L \quad (9)$$

де N — кількість смертей, L — середня кількість років життя, втрачених через передчасну смерть.

$$YLD = I \times DW \times L \quad (10)$$

де I — кількість випадків захворювання або травм, DW — коефіцієнт тяжкості інвалідності (від 0 до 1), L — середня тривалість захворювання або інвалідності (у роках).

Результати GBD використовуються для формування політик у сфері охорони здоров’я та профілактики, оцінки ефективності медичних заходів та втручань та порівняння стану здоров'я населення в різних країнах. GBD є одним з найбільш авторитетних джерел інформації щодо глобального здоров’я, що регулярно оновлюється і використовується міжнародними організаціями (ВООЗ, Світовий банк) та урядами країн.

Модель ONS (Office for National Statistics, Велика Британія) розраховує очікувану тривалість життя на основі детальних демографічних даних, зокрема показників смертності за віковими групами та статтю, які отримуються з національних реєстрів смертей та оцінок чисельності населення [17, 18]. Для підвищення точності результатів ONS застосовує згладжування, формуючи таблиці смертності на основі середніх даних за трирічні періоди.

Таблиця смертності моделює ймовірність смерті для кожної вікової групи, починаючи з умовної когорти 100 000 новонароджених. Для кожного віку обчислюється ймовірність смерті та кількість тих, хто виживає до наступного вікового інтервалу. Такі розрахунки виконуються послідовно для всіх вікових груп, що дозволяє визначити очікувану тривалість життя на будь-якому віці.

Очікувана тривалість життя обчислюється як середня кількість років, яку має прожити особа від певного початкового віку за умови, що поточні вікові показники смертності залишаться незмінними. Формула обчислення очікуваної тривалості життя:

$$e_x = \frac{T_x}{l_x} \tag{11}$$

де T_x — сумарна кількість років, яку ще проживуть всі особи, що дожили до віку x, а l_x — число тих, хто дожив до віку l_x.

ONS застосовує прогнозні моделі для екстраполяції тенденцій смертності у майбутнє. Основним методом прогнозування є модель «вік—період—когорта» (Age-Period-Cohort, APC), яка розділяє зміни смертності на вікові ефекти, періодні ефекти (загальні річні зміни смертності для всіх вікових груп) та когортні ефекти (особливості смертності конкретних поколінь). Модель використовує історичні дані про смертність населення за останні 50 років і прогнозує коефіцієнти смертності для кожної вікової групи на кілька десятиліть уперед.

Результати прогнозування регулярно публікуються ONS у вигляді національних таблиць смертності, які використовуються як державними органами для планування соціальної політики та охорони здоров'я, так і приватним сектором (наприклад, страховими компаніями). Також ONS надає інтерактивні калькулятори тривалості життя, що дозволяють користувачам оцінити очікувану тривалість життя залежно від віку та статі.

Крім моделей MPoRT, GBD та ONS, у світовій практиці також використовуються такі моделі прогнозування тривалості життя, як Sullivan (в рамках проєкту EHEMU) [19], моделі зроблені з даних UK Biobank [20, 21] або таблиць Eurostat [22]. Проте саме MPoRT, GBD та ONS були обрані для детального аналізу в даній роботі, їхні методології є найбільш доступними, ретельно валідованими і регулярно оновлюються, що робить їх надійними для застосування

в міжнародних дослідженнях, формуванні державної політики та прийнятті стратегічних рішень у сфері охорони здоров’я.

1.1.3 Аналіз методів машинного навчання та нейронних мереж для прогнозування

Завдання прогнозування тривалості життя традиційно вирішувались за допомогою статистичних та демографічних підходів, які мали на меті врахувати типові фактори ризику і загальні закономірності виживання [23-25]. Однак зростання обсягів і складності доступних даних, а також потреба в персоналізованих прогнозах, зумовили перехід до використання складніших математичних та обчислювальних підходів, зокрема сучасних методів математичного моделювання та алгоритмів машинного навчання.

Статистичні методи математичного моделювання в задачах прогнозування тривалості життя переважно представлені моделями аналізу виживання. В основі таких моделей лежить концепція оцінювання ймовірності виживання особи до певного віку чи моменту часу, враховуючи вплив ряду факторів. Однією з найбільш поширених є модель пропорційних ризиків Кокса (Cox proportional hazards model) [26], яка дозволяє визначати, як різні фактори ризику впливають на ризик смерті чи іншої критичної події з часом (6). Параметри β_i зазвичай оцінюють за допомогою методу максимальної правдоподібності (MLE). Цей метод дозволяє отримати максимально точні оцінки параметрів моделі, які найбільш імовірно відповідають спостережуваним даним.

Крім того, широко використовуються логістичні регресії, особливо для оцінки ймовірності смерті чи захворювання в межах визначеного часового проміжку [27]. Загальний вигляд логістичної регресії:

(12)

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}$$

де P — ймовірність настання події, наприклад смерті або виникнення захворювання, а x — фактори ризику, такі як вік, стать, спосіб життя, генетичні маркери тощо. Логістичні моделі особливо корисні тоді, коли йдеться про прогнозування події у чітко визначений період часу.

Водночас традиційні статистичні моделі мають суттєві обмеження щодо здатності враховувати складні нелінійні взаємозв'язки та взаємодію великої кількості факторів. Відповіддю на ці виклики стали сучасні алгоритми машинного навчання (Machine Learning) [28], які забезпечують потужні інструменти для аналізу багатовимірних, гетерогенних масивів даних.

Одним з таких підходів є метод випадкових лісів (Random Forest). Метод випадкового лісу є одним із найбільш потужних та ефективних сучасних алгоритмів машинного навчання, який застосовується в широкому спектрі задач прогнозування, включаючи прогнозування тривалості життя та смертності. Вперше запропонований Лео Брейманом (Leo Breiman) у 2001 році [29], випадковий ліс є ансамблевим методом, який базується на комбінації великої кількості незалежно побудованих дерев рішень (Decision Trees), а його ключовою перевагою є стабільність, точність прогнозів і стійкість до перетренування (overfitting).

З математичної точки зору, алгоритм випадкового лісу працює за таким принципом. З наявного набору даних формується множина бутстреп-вибірок (bootstrap samples) — це випадкові вибірки з поверненням, які використовуються для побудови окремих дерев рішень [30]. Кожне дерево рішення будується не на всій множині доступних ознак (факторів ризику), а на випадково вибраному піднаборі цих ознак. Такий підхід називають випадковим відбором підмножин ознак (random feature selection). Саме поєднання бутстрепу та випадкового вибору ознак забезпечує незалежність та різноманітність дерев, що дозволяє значно підвищити загальну точність і надійність прогнозу.

Після того, як усі дерева ансамблю побудовані, прогноз для нових даних визначається шляхом агрегування результатів усіх дерев. Якщо вирішується задача класифікації (наприклад, прогнозування ймовірності смерті у визначений період), остаточне рішення ухвалюється за принципом голосування (majority voting) — прогнозом вважається той клас, за який проголосувала більшість дерев. Якщо йдеться про регресійні задачі (наприклад, прогнозування конкретного значення очікуваної тривалості життя), результатом буде середнє арифметичне значення прогнозів усіх дерев. Формально це записується таким чином:

$$\hat{y} = \left(\frac{1}{K}\right)\sum_{k=1}^{K} f_{k(x)} \tag{13}$$

де $\hat{y}$ — це прогнозоване значення, K — кількість дерев в ансамблі, а $f_{k(x)}$ — результат прогнозу k дерева для даного набору факторів x.

Важливою перевагою випадкового лісу є можливість автоматичної оцінки важливості факторів ризику. Алгоритм здатний оцінити, наскільки кожна ознака сприяє підвищенню точності прогнозу. Це дозволяє виявити фактори, що мають найбільший вплив на очікувану тривалість життя або ризик смерті. Для цього використовують показники, такі як Mean Decrease Impurity (середнє зменшення неоднорідності вузлів дерева) або Mean Decrease Accuracy (середнє зменшення точності), які вказують, наскільки погіршується точність прогнозу при виключенні певної ознаки з моделі.

Попри численні переваги, метод випадкових лісів має і деякі обмеження. Зокрема, він часто сприймається як «чорний ящик» через складність інтерпретації результатів порівняно з класичними лінійними моделями. Це може створювати певні труднощі в ситуаціях, де важлива саме зрозумілість і прозорість моделі (наприклад, у клінічних рішеннях). Проте ця проблема успішно вирішується за допомогою додаткових інструментів інтерпретації, таких як SHAP (SHapley Additive exPlanations) [31], LIME (Local Interpretable Model-agnostic Explanations) [32] або PDP (Partial Dependence Plots) [33].

Таким чином, завдяки своїй гнучкості, високій точності та стійкості до помилок випадковий ліс став одним із провідних методів у задачах прогнозування тривалості життя, медичної діагностики та оцінки ризиків. Його застосування дозволяє ефективно працювати з великими наборами складних і неоднорідних даних, формувати надійні та деталізовані прогнози й ефективно визначати ключові фактори ризику, що впливають на очікувану тривалість життя.

Ще одним потужним алгоритмом є бустинг (Gradient Boosting Machines, GBM), особливо його сучасна реалізація — XGBoost. Алгоритми бустингу (Gradient Boosting Machines, GBM) належать до сімейства сучасних, потужних та ефективних методів машинного навчання, які широко застосовуються для прогнозування різноманітних подій, включаючи ймовірність смерті та очікувану тривалість життя. Ідея бустингу вперше була запропонована Робертом Шапіре (Robert Schapire) [34] та Йоавом Фройндом (Yoav Freund) [35, 36] ще у 1990-х роках, однак справжню популярність ці методи здобули завдяки розробці алгоритму градієнтного бустингу (Gradient Boosting), автором якого є Джером Фрідман (Jerome Friedman) у 2001 році [37]. Відмінною рисою цього методу є послідовне побудування ансамблю моделей (зазвичай дерев рішень), де кожна наступна модель спрямована на зменшення помилок, допущених попередніми моделями ансамблю.

З математичної точки зору градієнтний бустинг працює наступним чином. На першому етапі будується просте дерево рішень на основі початкових даних, після чого визначається його помилка прогнозування. Наступні дерева рішень створюються і додаються до ансамблю з метою мінімізувати ці помилки, поступово «виправляючи» похибки попередніх моделей. Це відбувається шляхом побудови кожного нового дерева за допомогою градієнтного спуску (gradient descent) по функції втрат (loss function), яка вимірює відхилення прогнозу моделі від реальних значень. Саме завдяки такій ітеративній процедурі алгоритм GBM ефективно навчається навіть на великих і складних наборах даних, точно враховуючи складні взаємозв'язки між факторами ризику.

Загальна формула прогнозу, яку використовує градієнтний бустинг, має вигляд:

$$\hat{y}_i = \sum_{k=1}^{K} f_{k(x_i)} \tag{14}$$

де $\hat{y}_i$ — прогноз для і-го спостереження, $f_{k(x_i)}$— прогноз k-ї моделі (дерева рішень) у складі ансамблю, а K — загальна кількість моделей у бустинговому ансамблі. На кожному кроці навчання нове дерево додається з коефіцієнтом швидкості навчання (learning rate), який регулює швидкість адаптації моделі до помилок. Малий learning rate забезпечує більш повільне, але точне та стабільне навчання, а великий — швидше навчання, яке може призвести до перетренування (overfitting).

Особливим проривом у розвитку методів градієнтного бустингу стала сучасна реалізація, яка отримала назву XGBoost (Extreme Gradient Boosting), розроблена Тяньці Ченом (Tianqi Chen) у 2016 році. XGBoost швидко став одним із найпопулярніших інструментів у сфері аналізу даних, зокрема завдяки винятковій точності, високій швидкості роботи та оптимізованій реалізації алгоритму. На відміну від класичного GBM, XGBoost використовує ряд суттєвих інновацій і вдосконалень: регуляризацію для зменшення перетренування, паралельні обчислення для прискорення роботи, можливість роботи з пропущеними значеннями даних, автоматичну оптимізацію глибини дерев і додаткові параметри, які дозволяють краще контролювати процес навчання. Завдяки цим особливостям, XGBoost став еталонним алгоритмом у задачах прогнозування, включаючи задачу прогнозування очікуваної тривалості життя.

Як і випадкові ліси, алгоритми бустингу, зокрема XGBoost, дозволяють автоматично оцінювати важливість факторів ризику, використовуючи такі показники, як Gain (збільшення точності прогнозу, що забезпечує фактор), Weight (частота використання фактору в розбиттях дерев) та Cover (обсяг даних, які

покриваються фактором). Це дозволяє отримувати чіткі, практично корисні оцінки факторів ризику, які впливають на прогнозовану тривалість життя або ризик смерті.

Попри численні переваги, бустингові алгоритми мають і певні обмеження. Основними з них є складність інтерпретації результатів (аналогічно до випадкових лісів), чутливість до параметрів моделі (які вимагають ретельного налаштування — hyperparameter tuning), а також можливість перетренування при неправильному виборі параметрів. Однак сучасні інструменти, такі як згадані вище SHAP або LIME, дозволяють частково подолати ці недоліки, надаючи прозорість та пояснюваність результатам навіть дуже складних моделей.

Таким чином, градієнтний бустинг і його сучасна реалізація XGBoost — це ефективні, точні та гнучкі інструменти машинного навчання, які здатні ефективно використовувати великі набори складних даних для побудови моделей прогнозування тривалості життя, здоров’я та інших важливих медичних і соціальних показників. Завдяки високій точності та можливості визначення значущих факторів, ці методи сьогодні широко застосовуються в медичній практиці, епідеміології, демографії та соціальному прогнозуванні.

Окрему роль у прогнозуванні тривалості життя відіграють нейронні мережі. Нейронні мережі — це сучасні та потужні моделі машинного навчання, які стали одним з найбільш ефективних інструментів для вирішення складних завдань, зокрема задач прогнозування тривалості життя, здоров’я та смертності [39-41]. В основі нейронних мереж лежить ідея моделювання роботи біологічного мозку людини: модель складається з багатьох простих обчислювальних елементів — нейронів, які об'єднуються у шари (layers) і взаємодіють між собою шляхом передачі сигналів. Завдяки такій багаторівневій структурі нейронні мережі здатні ефективно апроксимувати складні, нелінійні взаємозв'язки між великою кількістю вхідних факторів та прогнозованими значеннями.

Серед нейромережевих підходів, які широко використовуються для прогнозування тривалості життя та оцінювання факторів ризику, особливої уваги

заслуговують багатошарові нейронні мережі (Multilayer Perceptron — MLP) [42], згорткові нейронні мережі (Convolutional Neural Networks — CNN) [43] та рекурентні нейронні мережі (Recurrent Neural Networks — RNN) [44], включаючи їхні спеціалізовані варіанти — мережі довгої короткочасної пам'яті (Long Short-Term Memory — LSTM) та блоки з вентилями (Gated Recurrent Units — GRU).

Стандартна архітектура нейронної мережі включає три основних типи шарів: вхідний (input layer), приховані (hidden layers) та вихідний шар (output layer). Вхідний шар отримує дані, які представляють собою набір факторів або характеристик (наприклад, вік, стать, медичні показники, соціально-економічні фактори), приховані шари виконують основну роботу, моделюючи нелінійні залежності, а вихідний шар формує остаточний прогноз. Кожен нейрон у шарах обчислює зважену суму вхідних сигналів і застосовує до цієї суми нелінійну функцію активації (activation function), яка дозволяє моделі будувати складні залежності між змінними.

Багатошарові нейронні мережі (MLP) є одним із найпоширеніших і найефективніших типів нейронних мереж, що активно застосовуються для вирішення широкого спектру задач прогнозування, включаючи задачі прогнозування тривалості життя, аналізу виживання та визначення факторів ризику [45]. Їх широка популярність зумовлена універсальністю, простотою архітектури, а також здатністю ефективно моделювати складні нелінійні взаємозв'язки між численними факторами ризику та прогнозованими подіями.

Архітектурно MLP складається з набору нейронів, організованих у кілька послідовних шарів: вхідний шар (input layer), один або декілька прихованих шарів (hidden layers) та вихідний шар (output layer). Кожен нейрон являє собою елементарну обчислювальну одиницю, що обчислює зважену суму своїх вхідних сигналів з додаванням зміщення (bias), після чого застосовує нелінійну функцію активації. Завдяки наявності цих функцій активації (зокрема таких, як сигмоїда, гіперболічний тангенс, ReLU — Rectified Linear Unit, або Leaky ReLU),

багатошаровий перцептрон здатний ефективно апроксимувати нелінійні залежності різного ступеня складності.

Математично процес роботи одного шару нейронів у багатошаровій нейронній мережі MLP можна записати такою формулою:

$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{15}$$

де x_i — це вхідні сигнали, w_i— вагові коефіцієнти, що налаштовуються під час навчання мережі, b —зміщення (bias), а $\sigma(\cdot)$ —нелінійна функція активації (наприклад, ReLU чи сигмоїда). Вихід прихованих шарів передається далі до наступних шарів, поки остаточний прогноз не буде сформовано на виході мережі.

Процес навчання багатошарової нейронної мережі MLP найчастіше відбувається за допомогою алгоритму зворотного поширення помилки (backpropagation) у поєднанні з градієнтним спуском (gradient descent). Під час навчання модель спочатку робить попередній прогноз для кожного навчального прикладу, після чого порівнює його з реальним значенням, визначає похибку (функцію втрат, наприклад Mean Squared Error чи Cross-Entropy), та поступово коригує вагові коефіцієнти нейронів, мінімізуючи таким чином помилку прогнозу. Такий підхід дозволяє ефективно налаштовувати велику кількість вагових параметрів, забезпечуючи високу точність прогнозу навіть у складних багатовимірних задачах.

Однією з найбільш важливих переваг MLP є їх здатність ефективно обробляти та інтегрувати різні типи даних: числові (вік, показники крові, індекс маси тіла), категоріальні (стать, соціально-економічний статус, анамнез захворювань), а також змішані набори даних, що часто зустрічаються в медичній та соціальній сферах. Це дозволяє ефективно враховувати широкий спектр факторів ризику та отримувати індивідуалізовані, персоналізовані прогнози тривалості життя та інших клінічних подій.

При застосуванні багатошарових нейронних мереж для прогнозування тривалості життя та аналізу факторів ризику важливе значення має процедура налаштування гіперпараметрів моделі. До таких параметрів належать кількість прихованих шарів, кількість нейронів у кожному шарі, тип функції активації, швидкість навчання (learning rate), алгоритми регуляризації та ранньої зупинки (early stopping). Саме ретельне налаштування цих гіперпараметрів дозволяє уникнути перетренування (overfitting), коли модель надмірно пристосовується до навчальних даних і втрачає здатність до ефективного прогнозування на нових, раніше невідомих наборах даних.

Важливим аспектом застосування MLP в медичній сфері є також питання пояснюваності результатів моделі. Оскільки багатошарові нейронні мережі традиційно сприймаються як так звані «чорні ящики», особливо важливим є використання сучасних інструментів пояснення моделей, таких як SHAP (SHapley Additive exPlanations) та LIME (Local Interpretable Model-agnostic Explanations). Ці методи дозволяють визначати та інтерпретувати, які саме фактори ризику найбільше впливають на прогнозовані результати, значно підвищуючи прозорість моделі і забезпечуючи зрозумілість її прогнозів для медичних фахівців і кінцевих користувачів.

Проте, попри численні переваги, використання багатошарових нейронних мереж має й певні обмеження. До основних із них належать необхідність великих наборів даних для ефективного навчання, високі обчислювальні витрати, ризик перетренування за неправильного налаштування гіперпараметрів, а також складність у прямій інтерпретації результатів моделі без додаткових інструментів пояснення.

Отже, багатошарові нейронні мережі (MLP) становлять потужний, універсальний та ефективний інструмент для вирішення складних задач прогнозування тривалості життя, аналізу виживання та визначення ключових факторів ризику. Успішне їхнє використання вимагає ретельного підходу до налаштування архітектури, ефективної валідації та застосування сучасних методів

інтерпретації результатів, що дозволяє отримати точні, надійні та практично корисні прогнозні моделі.

Згорткові нейронні мережі (CNN) є одним з найбільш ефективних та широко використовуваних типів нейромережевих моделей, особливо для аналізу структурованих, просторових або зображувальних даних [46]. Спочатку CNN набули популярності у сферах комп'ютерного зору, розпізнавання зображень та відеоаналізу, але сьогодні вони успішно застосовуються також для вирішення складних медичних завдань, включаючи прогнозування тривалості життя, діагностику захворювань і аналіз медичних даних.

Архітектура CNN принципово відрізняється від класичних багатошарових нейронних мереж (MLP). Основною особливістю CNN є використання спеціальних шарів — згорткових (convolutional layers), які дозволяють автоматично витягувати локальні ознаки (features) із вхідних даних. Завдяки цій властивості CNN ефективно виявляють важливі патерни та просторові закономірності у великих обсягах складно структурованих даних, таких як медичні зображення (рентген, МРТ, КТ), генетичні карти або навіть часові та структуровані числові дані, що представляються у вигляді багатовимірних матриць.

Процес роботи згорткового шару можна представити математично наступним чином. Згортковий шар виконує операцію згортки (convolution) вхідного сигналу (наприклад, зображення) із спеціальними фільтрами (ядрами згортки). Формально, згорткова операція описується формулою:

$$y_{i,j} = f\left(\sum_m \sum_n x_{i+m,j+n} \cdot w_{m,n} + b\right) \tag{16}$$

де $y_{i,j}$ — значення вихідного сигналу (елемента карти ознак) у позиції i, j

$x_{i+m,j+n}$ — елементи вхідного сигналу (наприклад, матриця пікселів зображення), b — зміщення (bias), $w_{m,n}$ — елементи матриці ядра згортки (фільтр), $f(\cdot)$ — нелінійна функція активації (наприклад, ReLU).

Після проходження згорткового шару отримується набір карт ознак (feature maps), які представляють собою узагальнені характеристики вхідних даних. Наступним важливим шаром CNN зазвичай є шар підвибірки (пулінгу, pooling layer). Пулінг має на меті зменшити розмірність отриманих карт ознак, що дозволяє знизити обчислювальну складність та забезпечити певну інваріантність до локальних змін у вхідних даних. Найбільш поширеним є max-pooling, який обчислює максимальне значення у певному локальному вікні карти ознак.

В результаті послідовного застосування декількох пар згорткових і пулінгових шарів отримані ознаки передаються до одного або декількох повнозв'язних шарів (fully connected layers), які аналогічні за структурою до шарів у класичних нейронних мережах. На цьому етапі остаточно формується прогноз моделі (наприклад, прогнозований ризик смерті чи захворювання або прогнозований час життя).

Однією з головних переваг CNN є їх здатність до автоматичного виявлення важливих ознак із вхідних даних, що значно зменшує необхідність ручного попереднього відбору факторів ризику. Це особливо корисно в медичних задачах, де кількість потенційно важливих факторів є надзвичайно великою, а їх вплив на прогнозовану подію може бути прихованим або складним для визначення вручну.

Застосування CNN для оцінки ризиків захворювань вже продемонструвало високу ефективність. Наприклад, CNN активно використовуються для аналізу медичних зображень при діагностиці онкологічних, серцево-судинних захворювань, захворювань центральної нервової системи та інших патологій. Окрім того, CNN успішно застосовуються і для роботи з неструктурованими текстовими медичними даними, дозволяючи знаходити приховані закономірності та чинники ризику в історіях хвороб пацієнтів.

Водночас, успішне застосування CNN вимагає значних обчислювальних ресурсів та великої кількості якісних навчальних даних. Особливо важливою є процедура налаштування гіперпараметрів мережі: кількості шарів, розміру ядер

згортки, швидкості навчання, типу функції активації та стратегії регуляризації. Крім того, через складну структуру CNN часто сприймаються як «чорні ящики», що ускладнює безпосередню інтерпретацію результатів. Саме тому в медичних застосуваннях важливим є використання додаткових методик пояснення моделей, таких як Grad-CAM (Gradient-weighted Class Activation Mapping), SHAP або LIME, що дозволяють чітко зрозуміти, які саме області або характеристики даних найбільше впливають на прогнозовані результати.

Таким чином, згорткові нейронні мережі (CNN) становлять потужний і перспективний інструмент для вирішення складних задач діагностики хвороб та виявлення факторів ризику, особливо у випадках використання складно структурованих або зображувальних даних. Безпосередньо для прогнозування очікуваної тривалості життя вони є менш актуальними. Їх використання дозволяє значно підвищити точність прогнозів та забезпечити автоматичне виявлення важливих закономірностей, проте вимагає уважного підходу до налаштування та інтерпретації отриманих результатів, особливо у чутливих галузях, таких як медицина й охорона здоров'я.

Рекурентні нейронні мережі (RNN — Recurrent Neural Networks) є спеціалізованим типом нейронних мереж, які ефективно застосовуються для аналізу послідовних і часових даних [47]. Особливість архітектури RNN полягає в тому, що ці мережі здатні враховувати попередні стани або послідовні події, тобто володіють своєрідною «пам’яттю». Саме ця особливість робить їх надзвичайно корисними для вирішення задач, що мають часовий або послідовний характер, таких як прогнозування тривалості життя, аналіз виживання пацієнтів, виявлення динаміки факторів ризику та медичних подій.

З технічної точки зору, RNN мають структуру, яка дозволяє інформації рухатися не лише вперед, від вхідного до вихідного шару, але й передаватися назад або циклічно, забезпечуючи таким чином збереження інформації про попередні стани нейронів. Завдяки цьому рекурентні мережі можуть ефективно враховувати динамічні зміни факторів у часі та вплив попередніх станів на поточні події.

Математично функціонування стандартного рекурентного нейрона описується наступними формулами:

$$h_t = f(W\ x_t + U\ h_{t-1} + b) \quad (17)$$

де h_t— прихований стан нейрона в момент часу t, x_t— вхідні дані (ознаки) в момент часу t, h_{t-1}— прихований стан у попередній момент часу $t-1$, $W\ x_t$ — вагова матриця для вхідних значень, $U\ h_{t-1}$ — вагова матриця для прихованого стану, b — зміщення (bias), $f(\cdot)$ — нелінійна функція активації (часто використовуються тангенс гіперболічний, ReLU або сигмоїда).

Ця формула демонструє принципову особливість RNN: інформація з попередніх станів h_t^{-1} використовується для визначення поточного стану мережі h_t, що забезпечує врахування попередніх подій при прогнозуванні.

Однак звичайні рекурентні мережі мають суттєвий недолік, пов'язаний із так званою «проблемою зникнення або вибуху градієнта» (vanishing/exploding gradient problem), яка виникає при навчанні довгих часових послідовностей. Для подолання цієї проблеми були розроблені спеціалізовані типи рекурентних мереж, такі як LSTM (Long Short-Term Memory) і GRU (Gated Recurrent Unit).

LSTM — це спеціалізований тип RNN, що використовує складніші внутрішні механізми для ефективнішого збереження інформації протягом тривалих часових періодів. Основною особливістю LSTM є спеціальні вентильні механізми (gates): вхідні (input gate), вихідні (output gate) та вентилі забування (forget gate), які дозволяють мережі вибірково зберігати чи забувати інформацію на різних часових відрізках. Це робить LSTM особливо ефективними для аналізу довготривалих медичних послідовностей (наприклад, аналіз історії захворювань пацієнтів, динаміки здоров’я або впливу лікування у часі).

Формально робота LSTM описується наступними формулами:

$$f_t = \sigma\big(W_f[h_{t-1}, x_t] + b_f\big) \quad (19)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

$$h_t = o_t \odot \tanh(C_t)$$

де f_t, i_t, o_t — вентилі LSTM (забування, входу, виходу), C_t — стан комірки LSTM, $\tilde{C}_t$ — кандидат на оновлення стану комірки, h_t— вихідний прихований стан, W_f, W_i, W_o, W_C,— вагові матриці, b_f, b_i, b_C, b_o — зміщення,$\sigma(\cdot)$ — сигмоїдна функція, $\odot$ — поелементне множення,$[h_{t-1}, x_t]$— конкатенація прихованого стану та вхідних даних.

GRU є дещо простішим, але не менш ефективним різновидом рекурентних мереж. В GRU використовуються два вентилі (gate) —вентиль оновлення (update gate) і вентиль скидання (reset gate), що дозволяють керувати потоком інформації, яка проходить крізь нейрон. Це забезпечує ефективність роботи з часовими послідовностями за меншої кількості параметрів, ніж в LSTM.

Рекурентні мережі, зокрема LSTM та GRU, мають суттєві переваги для прогнозування тривалості життя, оскільки дозволяють враховувати динаміку стану здоров'я пацієнтів, зміни факторів ризику у часі, вплив попередніх медичних подій, результатів лікування та інші тимчасові залежності, які часто є критичними для точного прогнозу виживання.

Проте використання RNN також має певні недоліки. До них належать складність інтерпретації результатів, потреба у великих обсягах якісних даних та висока обчислювальна складність при навчанні моделей. Для ефективного застосування рекурентних нейронних мереж потрібен ретельний вибір архітектури, налаштування гіперпараметрів, таких як кількість прихованих шарів, розмірність прихованого стану, швидкість навчання та регуляризація.

Таким чином, рекурентні нейронні мережі (RNN), особливо у своїх сучасних різновидах LSTM та GRU представляють собою потужний інструмент скоріше для аналізу факторів ризику та вивчення динаміки стану здоров’я пацієнтів у часі, ніж для прогнозування тривалості життя.

Багатошарові нейронні мережі (MLP) завдяки своїй гнучкості та універсальності успішно використовуються для моделювання складних нелінійних взаємозв'язків між численними соціальними, медичними та демографічними факторами. Вони ефективні як у задачах класифікації (прогнозування настання конкретних подій), так і в регресійних задачах (прогнозування тривалості життя). Водночас їх ефективне застосування вимагає ретельного налаштування гіперпараметрів, що може створювати певні труднощі, особливо за наявності великих наборів неоднорідних даних.

Особливо важливе місце серед нейромережевих підходів займають рекурентні нейронні мережі (RNN), включаючи їх спеціалізовані варіанти — мережі довгої короткочасної пам’яті (LSTM) та блоки з вентилями (GRU). Ці моделі ефективно працюють із часовими та послідовними даними, завдяки чому вони особливо корисні для прогнозування перебігу захворювань і динаміки змін факторів ризику у часі. Завдяки наявності внутрішніх механізмів пам'яті ці нейромережі дозволяють враховувати вплив попередніх станів та подій на поточні прогнози, що значно підвищує точність та практичну корисність результатів.

Спільною проблемою всіх нейромережевих моделей залишається складність інтерпретації результатів, яка є особливо важливою в медичних та соціальних задачах прогнозування тривалості життя та факторів ризику. Для вирішення цієї проблеми рекомендовано застосовувати спеціальні методи пояснення моделей, такі як SHAP, LIME або Grad-CAM, які дозволяють зробити нейронні мережі більш прозорими та зрозумілими для медичних спеціалістів і кінцевих користувачів.

Таким чином, нейромережеві підходи мають високий потенціал для застосування, в першу чергу, у задачах аналізу факторів ризику, ніж прогнозування тривалості життя.

На основі здійсненого аналізу та огляду різних методів машинного навчання та нейронних мереж, для подальшого застосування у дослідженні обрано алгоритми XGBoost та Random Forest. Цей вибір зумовлений насамперед тим, що ці алгоритми довели свою ефективність у схожих задачах прогнозування тривалості життя та здоров'я населення.

Random Forest вирізняється високою стабільністю та стійкістю до шумів у даних, забезпечує стабільні результати завдяки усередненню прогнозів великої кількості незалежних дерев рішень, що робить його особливо придатним для медичних та демографічних даних. XGBoost має високу точність завдяки використанню градієнтного бустингу, ефективній оптимізації гіперпараметрів, а також можливості ефективно працювати з великими масивами даних та пропущеними значеннями.

Окремо на результатах застосування алгоритмічної моделі на реальних даних українського населення запропоновано здійснення навчання прогнозних моделей на основі алгоритмів Random Forest (далі — «Life Calculator Random Forest» або «LCR») та XGBoost (далі — «Life Calculator XGBoost» або «LCX») (Табл. 1). Це дозволяє максимально ефективно врахувати нелінійні взаємозв'язки між факторами ризику та прогнозованою тривалістю життя.

Таблиця 1. Порівняльна характеристика алгоритмів машинного навчання та обґрунтування їх вибору для прогнозування тривалості життя

Алгоритм	Переваги	Недоліки	Причини вибору
Random Forest	Висока стійкість до шумів, стабільні результати, простота у налаштуванні	Складна інтерпретація окремих рішень дерев	Надійні, стабільні результати для великих даних

XGBoost	Висока точність прогнозування, ефективність при роботі з великими масивами даних, швидкість роботи	Потребує ретельного налаштування гіперпараметрів	Високий потенціал в прогнозуванні здоров'я за попередніми дослідженнями

Інші методи, такі як нейронні мережі (MLP, CNN, RNN), хоча й демонструють високий потенціал та точність, не були обрані для основного практичного застосування через більшу актуальність для діагностики медичних факторів ризику, складність інтерпретації результатів, високі обчислювальні вимоги та потребу в більших обсягах даних для навчання, що обмежує їхню практичну цінність у контексті цього дослідження.

1.1.4 Порівняльний аналіз критеріїв оцінювання точності прогнозних моделей

Одним із ключових етапів розроблення та впровадження прогнозних моделей є оцінювання їхньої точності та ефективності. Вибір адекватних критеріїв для оцінювання точності прогнозів є важливим кроком, що впливає як на вибір конкретної моделі, так і на достовірність отриманих результатів. У сучасній практиці застосовуються різноманітні критерії, кожен з яких має свої переваги та обмеження залежно від типу моделі, особливостей наборів даних та задачі прогнозування [48, 49].

Прогнозування ймовірності подій (наприклад, ризику смерті або виникнення конкретного захворювання) є однією з поширених задач у медицині, демографії та соціальних дослідженнях. На відміну від задач прогнозування конкретних числових значень (регресії), такі задачі є задачами класифікації: модель повинна

визначити, чи настане подія для конкретної людини за певний період часу чи ні. Для коректного оцінювання якості таких моделей використовуються спеціальні критерії, які дозволяють виміряти точність прогнозу у різних аспектах.

Одним із базових підходів для оцінювання класифікаційних моделей є використання матриці помилок (confusion matrix) [50]. Ця матриця складається з чотирьох ключових показників: істинно позитивні (TP — true positive), істинно негативні (TN — true negative), хибнопозитивні (FP — false positive) та хибнонегативні (FN — false negative). Саме на основі цих чотирьох категорій можна обчислити цілий ряд важливих критеріїв оцінювання моделі.

Першим і найпростішим критерієм є точність (Accuracy), яка визначає загальну частку правильно класифікованих випадків з усіх спостережень:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{19}$$

Цей критерій досить простий та зрозумілий, але має суттєве обмеження— його значення може бути завищене у разі незбалансованості класів (тобто коли кількість позитивних і негативних випадків значно відрізняється).

Наступним важливим критерієм є чутливість (Sensitivity або Recall), що визначає, наскільки ефективно модель виявляє позитивні випадки (наприклад, випадки смерті чи розвитку захворювання). Вона вимірює частку правильно визначених позитивних подій з усіх реальних позитивних подій:

$$Sensitivity = \frac{TP}{TP + FN} \tag{20}$$

Цей критерій особливо важливий у медичних задачах, де ціна пропуску позитивного випадку (наприклад, недіагностованої хвороби) може бути надзвичайно високою.

У свою чергу, специфічність (Specificity) визначає, наскільки ефективно модель ідентифікує негативні випадки (відсутність події). Вона вимірює частку правильно визначених негативних подій серед усіх реальних негативних подій:

$$Specificity = \frac{TN}{TN + FP} \quad (21)$$

Специфічність має критичне значення у задачах, де хибнопозитивні прогнози можуть мати серйозні наслідки, такі як зайва тривога пацієнтів або непотрібні медичні процедури.

Ще одним важливим критерієм є точність прогнозу (Precision), яка визначає, яку частку з позитивно прогнозованих випадків складають справжні позитивні випадки. Це особливо важливо для задач, де є високі витрати на хибні позитивні результати:

$$Precision = \frac{TP}{TP + FP} \quad (22)$$

Щоб одночасно враховувати точність (precision) і повноту (recall), застосовується комплексний показник — F1-score (F1-міра), який є гармонійним середнім precision та recall:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (23)$$

Цей показник особливо ефективний у випадках незбалансованих наборів даних і дозволяє збалансовано оцінювати модель у ситуації, коли обидва типи помилок є небажаними.

Ще один популярний та потужний критерій — площа під кривою ROC (ROC-AUC), використовується для оцінювання загальної якості моделі, незалежно від вибору порогу класифікації [5]11. ROC-крива будується шляхом зміни порогового значення класифікації та відображає залежність між чутливістю (Sensitivity) та хибнопозитивним рівнем (1 - Specificity). Площа під цією кривою (AUC — Area Under Curve) характеризує загальну дискримінаційну здатність моделі. Значення AUC змінюється від 0,5 (модель випадкового вибору) до 1 (ідеальна модель). Високі значення AUC свідчать про високу здатність моделі ефективно розрізняти позитивні та негативні випадки незалежно від обраного порогу класифікації.

Регресійні моделі використовуються у задачах прогнозування числових значень, таких як прогноз тривалості життя, віку настання події, кількісних медичних показників, рівнів ризику або інших вимірних показників. На відміну від задач класифікації, де оцінюється ймовірність настання або ненастання певної події, у задачах регресії основною метою є визначення того, наскільки близько прогнозоване значення моделі відповідає реальному спостережуваному значенню. Відповідно, для оцінки точності таких моделей застосовуються спеціалізовані критерії, що дозволяють виміряти ступінь відхилення прогнозу від фактичних значень.

Одним із найпоширеніших критеріїв оцінювання точності регресійних моделей є середня абсолютна похибка (Mean Absolute Error, MAE) [52]. Цей критерій визначає середнє абсолютне значення відхилення прогнозованих значень від фактичних і розраховується за формулою:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{24}$$

де y_i — це реальне спостережуване значення, а $\hat{y}_i$ — прогнозоване значення для i-го спостереження. Основною перевагою MAE є його зрозуміла інтерпретація: це середнє значення абсолютних відхилень прогнозу від реальності, що має ті ж самі одиниці вимірювання, що й прогнозований показник (наприклад, роки життя). Водночас MAE має суттєве обмеження—він однаково ставиться до всіх помилок, не враховуючи, що окремі дуже великі помилки можуть мати суттєві негативні наслідки для практичного застосування моделі.

Для подолання цього обмеження застосовується інший критерій—середньоквадратична похибка (Mean Squared Error, MSE) [53]. На відміну від MAE, цей критерій надає більшу вагу більшим відхиленням прогнозу, оскільки помилки підносяться до квадрату. Формально критерій MSE визначається за формулою:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{25}$$

Саме через піднесення до квадрату MSE чутливий до великих похибок, що особливо корисно, коли вартість значних помилок дуже висока. Це робить його дуже ефективним у ситуаціях, коли потрібно зменшити саме великі відхилення. Проте квадрат одиниць вимірювання прогнозованого показника (наприклад, «роки²») може бути менш зрозумілим для кінцевих користувачів моделі, що ускладнює його інтерпретацію.

Для того, щоб повернутися до вихідних одиниць вимірювання та зробити інтерпретацію результату більш зрозумілою, використовують корінь середньоквадратичної похибки (Root Mean Squared Error, RMSE) [54], який є квадратним коренем з MSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{26}$$

RMSE має перевагу у зручності інтерпретації, оскільки значення виражене в тих самих одиницях вимірювання, що й прогнозовані значення. Як і MSE, RMSE чутливий до великих помилок і є ефективним критерієм, коли значні помилки є особливо небажаними. З цієї причини RMSE є одним із найпопулярніших критеріїв оцінювання точності регресійних моделей.

Ще одним цікавим критерієм для регресійних моделей є середня абсолютна процентна похибка (Mean Absolute Percentage Error, MAPE) [55], що визначає похибку у відсотковому виразі відносно реальних значень:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{27}$$

Основною перевагою цього критерію є його простота інтерпретації, оскільки значення MAPE показує, на скільки відсотків у середньому прогноз моделі

відрізняється від реального значення. Проте MAPE має значне обмеження: він не визначений, коли фактичні значення дорівнюють нулю, а також може бути нестабільним у разі, коли реальні значення близькі до нуля.

Додатково до зазначених вище критеріїв, для оцінки якості регресійних моделей часто застосовується коефіцієнт детермінації — R-квадрат (Coefficient of Determination, R^2) [56]. Він вимірює частку варіативності вихідної змінної, що пояснюється моделлю:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}\left(y_i - \hat{y}\right)^2} \tag{28}$$

де $\hat{y}$ — середнє значення вихідних (фактичних) значень. Значення R^2 знаходиться у межах від 0 до 1, де 1 означає ідеальну точність прогнозу, а 0 — повну відсутність прогнозної здатності моделі. R^2 особливо корисний, коли необхідно порівняти кілька регресійних моделей між собою.

Таким чином, для точного й об'єктивного оцінювання якості регресійних моделей рекомендується використовувати декілька критеріїв одночасно, кожен з яких має свої переваги й обмеження. Залежно від конкретної прикладної задачі та особливостей даних, можна обрати ті критерії, які найбільш відповідають специфіці задачі прогнозування та забезпечують найбільш адекватне розуміння точності моделі. Це дозволить не лише забезпечити математичну обґрунтованість моделей, а й зробити їх придатними для реального практичного застосування.

Для задач прогнозування виживання і тривалості життя специфічним і важливим є використання спеціальних критеріїв, таких як індекс узгодженості або C-індекс (Concordance index) [57]. C-індекс є спеціалізованим критерієм, що використовується для оцінювання точності прогнозних моделей у задачах аналізу виживання (survival analysis), таких як прогнозування тривалості життя, часу до настання певної події (наприклад, смерті або рецидиву хвороби) та інших подій, що мають чітко визначений часовий аспект. Особливість цього критерію полягає в

тому, що він оцінює здатність моделі правильно передбачати порядок настання подій (послідовність подій) серед спостережуваних осіб.

Інтуїтивно, C-індекс визначає ймовірність того, що для будь-якої випадково вибраної пари спостережень модель правильно прогнозує, хто з двох осіб раніше зазнає події (наприклад, помре або захворіє). Іншими словами, якщо модель передбачила, що подія (наприклад, смерть) для одного пацієнта настане раніше, ніж для іншого, то C-індекс оцінює, наскільки часто такі прогнози є правильними відносно реальних результатів.

Математично C-індекс можна визначити наступним чином:

$$C = \frac{\text{Кількість правильно впорядкованих пар}}{\text{Загальна кількість можливих пар}} \quad (29)$$

Точніше, якщо ми маємо два пацієнти, для яких відомі фактичні часи настання події T^1 та T^2 та їхні прогнозовані ризики або прогнозовані часи $\check{T}^1$ та $\check{T}^2$, то пара спостережень вважається правильно впорядкованою, якщо $(\check{T}^1 < \check{T}^2)$ і при цьому реально $(T^1 < T^2)$, або навпаки, $(\check{T}^2 < \check{T}^1)$ і $(T^2 < T^1)$.

C-індекс може приймати значення у діапазоні від 0 до 1:

- C = 0,5 означає, що модель не має жодної дискримінаційної здатності та працює на рівні випадкового прогнозування (модель фактично «вгадує»).
- C = 1 означає ідеальний прогноз: модель завжди правильно передбачає послідовність подій.
- Значення C > 0,7 вважаються хорошими, а C > 0,8—0,9 — відмінними та високонадійними.

Перевагою C-індексу є його незалежність від абсолютних значень прогнозованих ризиків або прогнозованих часів подій. Це означає, що він оцінює лише здатність моделі правильно визначити відносний порядок, а не точні абсолютні значення часу. Таким чином, C-індекс є особливо цінним і поширеним в медицині, де правильне визначення відносного ризику має ключове значення для

клінічних рішень (наприклад, визначення пріоритету лікування, ранньої діагностики або інших заходів).

При цьому важливо зазначити, що C-індекс особливо ефективний для даних з цензуруванням (censored data), які дуже поширені в аналізі виживання. Цензурування означає, що для деяких спостережень подія ще не відбулася на момент завершення дослідження, і точний час настання події є невідомим. C-індекс дозволяє враховувати такі спостереження, що є його вагомою перевагою порівняно з іншими критеріями.

Водночас C-індекс також має свої обмеження. Зокрема, він не враховує ступінь відмінності між прогнозованими і фактичними часами подій, оцінюючи лише їхній порядок. Це означає, що навіть якщо модель суттєво помиляється в абсолютних значеннях часу, але правильно передбачає послідовність, C-індекс залишиться високим. Саме тому C-індекс часто рекомендується використовувати в комбінації з іншими критеріями (наприклад, RMSE або MAE), які дозволяють оцінити абсолютну величину похибок прогнозу.

Останнім часом популярними стали модифікації C-індексу, які враховують додаткові характеристики прогнозів (наприклад, зважений C-індекс — weighted C-index), що дозволяє підсилити чутливість до великих відхилень у прогнозах. Такі модифікації можуть бути корисними в певних задачах, коли важлива не тільки послідовність подій, а й розмір відхилення прогнозів від фактичних значень.

Таким чином, C-індекс є потужним, широко застосовуваним та практично цінним критерієм для оцінювання моделей прогнозування тривалості життя, виживання та інших часових подій. Його застосування дозволяє ефективно оцінювати дискримінаційну здатність моделей, забезпечуючи надійний інструмент для вибору та оптимізації прогнозних підходів у медицині, охороні здоров'я та інших сферах, де важливим є точне прогнозування послідовності настання подій.

Підсумовуючи аналіз критеріїв оцінювання точності прогнозних моделей, можна констатувати, що вибір правильних критеріїв є ключовим етапом при

побудові та впровадженні будь-яких моделей прогнозування, включаючи задачі прогнозування тривалості життя та ймовірності настання певних подій. Як показав проведений аналіз, критерії оцінювання точності суттєво відрізняються залежно від типу задачі: для моделей, що прогнозують ймовірність подій (класифікаційних), важливими є такі критерії, як точність (Accuracy), чутливість (Sensitivity або Recall), специфічність (Specificity), F1-міра (F1-score), а також загальні інтегральні критерії, такі як площа під кривою ROC (ROC-AUC). Ці показники дозволяють всебічно та об'єктивно оцінити здатність моделі правильно класифікувати випадки, враховуючи при цьому можливі наслідки хибних прогнозів та специфіку використання в реальних умовах.

У випадку регресійних задач, коли необхідно оцінити точність прогнозу конкретних числових значень, застосовуються інші критерії. Середня абсолютна похибка (MAE), середньоквадратична похибка (MSE), корінь середньоквадратичної похибки (RMSE), середня абсолютна процентна похибка (MAPE), а також коефіцієнт детермінації (R^2) забезпечують комплексну та ефективну оцінку моделі. При цьому MAE характеризується простотою інтерпретації та стабільністю щодо незначних відхилень, тоді як MSE та RMSE ефективні для виявлення та контролю значних помилок прогнозування. MAPE додатково дозволяє виразити похибку у зручних відсотках, що може бути особливо корисним для комунікації результатів користувачам.

Окрему роль відіграє специфічний критерій C-індекс (Concordance index), який є надзвичайно важливим у задачах аналізу виживання, таких як прогнозування тривалості життя чи часу настання певних подій. Він дозволяє об’єктивно оцінити здатність моделі правильно передбачати послідовність подій, що є критично важливим у медичних і демографічних застосуваннях. Водночас, C-індекс має свої обмеження, зокрема він не враховує абсолютні відхилення між прогнозованими та реальними значеннями, тому рекомендується використовувати його у поєднанні з іншими критеріями (такими як RMSE чи MAE), що дозволяє оцінити повну якість моделі.

Специфіка задач прогнозування тривалості життя створює додаткові складнощі для перевірки точності моделей, оскільки справжню перевірку можна здійснити лише ретроспективно у майбутньому. З цієї причини використовуються спеціальні підходи: ретроспективна валідація на історичних даних, використання проміжних сурогатних критеріїв (наприклад, медичних подій, що передують смерті), перехресна перевірка, тестування моделей на незалежних вибірках, а також перспективний моніторинг і регулярне калібрування моделей. Комплексне використання цих методів дозволяє отримати достовірну оцінку точності прогнозних моделей вже на ранніх етапах їх впровадження, навіть за умов неможливості негайної верифікації кінцевих прогнозів.

Правильний вибір та поєднання критеріїв оцінювання точності є фундаментально важливими для успішного практичного впровадження прогнозних моделей. Враховуючи специфіку задачі прогнозування тривалості життя та інших подій, доцільно використовувати комплексну стратегію, яка об’єднує як класичні, так і спеціалізовані критерії. Такий інтегральний підхід гарантує не лише об'єктивність і точність отриманих результатів, але й створює передумови для ефективного застосування таких моделей у реальній медичній, демографічній та соціальній практиці.

У даному дослідженні для оцінки якості прогнозних моделей застосовуються виключно регресійні метрики. Класифікаційні або інші метрики, які використовуються для оцінки моделей, що передбачають категоріальний вихід (наприклад, точність, чутливість, специфічність, F1-score), не можуть бути використані через особливість набору еталонних даних Держстату. Цей набор даних не передбачає класифікацію результатів на дискретні класи (ті, хто живуть більше та менше середнього): очікувана тривалість життя за певним віком й статтю завжди буде вищою, ніж заявлена середня тривалість життя за статтю. Але цей недолік даних нівелюється специфікою поставленої задачі, яка полягає саме у прогнозуванні кількісної величини (тривалості життя), що належить до категорії задач регресії.

Таким чином, вибір саме регресійних метрик, таких як MAE, MSE, RMSE, MAPE, R^2 та додатково C-індекс, є найбільш логічним та методологічно обґрунтованим для задачі прогнозування тривалості життя.

1.1.5 Аналіз інтеграції цифрових медичних та персональних даних для прогнозування

У сучасних персоналізованих моделях прогнозування тривалості життя ключову роль відіграє інтеграція цифрових даних різної природи — медичних, поведінкових, соціальних, фізіологічних тощо [58]. Сутність такої інтеграції полягає в об’єднанні цих різнорідних джерел інформації у єдину систему, що забезпечує цілісний цифровий портрет індивіда. Завдяки цьому модель може врахувати повний спектр факторів, які впливають на тривалість життя, від клінічних показників до соціально-демографічних детермінант.

Ще донедавна поєднання настільки різних даних було складним завданням, однак розвиток технологій великих даних та хмарних обчислень значною мірою розв’язав цю проблему. Накопичені за роки великі масиви інформації тепер можуть оброблятися комплексно, генеруючи нові знання для моделювання тривалості життя. Інтеграція даних у єдині інформаційні платформи дозволяє прогнозним моделям виявляти складні взаємозв’язки між різними чинниками. Зокрема, багато показників здоров’я є взаємозалежними, тому їхня інтерпретація має здійснюватися в сукупності: наприклад, зміни у фізіологічних параметрах слід аналізувати разом із пов’язаними поведінковими чи медичними даними, щоб правильно оцінити ризики. Такий синергетичний підхід забезпечує більш нюансований аналіз, ніж розрізнений розгляд окремих параметрів. Об’єднання медичних даних з даними про спосіб життя і соціальне середовище, по суті, трансформує прогнозування з загальнопопуляційного (усередненого) на індивідуалізоване, пристосоване до унікального профілю конкретної особи.

Іншими словами, інтегрована система даних робить можливим перехід від «середньостатистичного» прогнозу до персоналізованого, що відображає особисті характеристики людини. Результатом якісної інтеграції даних є суттєве підвищення точності, адаптивності й загальної ефективності прогнозів. Дослідники відзначають, що збільшення кількості релевантних змінних, врахованих у моделі, як правило, прямо покращує точність оцінки очікуваної тривалості життя.

Крім того, інтегровані цифрові моделі є динамічними: у міру надходження нових даних (наприклад, оновлення медичних показників чи інформації про спосіб життя) прогностичний алгоритм коригує свої оцінки, підвищуючи їхню точність і оперативність. Така адаптивність означає, що прогноз постійно налаштовується під актуальний стан пацієнта, залишаючись достовірним на кожному етапі спостереження. Отже, інтеграція різнопланових цифрових даних виступає необхідною передумовою створення високоточних та гнучких моделей прогнозування тривалості життя, підвищуючи їхню практичну цінність у персоналізованій медицині.

Одним з перспективних напрямів підвищення точності персоналізованого прогнозування тривалості життя є інтеграція медичних даних, отриманих за допомогою носимих пристроїв, зокрема, електрокардіографічних (ЕКГ) сигналів. Як свідчать сучасні дослідження, зокрема розробка комбінованих підходів аналізу ЕКГ, заснованих на штучному інтелекті, дозволяє виявляти відхилення як від персональних, так і від загальнопопуляційних норм, що суттєво підвищує точність та актуальність прогнозів щодо здоров'я і тривалості життя

Дані з медичних пристроїв (носимі трекери, розумні годинники, кардіомонітори тощо) забезпечують безпрецедентний доступ до динамічних даних про повсякденну активність і фізіологічні параметри пацієнтів. Ці дані істотно збагачують моделі прогнозування, адже відображають поведінкові та середовищні фактори, які важко зафіксувати під час епізодичних візитів до лікаря. Об’єктивні показники фізичної активності, серцевого ритму, варіабельності пульсу,

артеріального тиску та якості сну суттєво підвищують точність прогнозів, дозволяючи врахувати реальний стан і спосіб життя пацієнта.

Максимальна точність прогнозування досягається при комплексному використанні всіх зазначених джерел даних. Кожен тип інформації висвітлює різні аспекти здоров'я, тому їхнє поєднання в єдиній моделі забезпечує повну картину і дозволяє врахувати складні взаємозв'язки між факторами. Вже існують інтегровані платформи (наприклад, UK Biobank, MIMIC-III [59]), які успішно об'єднують різні типи медичних даних, демонструючи ефективність такого підходу. Розвиток таких платформ та методів аналізу буде сприяти подальшому підвищенню точності прогнозування індивідуальної тривалості життя, створюючи нові можливості для прецизійної медицини та охорони здоров'я.

У контексті прогнозування тривалості життя дедалі більшого значення набувають дані, отримані за допомогою носимих пристроїв, мобільних застосунків, домашніх медичних приладів та мобільних опитувальників. Ці технології забезпечують безперервний моніторинг і збір інформації про стан здоров'я та поведінкові фактори у реальних умовах, що суттєво розширює набір параметрів для аналізу. Зокрема, носимі пристрої (наприклад, розумні годинники та фітнес-браслети) фіксують показники фізичної активності (кількість кроків, рівень активності), частоту серцевих скорочень, параметри сну (тривалість та якість), варіабельність ритму серця та інші фізіологічні сигнали.

Цифрові дані про спосіб життя та здоров'я вважаються перспективними для прогнозних моделей тривалості життя з кількох причин. По-перше, вони відкривають доступ до нових змінних, які раніше було важко або неможливо отримати у таких масштабах. По-друге, ці дані є більш точною та якісною альтернативою опитуванням: багато параметрів, що раніше оцінювалися суб'єктивно, тепер фіксуються об'єктивно сенсорами. По-третє, інформація збирається з високою частотою в реальному часі, що дозволяє виявляти навіть короткострокові зміни стану здоров'я та оперативно оновлювати прогноз. Завдяки цьому моделі можуть реагувати на динамічні зміни — приміром, погіршення

показників сну чи падіння рівня активності одразу впливатиме на оцінку ризику. Очікується, що залучення таких деталізованих даних сприятиме глибшому розумінню детермінант довголіття та вдосконаленню індивідуальних прогнозів тривалості життя.

Збирання та обробка таких даних здійснюється через інтегровані цифрові екосистеми та спеціалізовані платформи. Носимі пристрої зазвичай підключені до смартфонів і передають дані до централізованих сервісів: прикладами є Apple HealthKit [60] та Google Fit [61], які агрегують інформацію з різних сенсорів і застосунків на телефоні. Ці платформи накопичують широкий спектр показників — від числа кроків і частоти пульсу до калорійності раціону.

Однак використання таких даних супроводжується низкою викликів. Значний обсяг і різнорідність інформації потребують складної передобробки та інтеграції. Виникають проблеми стандартизації та інтероперабельності різних форматів даних. Необхідно також враховувати можливі похибки вимірювань і пропуски в даних. Для зберігання та аналізу великого обсягу високочастотних записів потрібні потужні обчислювальні ресурси та складні алгоритми обробки. Критичними є питання конфіденційності, захисту даних та уникнення систематичних зміщень вибірки.

Таким чином, цифрові дані з носимих пристроїв, мобільних застосунків та домашніх медичних сенсорів відкривають нові можливості для прогнозування тривалості життя, сприяють точнішій оцінці індивідуального ризику і, відповідно, покращенню якості медичних прогнозів. Подальший розвиток методів збору та обробки цих даних є перспективним напрямком для персоналізованої медицини та управління здоров’ям населення.

Для точного прогнозування тривалості життя необхідно цілісно враховувати медичні, поведінкові, соціальні та генетичні фактори. Сучасні системи прагнуть об’єднати ці різнорідні джерела інформації в єдину аналітичну платформу. Комплексне поєднання клінічних даних (електронних медичних записів,

результатів обстежень), поведінкових показників (фізична активність, харчові звички), соціально-демографічних характеристик (освіта, дохід, місце проживання) та генетичної інформації дозволяє врахувати всі ключові детермінанти здоров'я. Базовий принцип інтеграції — об'єднання різних типів цифрових даних про пацієнта у єдиному середовищі для всебічного аналізу. Таке об'єднання потребує уніфікованих моделей даних і суворого дотримання стандартів сумісності.

Захист персональних даних є критичним аспектом інтеграції цифрових медичних даних. Для забезпечення конфіденційності використовуються методи анонімізації та псевдонімізації [62]. Важливу роль може відігравати федеративне навчання (federated learning) [63], яке дозволяє аналізувати дані без їх централізованого збору, зменшуючи ризики втрати приватності. У поточній версії платформи федеративне навчання не реалізоване, але може бути перспективним напрямком для подальшого вдосконалення технології прогнозування з метою додаткового захисту конфіденційності користувачів.

Реалізація комплексної інтеграції даних стикається з низкою труднощів. Інтероперабельність між системами залишається складною задачею через наявність різних форматів і стандартів [64]. Фрагментованість даних у медичних установах, нестача повної стандартизації, а також проблеми якості інформації (неповнота, помилки, неточності) є значними перешкодами для створення цілісної системи [65]. Семантичні розриви виникають через неоднозначність у трактуванні одних і тих самих параметрів, що вимагає використання додаткових методів зіставлення і узгодження [66]. Організаційні та етичні бар'єри (обмеження обміну інформацією між установами) також перешкоджають швидкій інтеграції даних [67].

Важливий потенціал для подальшого покращення якості прогнозування мають також дані з популярних українських сервісів, таких як Helsi (цифрова медична карта, записи лікарів, медичні огляди та лабораторні дослідження) [68] та Дія (інформація щодо віку, статі, офіційних реєстрацій та інших соціальних та юридичних параметрів громадян) [69]. Включення цих даних може значно

розширити інформаційну базу моделей, підвищити їхню персоналізацію та релевантність прогнозів у локальному контексті України.

Таким чином, інтеграція даних з ЕКГ, HealthKit, Google Fit, а також потенційне використання Helsi та Дія створює широкі можливості для суттєвого покращення точності та якості прогнозування тривалості життя за допомогою сучасних моделей машинного навчання. Сучасні технології інтеграції та аналізу цифрових даних створюють потужну основу для побудови високоточних і глибоко персоналізованих моделей прогнозування тривалості життя, забезпечуючи комплексне розуміння стану здоров'я людини та ефективне управління ним у довгостроковій перспективі.

1.2 Концептуальні основи інтегрованої моделі прогнозування тривалості життя

На сучасному етапі розвитку демографічних і медичних досліджень очевидним є той факт, що прогнозування тривалості життя вимагає комплексного підходу, який би враховував множину взаємопов'язаних факторів [70, 71]. Традиційно демографічні, соціальні та медичні чинники розглядалися окремо в різних моделях, що дозволяло лише частково враховувати їхній сумарний вплив. Однак реальне життя демонструє значно складнішу взаємодію між цими групами факторів. Саме тому виникає нагальна необхідність у створенні інтегрованої моделі, яка була б здатна охоплювати та аналізувати демографічні, соціальні та медичні фактори у взаємодії, забезпечуючи таким чином більш точне та реалістичне прогнозування тривалості життя.

Інтегрована модель прогнозування має низку суттєвих переваг у порівнянні з окремими моделями, що використовують лише одну категорію чинників [72]. По-перше, демографічні фактори (вік, стать, регіон проживання тощо) створюють базову структуру оцінювання ризиків смертності, але не дозволяють враховувати індивідуальні особливості стану здоров'я або особливості життєвих умов [73, 74]. По-друге, медичні показники (хронічні захворювання, лабораторні аналізи,

генетична схильність) характеризують лише стан здоров’я конкретної людини на певний момент часу і можуть змінюватися залежно від соціальних умов чи поведінкових звичок [75, 76]. Нарешті, соціальні фактори (економічний статус, освіта, доступ до медичних послуг, поведінкові звички) суттєво впливають на стиль життя та можливості щодо профілактики захворювань [77, 78]. Водночас окреме використання цих груп факторів не дозволяє оцінити їх комплексний вплив і взаємне посилення або нівелювання ризиків. Інтегрований підхід, що поєднує всі ці категорії факторів, дозволяє створити повнішу та багатовимірну картину факторів ризику і таким чином підвищує точність прогнозування тривалості життя.

Виходячи з цього, базовими принципами побудови інтегрованої моделі прогнозування виступають багатофакторність, гнучкість і адаптивність. Багатофакторність моделі полягає в її здатності враховувати одночасно широкий спектр різнорідних показників, демографічних, медичних, соціально-економічних, поведінкових і генетичних, що дає змогу оцінювати ризики смертності з високою точністю та персоналізацією. Це дозволяє враховувати не лише прямий вплив окремих чинників, але й їхні непрямі взаємодії, які неможливо адекватно змоделювати, розглядаючи фактори ізольовано.

Гнучкість інтегрованої моделі передбачає можливість легко адаптувати її до змін у зовнішньому середовищі чи появи нових даних. Така модель повинна бути побудована на принципах модульності, що дозволяє включати нові фактори або виключати ті, які втратили свою актуальність чи інформативність. Це особливо важливо в умовах швидких соціально-економічних та епідеміологічних змін, як війни [79], пандемії COVID-19 [80] чи інших кризових ситуацій, які можуть суттєво змінювати профіль ризиків у популяції.

Адаптивність є ще одним важливим принципом побудови інтегрованої моделі. Вона полягає в можливості автоматично налаштовувати параметри моделі на основі накопичених даних та досвіду її застосування. Завдяки використанню сучасних методів машинного навчання, адаптивні алгоритми здатні постійно уточнювати коефіцієнти впливу факторів на основі поточних даних, що дозволяє

моделі залишатися актуальною навіть за умов непередбачуваних змін у зовнішньому середовищі. Крім того, адаптивність моделі забезпечує її персоналізацію — алгоритм може динамічно враховувати індивідуальні зміни в поведінці, здоров'ї чи соціальних умовах конкретної людини.

Загальний методологічний підхід до інтеграції даних у моделі прогнозування передбачає кілька етапів. На першому етапі проводиться визначення та вибір найбільш вагомих факторів з кожної категорії, що мають доведений вплив на тривалість життя. На другому етапі ці фактори формалізуються і перетворюються в кількісні змінні, які можуть бути введені в математичні або статистичні моделі. Третій етап передбачає визначення взаємозв'язків і взаємодій між окремими факторами за допомогою кореляційного аналізу, факторного аналізу або інших методів виявлення прихованих взаємозалежностей. Після цього розробляється безпосередньо математична структура інтегрованої моделі, яка дозволяє враховувати сумарний вплив відібраних факторів. Нарешті, інтегрована модель проходить етап валідації, на якому її параметри уточнюються і перевіряються на реальних вибірках даних з використанням методів перехресної валідації або інших підходів, що підтверджують її надійність та точність.

В якості основи запропоновано комбінований підхід: інтегрована факторна модель прогнозування тривалості життя, яка враховує персоналізовані коефіцієнти впливу демографічних, соціальних та медичних факторів.

Таким чином, запропонована концепція інтегрованої моделі прогнозування тривалості життя ґрунтується на усвідомленні багатогранності факторів, що визначають стан здоров'я та виживаність людей, і дозволяє здійснювати більш точні, гнучкі та персоналізовані прогнози. Впровадження такого підходу відкриває широкі можливості для практичного застосування в сфері охорони здоров'я, соціальної політики та індивідуального медичного консультування, сприяючи покращенню управління здоров'ям на рівні як окремих людей, так і суспільства загалом.

1.3 Вибір і формалізація демографічних, соціальних та медичних факторів

Вибір і формалізація факторів, які включаються в інтегровану модель прогнозування тривалості життя, є критично важливим етапом дослідження. Від правильності та повноти цього вибору безпосередньо залежить якість та точність прогнозів, які будуть отримані в результаті роботи моделі. Саме тому необхідно чітко обґрунтувати логіку включення саме демографічних, соціальних (Табл. 2) та медичних факторів як ключових груп параметрів, які визначають прогноз тривалості життя.

Таблиця 2. Перелік демографічних, звичкових та соціо-демографічних параметрів впливу

Назва параметру впливу	Тип	Додатковий опис
Вік	Неперервний	Діапазон в межах 18-110 років
Стать	Дихотомічний	Чоловіча/жіноча
Індекс маси тіла (ІМТ)	Неперервний	Розраховується на основі зросту та маси тіла
Фізична активність	Категоріальний	За часом (відсутність, мінімальна, оптимальна) та типом (висока, середня, низька) активності
Споживання здорової їжі	Категоріальний	За рівнем споживання — 4 градації
Захворювання та медичні стани	Дихотомічний	Наявність захворювань та медичних станів: всього 14 згідно Табл. 5
Взаємодія з лікарями	Категоріальний	Звернення до сімейного лікаря, виклик лікарів додому, стаціонарне лікування — 4 градації

Назва параметру впливу	Тип	Додатковий опис
Куріння	Категоріальний	Частота куріння — 5 градацій, вік початку, вік кидання.
Вживання алкоголю	Категоріальний	Частота вживання — 6 градацій
Освіта	Категоріальний	6 рівнів
Щомісячний дохід	Категоріальний	5 рівнів
Статус внутрішньо переміщеної особи	Дихотомічний	Так/ні

Демографічні фактори (вік, стать, територія проживання тощо) є фундаментальними і визначають базові тенденції смертності на популяційному рівні. Вони дають змогу враховувати загальні вікові і статеві закономірності та регіональні особливості, що мають безпосередній вплив на стан здоров’я та ймовірність виживання. Саме демографічні фактори створюють структурну основу моделі, яка забезпечує загальні рамки оцінки індивідуальних ризиків і є відправною точкою для подальшого уточнення прогнозу з використанням додаткових показників.

Соціальні фактори були обрані для включення в модель завдяки їхньому значному впливу на умови життя, доступність ресурсів та медичної допомоги, поведінкові патерни і, відповідно, на загальний рівень здоров'я і тривалість життя. До цієї групи належать такі чинники, як економічний статус, рівень освіти, характер зайнятості, соціальна підтримка та середовище проживання. Соціальні фактори дозволяють поглибити модель, надаючи їй змогу відображати суттєві відмінності в умовах життя людей і враховувати, як ці умови можуть модифікувати ризики розвитку хвороб або їх наслідки.

Коефіцієнти впливу (Табл. 3-4) вказують на те, як кожен конкретний параметр чи стан змінює очікувану тривалість життя людини, зменшуючи або збільшуючи її. Позитивні коефіцієнти відповідають факторам, які потенційно збільшують

прогнозовану тривалість життя (наприклад, високий рівень фізичної активності, відсутність шкідливих звичок), тоді як негативні коефіцієнти вказують на фактори ризику, що знижують очікувану тривалість життя (наприклад, паління, хронічні захворювання, низький рівень доходу). Коефіцієнти впливу факторів на ймовірність виживання були визначені на основі офіційних статистичних даних Державної служби статистики України, а також за допомогою інтерпретації даних та екстраполяції моделей ризику MPoRT (Mortality Population Risk Tool). Останні дозволили розширити аналітичні можливості, уточнити коефіцієнти для окремих факторів ризику та здійснити прогнозування для вікових груп старше 100 років, які не охоплені даними офіційної статистики.

Таблиця 3. Перелік значень звичкових та соціо-демографічних параметрів впливу

Параметр	Градації та коефіцієнти впливу					
Споживання здорової їжі	Відсутнє	Низьке	Помірне	Опти-мальне		
	-390	-70	+420	+910		
Фізична активність	відсутня					
	-810					
Фактор ІМТ	0.5	0.75	1.25	1.5	2.0	2.5
	+1700	+200	-300	-380	-660	-980
Виклики швидкої допомоги протягом року	відсутні	один	час від часу	часті		
	+800	-30	-300	-700		
Госпіталізації протягом року	відсутні	один раз	час від часу	часті		
	+800	-30	-300	-700		

Медична допомога сімейного лікаря протягом року	відсутня	один раз	час від часу	часта		
	+800	-30	-300	-700		
Рівень освіти	відсутня	неповна середня	сер. спе-ціальна	бака-лавр	магістр	доктор
	-190	-40	-30	+170	+190	+400
Фізична активність високої інтенсивності	відсутня	міні-маль-на	оптима-льна			
	-120	+110	+310			
Фізична активність середньої інтенсивності	відсутня	міні-маль-на	оптима-льна			
	-140	+170	+480			
Фізична активність низької інтенсивності	відсутня	міні-маль-на	оптима-льна			
	-220	+270	+720			
Відсутність захворювань	+2700					
Вживання алкоголю	відсутнє	ніколи не вжи-вав	час від часу	часте	Більше 3 пляшок міцного алкоголю на тиждень	7

	+970	+270	-30	-180	-480	-1310
Куріння	ніколи не курив	в минулому	час від часу	менше пачки	пачка і більше	
				цигарок на добу		
	+1590	-110	-170	-430	-1870	
				коефіцієнт куріння		
Дохід	за межею бідності	низький	нижче середнього	середній	вище середнього	
	-170	-120	-70	-10	+240	
Статус ВПО	Наявність					
	-120					

Окремо відзначений вплив по регіонах (Табл. 4). Коефіцієнти впливу в залежності від регіону були розраховані на підставі офіційних даних Державної служби статистики України [81]. Ці дані охоплюють населення України за винятком тимчасово окупованих територій Автономної Республіки Крим, м. Севастополя та частини Донецької та Луганської областей.

Таблиця 4. Перелік коефіцієнтів впливу в залежності від регіону респондента

Область	Вплив	Область	Вплив	Область	Вплив
Вінницька	890	Кіровоградська	-1240	Харківська	-240
Волинська	-60	Львівська	1070	Херсонська	-1300
Дніпропетровська	-1180	Миколаївська	-310	Хмельницька	450
Житомирська	-1630	Одеська	-40	Черкаська	420

Закарпатська	-880	Полтавська	-190	Чернівецька	1460
Запорізька	-520	Рівненська	-110	Чернігівська	-1270
Івано-Франківська	1480	Сумська	-130	м.Київ	2150
Київська	-1630	Тернопільська	1700		

Таким чином, наведені таблиці забезпечують чітке уявлення про кількісну оцінку впливу кожного фактору, що дозволяє персоналізовано та точно прогнозувати тривалість життя людини з урахуванням її індивідуальних характеристик.

Параметри медичного стану або захворювань (Табл. 5) було вирішено включити до моделі через їх прямий зв'язок зі станом здоров'я та смертністю. Ця категорія факторів охоплює як хронічні захворювання та патологічні стани (серцево-судинні захворювання, діабет, онкологічні захворювання тощо), так і фізіологічні маркери здоров’я (рівень артеріального тиску, холестерину, маркери запалення тощо). Медичні фактори надають моделі клінічної релевантності та здатності враховувати індивідуальні особливості стану здоров’я кожної людини. Важливим аспектом є те, що ця група факторів найбільш чутлива до змін у часі, що дозволяє моделі бути адаптивною до зміни стану здоров’я особи.

Таблиця 5. Перелік параметрів медичного стану або захворювань

Назва параметру впливу	Значення коефіцієнту впливу
Діабет	-670 * фактор ІМТ
Захворювання серця	-410 * фактор ІМТ
Інсульт	-400 * фактор ІМТ
Гіпертонічна хвороба	-150 * фактор ІМТ

Деменція	-710
Онкологічні захворювання	-1290
Хронічні обструктивні захворювання легень (ХОЗЛ)	-390
Епілепсія	-660
Депресія	-360
Туберкульоз	-650
ВІЛ/СНІД	-700
Посттравматичний стресовий розлад (ПТСР)	-360
COVID-19 помірного перебігу	-370
COVID-19 тяжкого перебігу	-700

Процес формалізації цих факторів передбачає їх кількісне або якісне кодування у формі, придатній для статистичного або математичного аналізу. Це необхідно для того, щоб інтегрована модель могла враховувати не лише наявність того чи іншого фактора, але і його ступінь вираженості чи рівень впливу. Формалізація здійснюється через визначення стандартних шкал, категорій чи бінарних індикаторів, які дозволяють однозначно інтерпретувати значення кожного фактора в межах прогнозної моделі.

Соціальні, поведінкові та економічні дані відіграють ключову роль у персоналізованому прогнозуванні тривалості життя. Загальновизнано, що самі лише біологічні та медичні показники не дають повної картини про майбутню тривалість життя людини. На індивідуальний прогноз суттєво впливають умови та стиль життя, психологічний стан, рівень матеріального добробуту, а також соціальне оточення й підтримка.

До соціальних чинників належить соціальне середовище, в якому живе індивід: сім’я, спільнота, рівень освіти, доступ до медичних послуг, якість інфраструктури та система соціального захисту. Наприклад, наявність підтримки з боку родини і друзів, безпечні житлові умови та доступність якісної охорони здоров’я здатні підвищувати прогнозовану тривалість життя. Водночас соціальна ізоляція, хронічний стрес через нестабільність в суспільстві або брак медичних послуг можуть призводити до погіршення здоров’я і скорочення життя. Економічні аспекти тісно переплітаються із соціальними: рівень доходу, зайнятості та життєві можливості визначають якість харчування, умови праці й проживання, що в сукупності впливає на ризики захворювань і смертності. Так, бідність і безробіття часто асоціюються з гіршим станом здоров’я через обмежений доступ до ресурсів і медичної допомоги, тоді як фінансова стабільність і вищий соціально-економічний статус створюють сприятливі умови для довголіття.

Поведінкові (або спосіб життя) фактори включають індивідуальні звички та щоденну поведінку, які безпосередньо впливають на здоров’я. Серед них — харчові вподобання, рівень фізичної активності, режим сну, вживання тютюну, алкоголю чи інших психоактивних речовин, дотримання медичних рекомендацій та профілактичних заходів. Наприклад, здорове харчування і регулярні фізичні вправи сприяють зміцненню серцево-судинної системи та зниженню ризику хронічних захворювань, що в результаті подовжує життя. Натомість тривале паління, зловживання алкоголем або сидячий спосіб життя статистично пов’язані зі зменшенням очікуваної тривалості життя. Важливим поведінковим аспектом є також дотримання медичних приписів: людина, яка відповідально приймає призначені ліки та регулярно проходить медичні огляди, має вищі шанси на раннє виявлення і лікування небезпечних станів, ніж та, що ігнорує симптоми або рекомендації лікарів.

Окремо слід відзначити психоемоційні чинники, які хоча й часто розглядаються в межах поведінкових або соціальних, заслуговують на особливу увагу. Психічне здоров’я та емоційний добробут — депресія, тривожність, рівень

стресу, відчуття щастя та життєвої задоволеності — мають доведений вплив на фізичне здоров'я і довголіття. Хронічний психологічний стрес чи травматичний досвід можуть спричинити фізіологічні зміни (наприклад, гормональні збої, постійне запалення), що підвищують ризик серцево-судинних захворювань, імунних розладів і, відповідно, передчасної смерті. Натомість позитивний психоемоційний стан, наявність підтримки та навички справлятися зі стресом асоціюються з кращим прогнозом здоров'я. Таким чином, включення індикаторів психічного здоров'я (рівень стресу, емоційне благополуччя) до моделі прогнозування додає важливий вимір персоніфікації оцінки тривалості життя. Значення для персоналізованого прогнозування. У контексті персоналізованого прогнозування тривалості життя врахування соціальних, поведінкових та економічних факторів є необхідним для досягнення точної та індивідуально релевантної оцінки. Двоє людей однієї статі й віку можуть мати цілком різні прогнози тривалості життя саме через відмінності у способі життя та соціально-економічних умовах. Медичні дані (наприклад, діагнози, результати аналізів, генетичні маркери) відображають поточний стан здоров'я та схильність до певних захворювань. Проте без інформації про те, в яких умовах живе пацієнт і як він поводиться щодня, прогноз може бути неповним. Соціально-економічні дані дозволяють моделі "зрозуміти", чи має людина належний доступ до лікування і здорового середовища, а поведінкові — чи підтримує вона здоровий стиль життя або наражає себе на додаткові ризики. Персоналізація прогнозу означає, що модель враховує унікальне поєднання цих чинників в кожної окремої особи.

Таким чином, інтеграція соціальних, поведінкових та економічних показників поряд із медичними даними дає змогу моделювати більш реалістичний індивідуальний сценарій старіння і тривалості життя. Поєднання з медичними та цифровими даними. Соціально-поведінкові та економічні показники найкраще працюють у прогнозній моделі в поєднанні з медичними і цифровими даними, оскільки всі ці джерела інформації взаємодоповнюють одне одного. Медичні дані забезпечують клінічну основу прогнозу — інформацію про хвороби, фізіологічні

параметри, стан органів і систем. Цифрові дані, отримані з носимих пристроїв, смартфонів чи електронних медичних записів, можуть безперервно фіксувати аспекти поведінки і здоров'я (наприклад, кількість кроків, частоту серцевих скорочень, якість сну), які раніше було важко врахувати. Ці цифрові сліди здоров'я та активності фактично конкретизують поведінкові фактори в режимі реального часу. Коли всі три категорії даних об'єднані, модель прогнозування отримує цілісний огляд життя людини.

Таким чином, вибір і формалізація демографічних, соціальних та медичних факторів для подальшого використання в інтегрованій моделі прогнозування тривалості життя ґрунтується на необхідності врахування комплексності людського життя та здоров'я. Саме ці три категорії факторів забезпечують багатовимірний підхід до оцінки індивідуальних ризиків, що дозволяє побудувати модель, яка не тільки є високоточною, але й здатна адаптуватися до змін у реальних умовах життя людини, що є надзвичайно важливим у контексті сучасних соціально-економічних та медичних викликів.

1.4 Математична модель та алгоритми прогнозування

1.4.1 Розробка математичної моделі прогнозування з використанням коефіцієнтів впливу факторів на ймовірність виживання

Запропонована інтегрована математична модель прогнозування тривалості життя ґрунтується на персоналізованому підході та враховує широкий спектр демографічних, медичних та соціально-поведінкових факторів, що є особливо актуальними для українського населення. На відміну від класичних і сучасних зарубіжних моделей, запропонована модель «Калькулятора життя» використовує єдиний концептуальний підхід для визначення коефіцієнтів впливу, що застосовуються до всіх вікових категорій.

Формально модель визначається наступною математичною залежністю:

$$P_i = B_{si} + K_{si} \cdot \left(\sum_{j=1}^{N_{nd}} P_{nd_j} + C \cdot N_d \sum_{k=1}^{N_d} P_{d_k} \right) \quad (31)$$

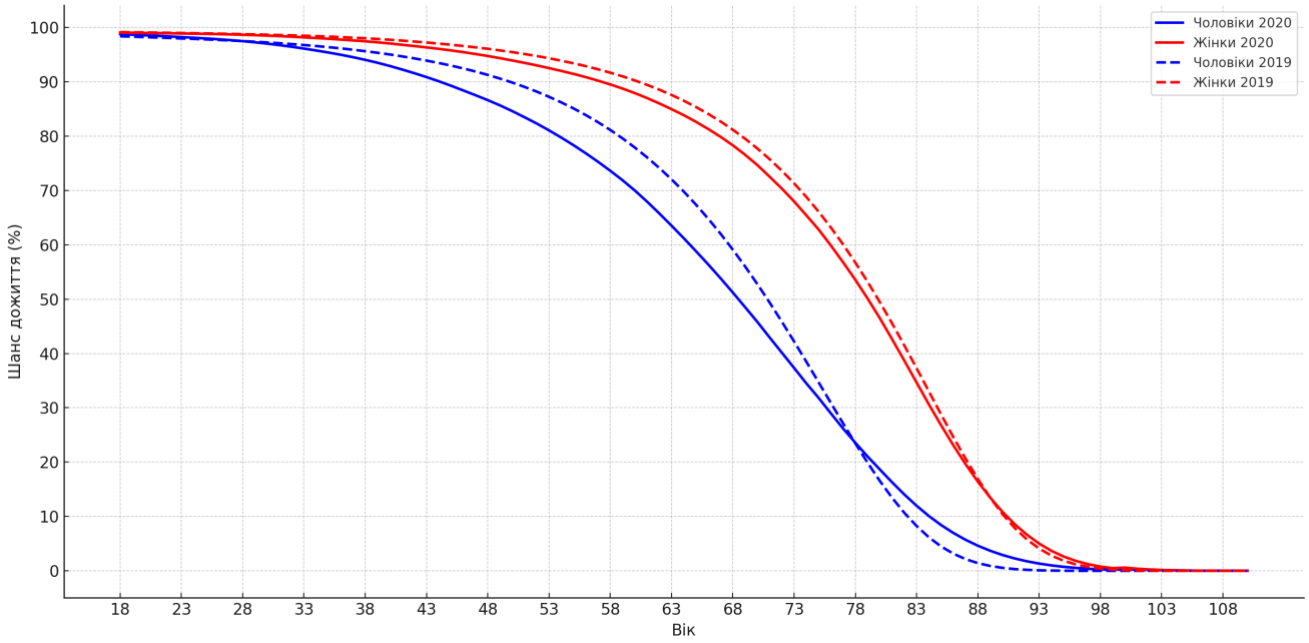
де P_i — персональна ймовірність виживання у віці i, B_{si} — базова ймовірність виживання за статтю s та віком i, K_{si} — коефіцієнт впливу факторів за статтю s та віком i, P_{nd_j} — вплив окремого недіагностичного фактора, C — коефіцієнт прихованих ризиків (КПР), N_d — кількість діагностичних факторів, N_{nd} — кількість недіагностичних факторів, P_{d_k} — вплив окремого діагностичного фактора.

Запропонована модель має унікальну розраховану базову ймовірність виживання для українського населення (Рис. 1) та коефіцієнт впливу для кожного віку та статі (Рис. 2). Коефіцієнти впливу коригують абсолютні значення персональних факторів впливу окремо для кожного віку та статі. Мінімальний вік для визначення ймовірності виживання становить 18 років, а максимальний становить 110 років. У даному дослідженні використовуються дві базові таблиці смертності, розроблені спеціально для популяції України. Перша базова таблиця побудована на підставі даних ВООЗ за 2019 рік (далі — «Life Calculator 2019» або «LC19»). Друга базова таблиця створена на основі даних Державної служби статистики України за 2020 рік (далі — «Life Calculator 2020» або «LC20»).

Наявність двох базових таблиць є важливою для забезпечення порівняльного аналізу та підвищення валідності отриманих результатів, оскільки вони дозволяють враховувати специфіку українських демографічних і медичних умов на основі різних джерел та періодів часу. Основні дані для розрахунків базової ймовірності виживання 2019 року (LC19) були взяті зі статистики Глобальних оцінок здоров'я (Global Health Estimates, GHE) ВООЗ щодо середньої очікуваної тривалості життя в Україні, яка на момент валідації за даними 2019 року становила 68 років для чоловіків та 78 років для жінок [82]. Базові ймовірності виживання за 2020 рік (LC20) були розраховані на підставі офіційних даних Державної служби статистики України, викладених у статистичному збірнику «Таблиці народжуваності,

смертності та середньої очікуваної тривалості життя за 2020 рік». Зокрема, були використані повні таблиці смертності та середньої очікуваної тривалості життя населення України за статтю та віковими групами, побудовані відповідно до затвердженої Держстатом методики (затверджено наказом Держкомстату від 26.04.2011 № 103 у редакції наказу Держстату від 18.02.2013 № 52 [83]). Оскільки у вихідних даних Державної служби статистики максимальний вік для розрахунку ймовірностей виживання обмежується 100 роками, у даному дослідженні ці значення були екстрапольовані до 110 років для забезпечення більш повної оцінки тривалості життя та прогнозування для осіб старших вікових груп.

Рисунок 1. Відображення розроблених таблиць базової ймовірності дожиття за віком та статтю 2019 та 2020 років (LC19 та LC20)

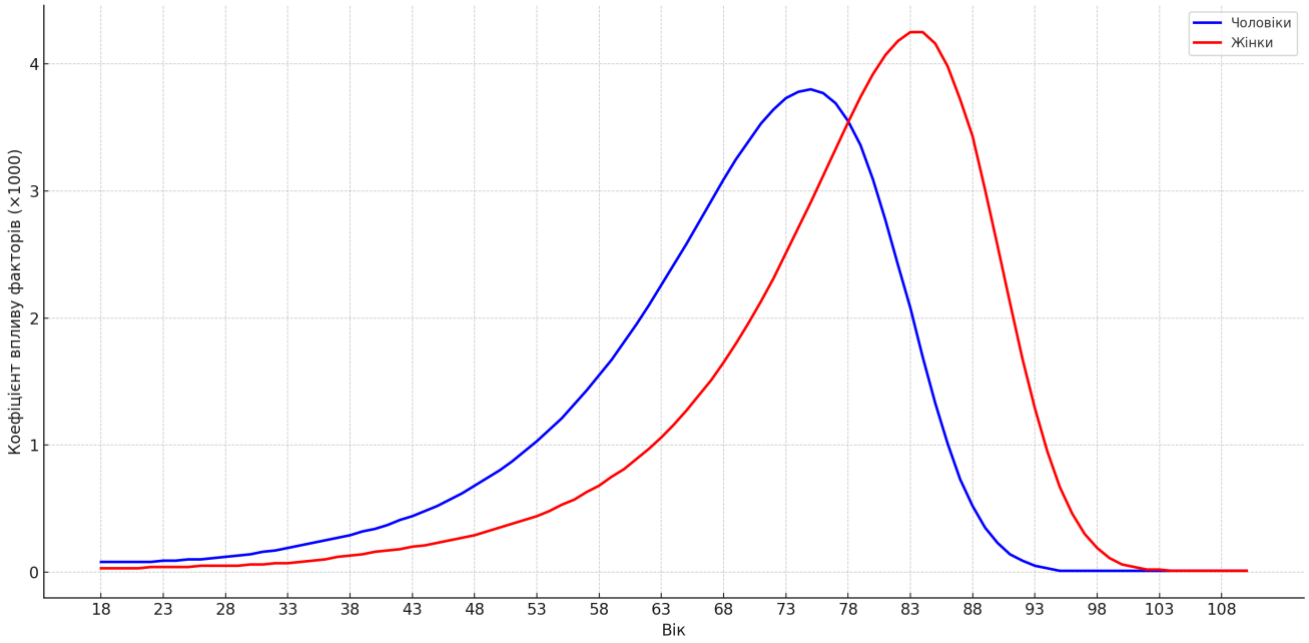


Таблиця коефіцієнтів впливу факторів за статтю та віком (Рис. 2) були розраховані на підставі поєднаних, валідованих та екстрапольованих на українську популяцію даних з MPoRT. Рішення використати міжнародно визнану модель MPoRT, як базовий орієнтир, було прийнято оскільки в Україні відсутні достатньо масштабні й деталізовані дані, які б дозволили надійно визначити власні національні коефіцієнти впливу факторів ризику.

Обрання саме цієї моделі обумовлено її широким визнанням, відкритістю методології та можливістю адаптації до місцевих умов. Для підвищення валідності застосування цієї моделі в українському контексті була виконана адаптація коефіцієнтів шляхом їх калібрування на наявних обмежених локальних даних та даних з офіційних українських демографічних і медичних звітів (зокрема даних Держстату).

Водночас усвідомлені методологічні обмеження такого підходу, оскільки соціально-демографічні умови, система охорони здоров'я, поведінкові та культурні особливості населення України можуть суттєво відрізнятися від канадських реалій. Тому отримані коефіцієнти слід розглядати як початкові орієнтири, які можуть уточнюватися з накопиченням локальних даних в процесі довгострокового використання розробленої технології.

Рисунок 2. Відображення таблиці коефіцієнтів впливу факторів за статтю та віком



Розроблена методика передбачає поділ факторів впливу на дві умовні групи: діагностичні (пов'язані з конкретними захворюваннями) та недіагностичні (поведінкові, соціально-економічні). Цей поділ обумовлений різницею у механізмах

їх впливу на ймовірність виживання. Діагностичні фактори мають прямий клінічний вплив, що пропорційно зростає з кількістю захворювань. Недіагностичні ж фактори, такі як куріння, фізична активність, освіта та рівень доходу, визначають загальний фон ризиків через зміну поведінки або умов життя людини.

Значення коефіцієнта прихованих ризиків (КПР) для діагностичних факторів було визначене на етапі розробки моделі та становлює 0,61. Даний коефіцієнт є узагальненим параметром, що кількісно оцінює вплив групи діагностичних факторів, які зазвичай складно точно виміряти. Він був отриманий шляхом проведення попереднього аналізу доступних міжнародних статистичних даних і використання методу калібрування коефіцієнтів моделі шляхом оптимізації похибки прогнозу (методом найменших квадратів) для загальнопопуляційних вибірок. Таким чином, КПР зі значенням 0,61 слугує апріорним параметром, який встановлює початкове співвідношення між впливом недіагностичних та діагностичних факторів у загальній прогнозній моделі

Кожен фактор, включений у модель, проходить процес формалізації та кількісного оцінювання у вигляді коефіцієнта, який розраховується на основі медико-статистичних та епідеміологічних даних. Запропонована модель прогнозування враховує не лише окремі фактори ризику незалежно один від одного, але й передбачає їх взаємну модифікацію. Зокрема, вплив факторів, пов’язаних із такими захворюваннями як цукровий діабет, серцеві захворювання, інсульт і гіпертонія, у моделі додатково модифікується залежно від індексу маси тіла (ІМТ).

Індекс маси тіла розраховується за стандартною формулою на основі персональних даних респондента (зросту та ваги) [84] і виступає мультиплікативним модифікатором, який посилює або послаблює сумарний вплив вказаних захворювань. Таким чином, модель не є простою лінійною комбінацією коефіцієнтів, а враховує складніші взаємодії між факторами ризику та фізичним станом респондентів. Це дозволяє врахувати індивідуальний ступінь ризику,

пов'язаний із зайвою вагою або її дефіцитом, оскільки фактор ІМТ змінюється в діапазоні від 0,5 до 2,5 залежно від відхилення від норми.

Очікувана тривалість життя визначається послідовним перебором персональних імовірностей виживання з таблиці для кожного віку та визначенням найменшого віку, для якого імовірність виживання вперше стає нижчою за 60%. Такий підхід базується на повних таблицях смертності Держстату України (2020 р.), і забезпечує практично аналогічний результат до офіційно опублікованої середньої очікуваної тривалості життя. Хоча 60% не є стандартом по оцінці очікуваної тривалості життя, в контексті розробленої моделі це робить показник більш наочним і зрозумілим для широкого кола користувачів при мінімальних відхиленнях від офіційних статистичних стандартів. Алгоритм можна виразити наступним чином:

$$E_x = \min\{age : P(survival) < 60\%\} \quad (32)$$

де E_x — очікувана тривалість життя для особи віком x, $P(survival)$ — персональна ймовірність виживання на відповідний вік (береться з персональної таблиці виживання), 60% — порогове значення, визначене для розрахунку очікуваної тривалості життя відповідно до методики.

Таким чином, запропонована математична модель, побудована з використанням чітко визначених коефіцієнтів впливу, дозволяє отримати високоточні персоналізовані прогнози ймовірності виживання. Такий підхід не лише забезпечує деталізацію прогнозів на рівні окремої особи, але й дозволяє враховувати динамічні зміни у стані здоров’я та умовах життя населення. У подальшому використання цієї моделі може слугувати ефективним інструментом для прийняття рішень у сфері охорони здоров’я, соціального планування та індивідуальної медичної профілактики.

1.4.2 Розробка алгоритму адаптивного прогнозування з урахуванням специфіки українського населення та врахування особливих категорій населення у моделі

Важливою складовою інтегрованої математичної моделі прогнозування тривалості життя є алгоритм, що здатний враховувати особливості демографічного, соціального та медичного профілю українського населення. Це особливо актуально в умовах сучасних викликів, таких як війна [85], пандемія COVID-19 [86], широке поширення посттравматичних стресових розладів (ПТСР) [87], а також значна кількість внутрішньо переміщених осіб (ВПО) [88]. Ці фактори істотно змінюють ризики для здоров'я та тривалості життя українців і мають бути враховані для отримання релевантних відповідних прогнозів.

Розроблений алгоритм базується на принципі динамічної корекції коефіцієнтів впливу факторів у процесі накопичення нових статистичних даних. Він включає кілька етапів, що послідовно забезпечують адаптивність та індивідуалізацію прогнозу.

На першому етапі алгоритм виконує ініціалізацію базових значень параметрів моделі, що відповідають середньопопуляційним характеристикам населення України за даними статистичних обстежень. Цей крок формує початкову основу для подальших адаптивних змін.

Другий етап включає інтеграцію персоналізованих даних користувача в модель прогнозування. Користувачем заповнюється відповідна анкета, в якій відображаються медичні діагнози, соціально-демографічні умови, поведінкові особливості та інші індивідуальні характеристики. На цьому етапі здійснюється автоматична перевірка якості та повноти введених даних, що забезпечує достовірність отриманих прогнозів.

На третьому етапі здійснюється адаптивне уточнення коефіцієнтів впливу факторів на основі нових надходжень даних. Якщо, наприклад, накопичуються нові дані щодо впливу COVID-19 або ПТСР на смертність, то відповідні коефіцієнти в моделі автоматично перераховуються.

Четвертий етап полягає в реалізації безпосереднього прогнозування тривалості життя з використанням адаптованих параметрів. На цьому етапі модель застосовує скориговані коефіцієнти для обчислення персоналізованої ймовірності виживання з урахуванням особливостей конкретної особи. В результаті кожен користувач отримує індивідуалізований прогноз, що максимально точно відображає його актуальний стан здоров'я та соціально-демографічні умови.

Останній етап алгоритму передбачає регулярний зворотний зв'язок та переоцінку результатів. При отриманні нових даних, що свідчать про зміну ситуації у користувача (наприклад, зміна соціально-економічного статусу, лікування захворювання чи зміна способу життя), модель повторно застосовує алгоритм адаптації, оновлюючи прогноз. Таким чином, прогноз не є статичним, а постійно оновлюється, відображаючи реальні зміни в житті користувача. Особливо важливим етапом у створенні інтегрованої математичної моделі прогнозування тривалості життя є врахування специфічних категорій населення, які характеризуються підвищеним рівнем ризиків через особливі життєві обставини. В умовах сучасної України такими категоріями є особи, які мають посттравматичний стресовий розлад (ПТСР), внутрішньо переміщені особи (ВПО), а також особи, які перенесли COVID-19. Урахування цих груп населення є принципово важливим, адже саме ці чинники найбільшою мірою впливають на зниження очікуваної тривалості життя, підвищення ризиків хронічних захворювань і суттєво погіршують якість життя.

Посттравматичний стресовий розлад (ПТСР) має суттєвий вплив на стан здоров’я людини через довготривалий психологічний стрес і супутні фізіологічні зміни. ПТСР часто виникає у результаті пережитих військових дій, насильства чи інших травматичних подій, які сьогодні є особливо актуальними в умовах війни в Україні. Особи з ПТСР мають підвищений ризик розвитку депресії, тривожних розладів, серцево-судинних захворювань, а також схильні до шкідливих поведінкових звичок (зловживання алкоголем, куріння). У запропонованій моделі прогнозування коефіцієнт впливу ПТСР визначається окремо, на основі даних

спеціальних досліджень українських ветеранів та цивільних осіб, які пережили травматичні події. Це дозволяє точніше оцінити додатковий ризик для цієї категорії населення і створити ефективніші рекомендації з метою продовження життя.

Внутрішньо переміщені особи (ВПО) — ще одна важлива категорія, специфічна для сучасних умов України. Вимушене переміщення та зміна звичного середовища проживання часто супроводжуються погіршенням соціально-економічних умов, доступу до якісної медичної допомоги та зростанням стресового навантаження. Це в свою чергу призводить до підвищення ризиків розвитку хронічних захворювань, особливо серцево-судинних та психічних розладів. У рамках моделі було проведено аналіз специфічних ризиків, пов'язаних із внутрішнім переміщенням, і визначено окремий коефіцієнт впливу статусу ВПО, що враховує тривалість перебування у статусі ВПО, доступність соціальних і медичних ресурсів та психологічні наслідки переміщення.

COVID-19 став безпрецедентним медико-соціальним викликом для всього світу, включаючи Україну. Особи, які перенесли COVID-19, особливо середньої і тяжкої форми, стикаються з численними довготривалими наслідками для здоров'я. До таких наслідків належать серцево-судинні ускладнення, хронічна втома, порушення дихальної функції, неврологічні порушення, а також загальне погіршення якості життя. У запропонованій моделі особливий акцент зроблено на оцінці впливу COVID-19 на прогноз виживання з огляду на тяжкість перебігу хвороби та наявність або відсутність ускладнень (так званого «довгого COVID-19»). Завдяки аналізу сучасних українських і міжнародних даних було визначено додаткові коефіцієнти ризику для осіб, які перенесли COVID-19, що дозволяє точніше оцінити довготривалий прогноз тривалості життя.

Додатковим підтвердженням актуальності вибору критеріїв оцінки фізичної та психоемоційної активності як факторів, що впливають на тривалість життя, є результати дослідження, проведеного за участю офіцерів Сил оборони України в умовах бойових дій. За допомогою портативних електрокардіографічних пристроїв було виявлено чіткі закономірності зміни фізіологічних резервів залежно від

характеру та інтенсивності навантажень. Встановлено, що показники варіабельності серцевого ритму (ВСР) дозволяють ефективно оцінити фізіологічну вартість діяльності та психоемоційні стани військовослужбовців у динаміці. Важливо відзначити, що високий рівень фізичного або психоемоційного навантаження призводить до суттєвого зменшення функціональних резервів організму, що підвищує фізіологічну ціну активності та потенційно скорочує тривалість життя [89, 90].

Ці дані корелюють із результатами нашого дослідження та підтверджують обґрунтованість застосування подібних показників при побудові моделі прогнозування тривалості життя.

Завдяки такому інтегрованому підходу, модель здатна враховувати специфічні для українського середовища фактори, зокрема ті, що виникли у зв'язку з поточними соціально-економічними та медичними викликами, такими як статус внутрішньо переміщених осіб (ВПО), посттравматичний стресовий розлад (ПТСР), та перенесений COVID-19. Ці фактори мають окремі коефіцієнти впливу, які були визначені спеціально для українського населення з використанням локальних даних, що дозволило значно підвищити точність та релевантність прогнозів.

Таким чином, створений алгоритм є ключовим елементом інтегрованої математичної моделі прогнозування, що забезпечує точність, динамічність і персоналізованість прогнозів тривалості життя для населення України в сучасних умовах. Урахування особливих категорій населення, таких як особи з ПТСР, ВПО та ті, хто переніс COVID-19, дозволяє зробити інтегровану модель максимально адаптованою до реальних умов сучасної України. Завдяки цьому модель не лише забезпечує високу точність і релевантність прогнозів тривалості життя, але й надає важливу інформацію для розробки соціально-економічних стратегій підтримки та медичних програм реабілітації, що спрямовані на покращення здоров’я і добробуту цих особливо вразливих категорій населення.

1.5 Висновки до першого розділу

У результаті аналізу теоретичних та методичних підходів до прогнозування тривалості життя встановлено, що традиційні демографічні моделі, такі як Lee-Carter та Gompertz-Makeham, мають суттєві обмеження для персоналізованих прогнозів через неможливість врахування індивідуальних медичних, соціальних та поведінкових факторів. Тому для задач персоналізованого прогнозування рекомендовано використовувати сучасні інтегровані підходи, що поєднують демографічні, соціальні та медичні фактори з застосуванням алгоритмів машинного навчання.

Проведений порівняльний аналіз сучасних інтегрованих моделей прогнозування тривалості життя, таких як MPoRT, ONS та GBD, показав, що вони ефективно враховують різнорідні фактори та забезпечують високу точність прогнозування завдяки застосуванню складних статистичних та машинних методів обробки великих масивів даних. Водночас для адаптації таких моделей до специфіки українського населення важливим є уточнення вагомості окремих факторів на основі місцевих соціально-демографічних умов та медичних даних.

З огляду на проведений аналіз сучасних методів машинного навчання, для практичного застосування у даному дослідженні рекомендовано використовувати алгоритми Random Forest та XGBoost, які вирізняються високою точністю, стабільністю результатів та можливістю роботи з великими й неоднорідними наборами даних. Ці алгоритми забезпечують ефективне визначення впливу індивідуальних факторів ризику на тривалість життя та дозволяють реалізувати динамічне оновлення прогнозів.

Комплексний аналіз різних критеріїв оцінювання точності прогнозних моделей визначив, що для задач прогнозування тривалості життя найбільш релевантними є регресійні метрики, такі як MAE, RMSE, MSE та коефіцієнт детермінації (R^2). Ці метрики дозволяють об’єктивно оцінити точність прогнозу та забезпечити надійність результатів у практичних застосуваннях.

Особливу увагу необхідно приділити інтеграції цифрових медичних та персональних даних, які отримуються за допомогою сучасних носимих пристроїв та цифрових платформ (HealthKit, Google Fit). Це дозволяє суттєво підвищити точність та актуальність персоналізованих прогнозів тривалості життя шляхом включення до моделі широкого спектру динамічних показників, які раніше були недоступні для аналізу. Подальший розвиток таких інтегрованих технологій створює передумови для їх широкого впровадження в системі охорони здоров'я України.

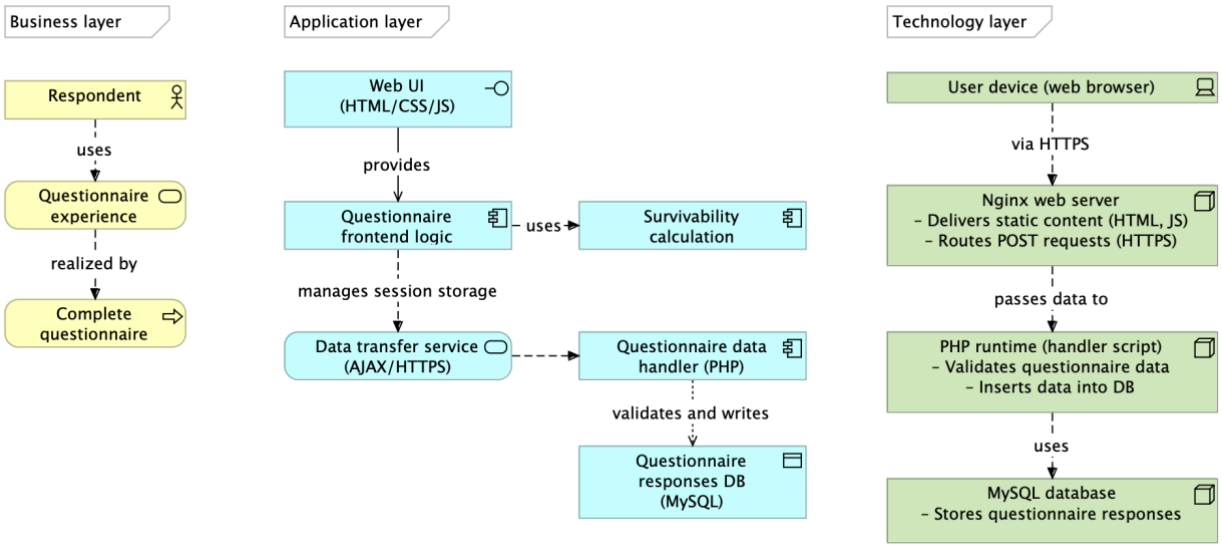
РОЗДІЛ 2 РОЗРОБКА ТА ПРАКТИЧНА РЕАЛІЗАЦІЯ ВЕБ-ЗАСТОСУНКУ ДЛЯ ПРОГНОЗУВАННЯ ТРИВАЛОСТІ ЖИТТЯ

2.1 Архітектура веб-застосунку «Калькулятор життя»

Веб-застосунок «Калькулятор життя» [91] реалізує інтегровану технологію персоналізованого прогнозування тривалості життя і має у своїй основі математичну модель, яка враховує широкий спектр демографічних, соціальних та медичних факторів, адаптованих до особливостей українського населення. Створення цієї технології було обумовлене потребою в ефективному і простому у використанні інструменті, здатному працювати в умовах високого навантаження і обмежених обчислювальних ресурсів, гарантуючи при цьому високий рівень конфіденційності і безпеки даних.

Архітектура веб-застосунку відповідає стандарту ISO 42010, що забезпечує чіткість взаємодії його компонентів [92]. Узагальнена модель архітектури була реалізована з використанням мови ArchiMate [93], що дозволило чітко розмежувати три основні рівні: бізнес-рівень, рівень застосунків та технологічний рівень (Рис. 3).

Рисунок 3. Основні рівні Archi-моделі архітектури Калькулятору Життя



2.1.1 Бізнес-рівень: модель взаємодії користувача

Бізнес-рівень архітектури веб-застосунку «Калькулятор життя» є найвищим рівнем у структурі, на якому безпосередньо реалізується взаємодія кінцевого користувача із системою. Основна мета цього рівня полягає у забезпеченні максимально зручного, ефективного та інтуїтивно зрозумілого процесу взаємодії, який дозволяє отримувати персоналізований прогноз тривалості життя з мінімальними зусиллями користувача та без додаткової підготовки [94, 95].

Ключовим учасником на бізнес-рівні є кінцевий користувач (респондент), яким може бути як пересічна людина без спеціалізованих знань у галузі медицини чи статистики, так і професійний медичний працівник, що використовує інструмент для консультативних цілей. Враховуючи різні рівні компетенції користувачів, була розроблена універсальна модель взаємодії, яка забезпечує доступність, простоту й інформативність для широкого кола аудиторії.

Основна взаємодія відбувається у вигляді покрокового анкетування, що дозволяє користувачеві поступово вводити необхідні дані.

З метою підвищення ефективності збору даних та зменшення помилок введення кожен пункт анкети супроводжується чіткими підказками та поясненнями, що дозволяє мінімізувати непорозуміння та забезпечує високий

рівень якості введеної інформації. Для зручності користувачів реалізовано можливість автоматичного заповнення частини анкетних даних завдяки інтеграції з зовнішніми цифровими платформами.

Після завершення заповнення анкети користувач отримує персоналізований результат прогнозу тривалості життя, який формується миттєво завдяки застосуванню клієнтського обчислення даних безпосередньо в браузері. Результат містить як базові числові показники (очікувану тривалість життя, ризики смертності у певному віці), так і розгорнуту пояснювальну інформацію про фактори, які найбільше впливають на отриманий прогноз. Додатково користувач отримує конкретні персоналізовані рекомендації щодо зміни поведінки або способу життя, які можуть допомогти йому покращити показники прогнозу в майбутньому.

Усі отримані дані можуть бути анонімно використані для подальшого статистичного аналізу та вдосконалення прогнозних моделей, що чітко повідомляється користувачу, і він має можливість вибору стосовно згоди на таке використання. Користувачі веб-застосунку були проінформовані про обробку персональних даних для дослідницьких цілей через посилання на Політику конфіденційності ДУ «Центр громадського здоров’я МОЗ України», розміщене безпосередньо в інтерфейсі опитувальника. Таким чином забезпечується прозорість щодо використання персональних даних, а також дотримуються сучасні вимоги захисту даних, такі як GDPR (General Data Protection Regulation) .

Особливу увагу було приділено дизайну інтерфейсу веб-застосунку, який реалізовано за принципами UX/UI-дизайну [98]. Це дозволяє користувачеві з мінімальними зусиллями і максимально комфортно проходити через усі етапи взаємодії з застосунком, що позитивно впливає на користувацький досвід (user experience) та підвищує ймовірність повторного використання сервісу. Слід зазначити, що прогнозована тривалість життя, яку отримують користувачі, є виключно оціночною інформацією, не є медичним діагнозом та не замінює консультації лікаря. Для ухвалення важливих медичних рішень або у разі будь-яких

питань щодо власного здоров'я користувачам рекомендується звертатися до професійних медичних фахівців.

Таким чином, бізнес-рівень веб-застосунку «Калькулятор життя» забезпечує ефективну модель взаємодії користувача, яка відповідає сучасним вимогам зручності, швидкості, точності і конфіденційності. Завдяки цьому вдається не лише ефективно збирати необхідні для прогнозування дані, але й забезпечувати користувачам зручний і зрозумілий інструмент, який відповідає їхнім індивідуальним потребам і водночас сприяє свідомому ставленню до власного здоров'я та способу життя.

2.1.2 Рівень застосунків: фронтенд та бекенд-логіка

Рівень застосунків у веб-застосунку «Калькулятор життя» реалізує безпосередню обробку та відображення даних, забезпечує ефективну взаємодію між клієнтською частиною (фронтендтом) та серверною частиною (бекендом). Вибір технологій та підходів для реалізації цього рівня було здійснено з урахуванням необхідності забезпечення високої продуктивності, масштабованості, безпеки та зручності використання на різних платформах і пристроях.

Фронтенд веб-застосунку реалізовано з використанням чистого (нативного) JavaScript без залучення зовнішніх фреймворків чи бібліотек. Вибір нативного JavaScript був обумовлений необхідністю максимальної швидкодії застосунку, високої оптимізації продуктивності, а також зменшення часу завантаження сторінок завдяки скороченню загального обсягу коду. Окрім швидкодії, відсутність залежностей від бібліотек значно спрощує підтримку та довгостроковий розвиток проєкту, забезпечуючи більшу контрольованість коду та незалежність від оновлень зовнішніх компонентів.

Функціональність фронтенду передбачає покрокову навігацію користувача по анкетах, яка супроводжується миттєвими клієнтськими обчисленнями і негайним представленням результатів. Для цього використовується ретельно

оптимізований JavaScript-код, який забезпечує високу швидкість реакції інтерфейсу без необхідності частих запитів до серверної частини. Це значно покращує користувацький досвід, особливо при роботі зі складними математичними обчисленнями, які реалізуються безпосередньо на стороні браузера.

Користувацький інтерфейс веб-застосунку створений з використанням повністю кастомного CSS-коду без застосування фреймворків на кшталт Bootstrap. Такий підхід дозволив розробити максимально легкий, гнучкий та адаптивний дизайн, що ефективно працює на будь-яких типах пристроїв — від стаціонарних комп'ютерів до мобільних пристроїв. Окрім того, кастомна CSS-реалізація дозволяє уникнути зайвого завантаження непотрібних стилів, суттєво скорочуючи час завантаження веб-сторінок і покращуючи загальну швидкодію інтерфейсу.

Відмова від використання зовнішніх бібліотек для побудови графіків та діаграм на користь нативного JavaScript дозволила ще більше зменшити кількість коду, що завантажується у браузер користувача. Візуалізація результатів прогнозу тривалості життя та ймовірності виживання реалізована з використанням вбудованих можливостей браузерних API (наприклад, SVG і Canvas), що забезпечує швидке, точне і наочне відображення результатів без потреби у додаткових бібліотеках.

Серверна частина веб-застосунку написана повністю на чистому PHP, без застосування будь-яких додаткових фреймворків чи сторонніх бібліотек. Такий підхід забезпечує максимально легку і швидку роботу бекенду, зменшує час обробки запитів і суттєво спрощує подальшу підтримку і масштабування.

Бекенд веб-застосунку функціонує як ізольований компонент, не інтегрований безпосередньо з жодними зовнішніми сервісами. Це рішення було прийнято з метою мінімізації ризиків для конфіденційності та безпеки користувацьких даних. Основні завдання серверної частини полягають у:

- Прийомі та валідації даних, які надходять із фронтенду після завершення процесу анкетування. Реалізована комплексна система перевірки валідності введених даних на сервері.
- Псевдонімізації та анонімізації даних, що гарантує високий рівень конфіденційності відповідно до стандартів GDPR.
- Безпечному зберіганні даних у базі даних MySQL, з використанням підготовлених (prepared) SQL-запитів, що забезпечує додатковий захист від атак типу SQL Injection.

Оскільки бекенд не взаємодіє із зовнішніми сервісами напряму, будь-яка комунікація між сервером і фронтендом реалізована через стандартизований RESTful API з обов'язковим використанням HTTPS-протоколу, що додатково захищає передачу інформації та унеможливлює несанкціонований доступ до чутливих даних.

PHP-код бекенду структурований за принципом максимальної простоти, з чітким розмежуванням логіки обробки, зберігання та надання даних. Завдяки використанню чистого PHP-коду без фреймворків, серверна частина є надзвичайно легкою для підтримки, швидкою у роботі та практично повністю позбавленою ризиків, пов’язаних із вразливостями зовнішніх бібліотек.

Таким чином, рівень застосунків веб-застосунку «Калькулятор життя» поєднує у собі швидкий, ефективний і мінімалістичний фронтенд, повністю побудований на нативних технологіях JavaScript та CSS, із серверною частиною, яка реалізована на чистому PHP-коді без жодних додаткових фреймворків. Вибраний підхід забезпечує максимальну продуктивність, зручність використання, масштабованість і надійний захист даних, що дозволяє досягати високих показників ефективності роботи веб-застосунку в умовах реального використання.

2.1.3 Технологічний рівень: серверна інфраструктура, захист даних та інформаційна безпека

Технологічний рівень веб-застосунку «Калькулятор життя» відповідає за забезпечення надійної, продуктивної та безпечної роботи програмного комплексу, який обробляє чутливі персональні, медичні й соціально-демографічні дані користувачів. Створення такого застосунку потребує ретельно підібраної серверної інфраструктури та сучасних механізмів захисту інформації, що відповідають міжнародним стандартам, зокрема GDPR [99] та ISO/IEC 27001 [100].

Серверна інфраструктура веб-застосунку «Калькулятор життя» створена відповідно до сучасних вимог безпеки, надійності та конфіденційності для обробки чутливих медичних та соціальних даних користувачів в Україні. Інфраструктура побудована на базі внутрішніх ресурсів ЦГЗ МОЗ України, що гарантує повний контроль над зберіганням та обробкою персональних даних користувачів.

Ключовими компонентами інфраструктури є:

- Веб-сервер Nginx, який приймає і розподіляє вхідні HTTP-запити, обробляє статичний контент (HTML, CSS, JavaScript, зображення), забезпечує маршрутизацію запитів до серверів застосунків та термінує HTTPS-з’єднання для додаткового захисту.
- Сервер застосунків PHP-FPM забезпечує виконання серверної логіки веб-застосунку, яка написана виключно на чистому PHP без використання будь-яких сторонніх фреймворків. Цей підхід забезпечує максимальну продуктивність, легкість підтримки та мінімальні ризики вразливостей. PHP-FPM працює в режимі ефективного пулу процесів, що дозволяє швидко обробляти численні запити одночасно.
- Реляційна база даних MySQL використовується для зберігання структурованих персональних і статистичних даних. Сервери баз даних функціонують у режимі реплікації master-slave, забезпечуючи надійність, швидке відновлення даних після збоїв та високу продуктивність обробки запитів.
- Внутрішнє безпечне сховище даних ЦГЗ МОЗ України застосовується для зберігання резервних копій (бек-апів), логів та архівних записів. Це сховище гарантує відповідність усім державним нормам безпеки і конфіденційності

персональних даних, забезпечуючи найвищий рівень захисту та контроль доступу до інформації.

Система використовує балансування навантаження на базі спеціалізованого рішення (Load Balancer), розгорнутого в межах інфраструктури ЦГЗ МОЗ України. Це рішення розподіляє запити між кількома серверами застосунків, забезпечуючи рівномірний розподіл навантаження, підвищуючи відмовостійкість і продуктивність системи.

Передбачено можливість горизонтального масштабування серверних ресурсів — при збільшенні навантаження на веб-застосунок адміністратори ЦГЗ можуть оперативно розгортати додаткові серверні вузли, що забезпечує стабільність роботи застосунку навіть за умов пікових навантажень.

Система забезпечує постійний моніторинг серверної інфраструктури з використанням внутрішньої системи моніторингу та оповіщення, яка контролює роботу всіх критичних компонентів (веб-серверів, серверів застосунків, баз даних). У разі виникнення збоїв чи критичних подій, відповідальні адміністратори негайно отримують автоматичні повідомлення.

Усі події на сервері, включаючи HTTP-запити, помилки та інші операції, логуються і зберігаються у захищеному лог-сховищі ЦГЗ МОЗ України. Це дозволяє проводити оперативний аудит діяльності застосунку та забезпечує необхідну прозорість у разі проведення перевірок.

Зберігання резервних копій реалізоване у внутрішньому безпечному сховищі ЦГЗ МОЗ України. Резервне копіювання відбувається регулярно (щодня, щотижня, щомісячно) згідно з відповідними політиками ЦГЗ. Це гарантує можливість швидкого відновлення роботи системи та збереження критичних даних навіть за умов серйозних технічних збоїв або інших форс-мажорних ситуацій.

План аварійного відновлення чітко визначає порядок дій у випадку непередбачуваних ситуацій, таких як вихід із ладу серверів чи інших критичних компонентів. Це дозволяє мінімізувати час простою системи та швидко відновити її нормальну роботу.

Таким чином, серверна інфраструктура веб-застосунку «Калькулятор життя», побудована на внутрішніх ресурсах Центру громадського здоров'я МОЗ України, повністю відповідає сучасним державним і міжнародним стандартам надійності, безпеки та конфіденційності. Використання внутрішнього безпечного сховища та повна інтеграція інфраструктури в захищене середовище ЦГЗ забезпечує високу ефективність, надійний захист інформації, швидкість роботи системи та гарантує довіру користувачів до веб-застосунку.

Серверна інфраструктура веб-застосунку побудована з використанням сучасних рішень, які дозволяють забезпечити високу доступність, масштабованість та стійкість до навантажень.

Ураховуючи специфіку веб-застосунку «Калькулятор життя», що оперує конфіденційними медичними та соціально-демографічними даними користувачів, питання інформаційної безпеки є одним із ключових аспектів його розробки та експлуатації. Застосунок побудований у відповідності до національних і міжнародних стандартів безпеки, зокрема вимог GDPR, ISO/IEC 27001, а також внутрішніх регламентів ЦГЗ МОЗ України.

Уся взаємодія між користувачами та сервером здійснюється виключно за захищеним протоколом HTTPS із використанням сертифікатів SSL/TLS, які відповідають найвищим сучасним стандартам шифрування (наприклад, TLS 1.2 та TLS 1.3). Це повністю виключає можливість перехоплення чи несанкціонованого доступу до даних під час їх передачі через інтернет. Сертифікати регулярно оновлюються та підлягають суворій перевірці й контролю.

Серверна частина веб-застосунку працює у повністю ізольованому середовищі ЦГЗ МОЗ України, що не має зовнішніх інтеграцій або прямих комунікацій з іншими зовнішніми ресурсами чи сервісами. Це дозволяє максимально мінімізувати ризик витоку чи компрометації даних. Усі сервери розташовані у захищеному дата-центрі, доступ до якого суворо обмежений і контрольований.

Регулярне резервне копіювання даних (backup) є невід’ємною частиною політики безпеки. Резервні копії створюються автоматично за встановленим розкладом та зберігаються в окремому захищеному сховищі ЦГЗ МОЗ України, з обмеженим фізичним та логічним доступом. Існує чітко прописаний план аварійного відновлення (disaster recovery), що дозволяє оперативно відновити повноцінну роботу системи у разі серйозних збоїв чи втрати даних.

Система оснащена комплексним механізмом моніторингу, який забезпечує реєстрацію всіх важливих подій: входу користувачів, спроб доступу до захищених ресурсів, помилок, а також будь-яких нетипових або підозрілих дій. Внутрішня система моніторингу, що базується на рішеннях ЦГЗ МОЗ України, забезпечує оперативне оповіщення відповідальних осіб у випадку виявлення будь-яких аномалій або підозрілих інцидентів.

Регулярно здійснюється аудит журналів подій (логів), що дозволяє виявляти потенційні вразливості або спроби несанкціонованого доступу на ранній стадії та оперативно вживати відповідних заходів.

Система захисту застосунку включає реалізацію на рівні серверної інфраструктури таких механізмів, як захист від SQL-ін'єкцій (шляхом використання підготовлених SQL-запитів), Cross-Site Scripting (XSS), Cross-Site Request Forgery (CSRF) та інших поширених веб-атак. Веб-сервер Nginx додатково оснащений спеціальними конфігураціями для фільтрації та блокування підозрілих запитів, що значно знижує ризик успішних атак.

Усі співробітники, які мають доступ до серверної інфраструктури або персональних даних, проходять регулярні тренінги з інформаційної безпеки та мають чітко визначені й документовані рівні доступу. Доступ до особливо чутливих даних надається лише обмеженій кількості авторизованих працівників відповідно до суворої політики управління доступом ЦГЗ МОЗ України.

Таким чином, система захисту даних та інформаційної безпеки веб-застосунку «Калькулятор життя» є комплексною та багаторівневою. Вона враховує всі сучасні вимоги безпеки, повністю відповідає державним і міжнародним

стандартам, гарантує надійний захист персональних і медичних даних користувачів, а також забезпечує постійний контроль та оперативне реагування на можливі загрози.

2.1.4 Компресія та обробка зображень інтерфейсу за допомогою ШІ

Однією з ключових технічних проблем було завантаження важких медіафайлів, що безпосередньо впливало на швидкість завантаження вебсторінок, ефективність взаємодії з інтерфейсом та загальний досвід користувача. Особливо гостро ця проблема стояла перед розробленим інструментом прогнозування тривалості життя, оскільки опитувальник може демонструвати до 21 високоякісного зображення протягом одного проходження. Загальний розмір цих медіафайлів початково становив близько 16 мегабайт, що значно ускладнювало швидке завантаження контенту, негативно впливало на зручність користування інструментом і могло призводити до зростання частки незавершених проходжень.

У зв’язку з цим була поставлена амбітна задача максимальної оптимізації цих файлів, що вимагало розробки спеціалізованих технічних рішень. Для вирішення цього складного технічного виклику були застосовані технології, на які подані патетні заявки в США, зокрема «Adaptive Content Delivery Systems and Methods» та «Systems for Reducing Media Data Size while Maintaining Perceptual Quality». Ці технології базуються на використанні спеціалізованих психовізуальних метрик та індексу психологічного стиснення (Psychovisual Compression Index — PVCI), які дозволяють аналізувати медіаконтент з урахуванням особливостей людського сприйняття. Завдяки застосуванню тренованих нейронних мереж вдалося ідентифікувати психологічно значущі особливості зображень, які необхідно зберегти, та визначити ділянки, що можуть бути оптимізовані без втрати візуальної якості.

Завдяки застосуванню розроблених технологій вдалося досягти значного ступеня оптимізації, зменшивши розмір медіаданих на 96%, при цьому повністю

зберігаючи сприйняття візуальної якості. Оптимізовані зображення були перетворені у сучасні формати, такі як WebP та AVIF, та підготовлені в кількох роздільних здатностях, що забезпечило повну адаптивність контенту під різні пристрої користувачів, від мобільних телефонів до високороздільних моніторів.

Зниження часу завантаження медіаконтенту, отримане завдяки оптимізації, відіграє надзвичайно важливу роль у забезпеченні комфортного проходження опитувальника, особливо в умовах повільного або нестабільного інтернет-з'єднання. Крім того, реалізована компресія суттєво знизила навантаження на серверну інфраструктуру Центру громадського здоров'я, зменшивши вимоги до ресурсів передачі даних медіаконтенту, що дозволило оптимізувати витрати й покращити стабільність роботи платформи. Таким чином, оптимізація медіаконтенту суттєво підвищила зручність і швидкість взаємодії з інструментом прогнозування, значно покращивши загальний користувацький досвід і підвищивши ефективність всього рішення.

2.2 Інтерфейсні рішення для ефективного збирання та аналізу даних

Для успішної роботи технології прогнозування тривалості життя важливе значення має зручність і зрозумілість користувацького інтерфейсу веб-застосунку. Інтерфейс «Калькулятора життя» був розроблений з урахуванням сучасних стандартів UX/UI-дизайну, що дозволяє легко й ефективно збирати інформацію та наочно презентувати результати аналізу.

Одним з основних принципів реалізації інтерфейсу є його адаптивність і орієнтація на користувача. Дизайн побудований з використанням мінімалістичного підходу, без залучення сторонніх бібліотек чи фреймворків, що дозволяє максимально швидко завантажувати сторінки та ефективно працювати на будь-яких типах пристроїв (від мобільних телефонів до настільних комп'ютерів). Простота оформлення, чистий візуальний стиль та інтуїтивно зрозуміле розташування

елементів сприяють комфортному та безперешкодному взаємодіянню користувача з системою.

Веб-інтерфейс опитувальника складається з 6 розділів:

1. Вітальна сторінка з інформацією про середню тривалість проходження та коротким описом анкети, яка запрошує розпочати опитування (Рис. 4).

2. Інформаційний розділ, що містить політику конфіденційності і збирає базові дані: вік і стать. Наступний слайд включає мотиваційне питання щодо досягнення обраної або самостійно сформульованої мрії у визначений рік. Ця інформація не використовується у розрахунках, не зберігається і служить лише мотивацією для повного проходження анкети (Рис. 5).

3. Розділ звичок: збирає інформацію щодо куріння, вживання алкоголю, харчування та фізичної активності (Рис. 6).

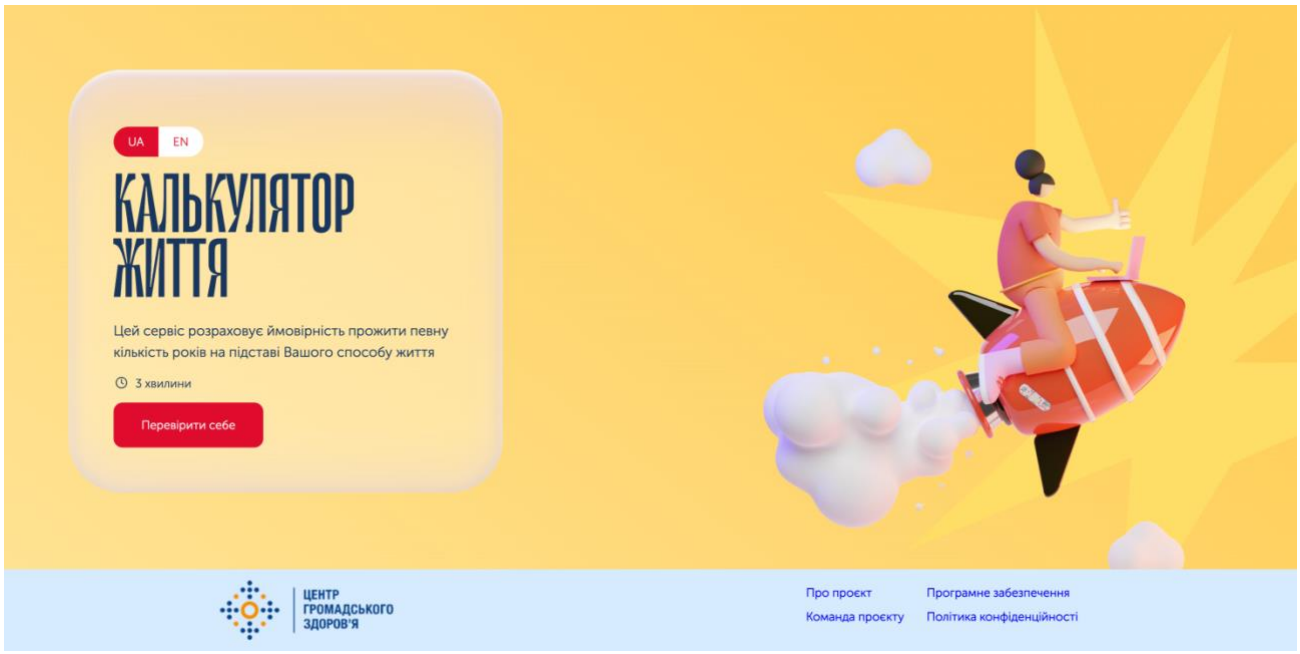
4. Розділ медичних даних, що включає зріст, вагу, перелік захворювань та медичних станів, інформацію про частоту звернень до сімейного лікаря, викликів швидкої допомоги та госпіталізацій на рік (Рис. 6).

5. Сторінка додаткової інформації про місце проживання, розмір населеного пункту, освіту, місячний дохід та статус внутрішньо переміщеної особи. Регіон і розмір населеного пункту наразі є статистичними і не впливають на персональні результати (Рис. 7).

6. Сторінка результатів: містить інформацію про очікувану тривалість життя, ймовірність досягнення визначеного віку у певному році та інтерактивний слайдер, який демонструє ймовірність виживання від поточного віку до 110 років (Рис. 8).

На головній сторінці користувача зустрічає лаконічний дизайн із чітко сформульованою інформацією про мету та завдання застосунку (Рис. 5). Користувачу одразу пропонується швидкий перехід до проходження анкети, яка займає орієнтовно 3 хвилини. Цей підхід мінімізує бар'єр входу та сприяє підвищенню конверсії серед користувачів.

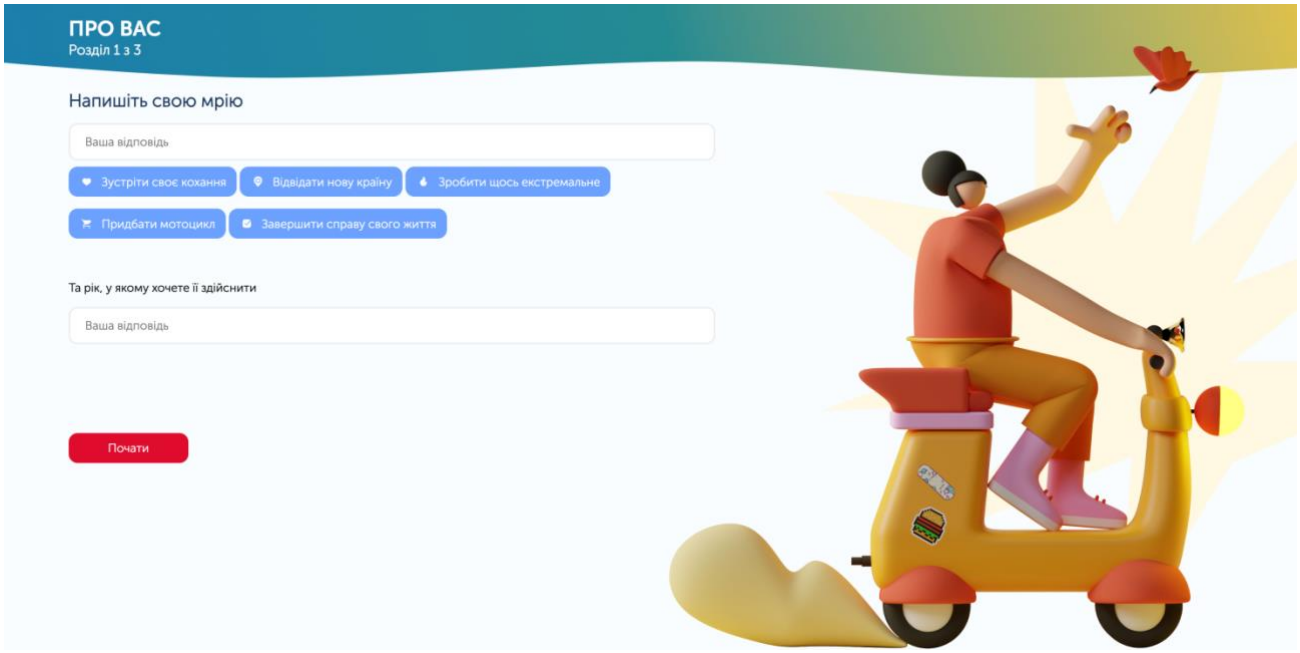
Рисунок 4. Загальний вигляд головної сторінки веб-інтерфейсу Калькулятора для настільних пристроїв



Покрокова анкета побудована з максимальною орієнтацією на простоту та швидкість взаємодії. Користувач поступово проходить три основні розділи, кожен з яких зосереджується на певній групі питань: персональні бажання («Про вас»), шкідливі та корисні звички («Ваші звички») та інформація про здоров'я («Ваше здоров'я»). Питання сформульовані максимально просто і зрозуміло, що дозволяє користувачам легко орієнтуватися в процесі заповнення анкети, незалежно від рівня їх цифрових навичок.

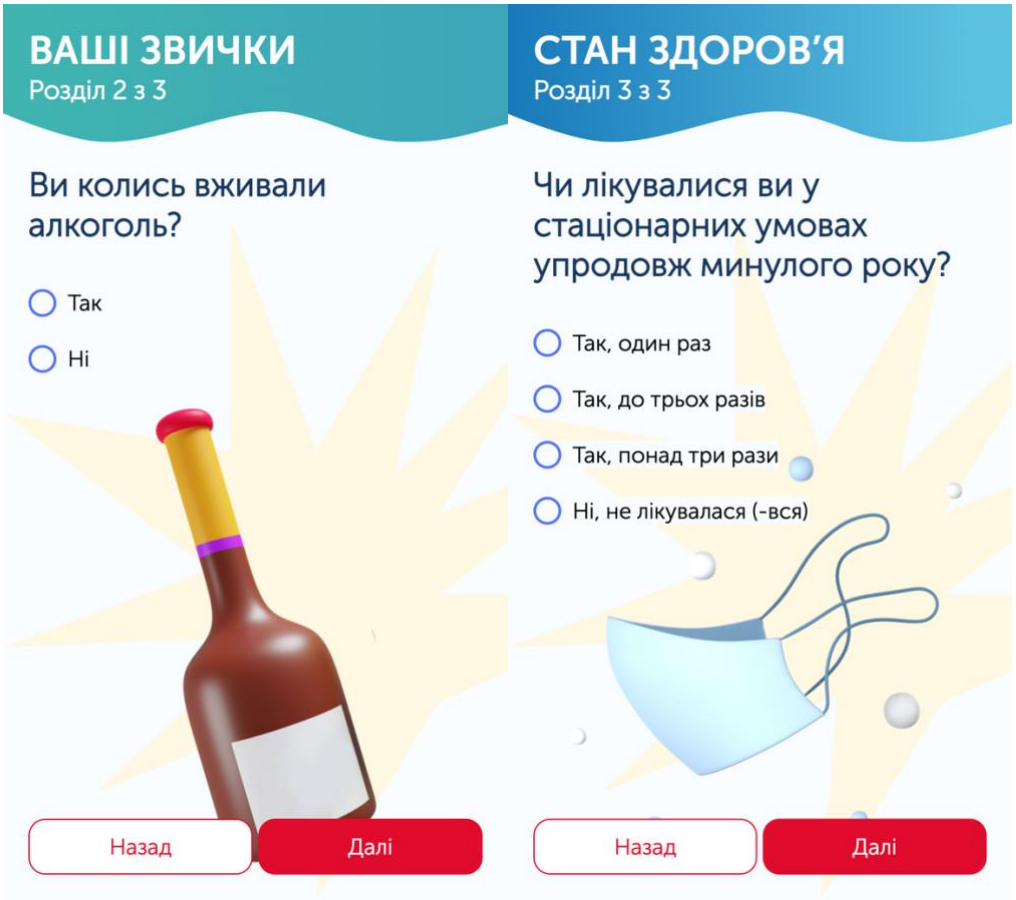
Інноваційним рішенням стало інтегрування в анкету питань, які стимулюють користувача замислитися над власними життєвими цілями та мріями (Рис. 5). Це підвищує мотивацію користувача пройти опитування до кінця, формуючи додатковий особистий інтерес до кінцевого результату.

Рисунок 5. Загальний вигляд інформаційного розділу «Про Вас»



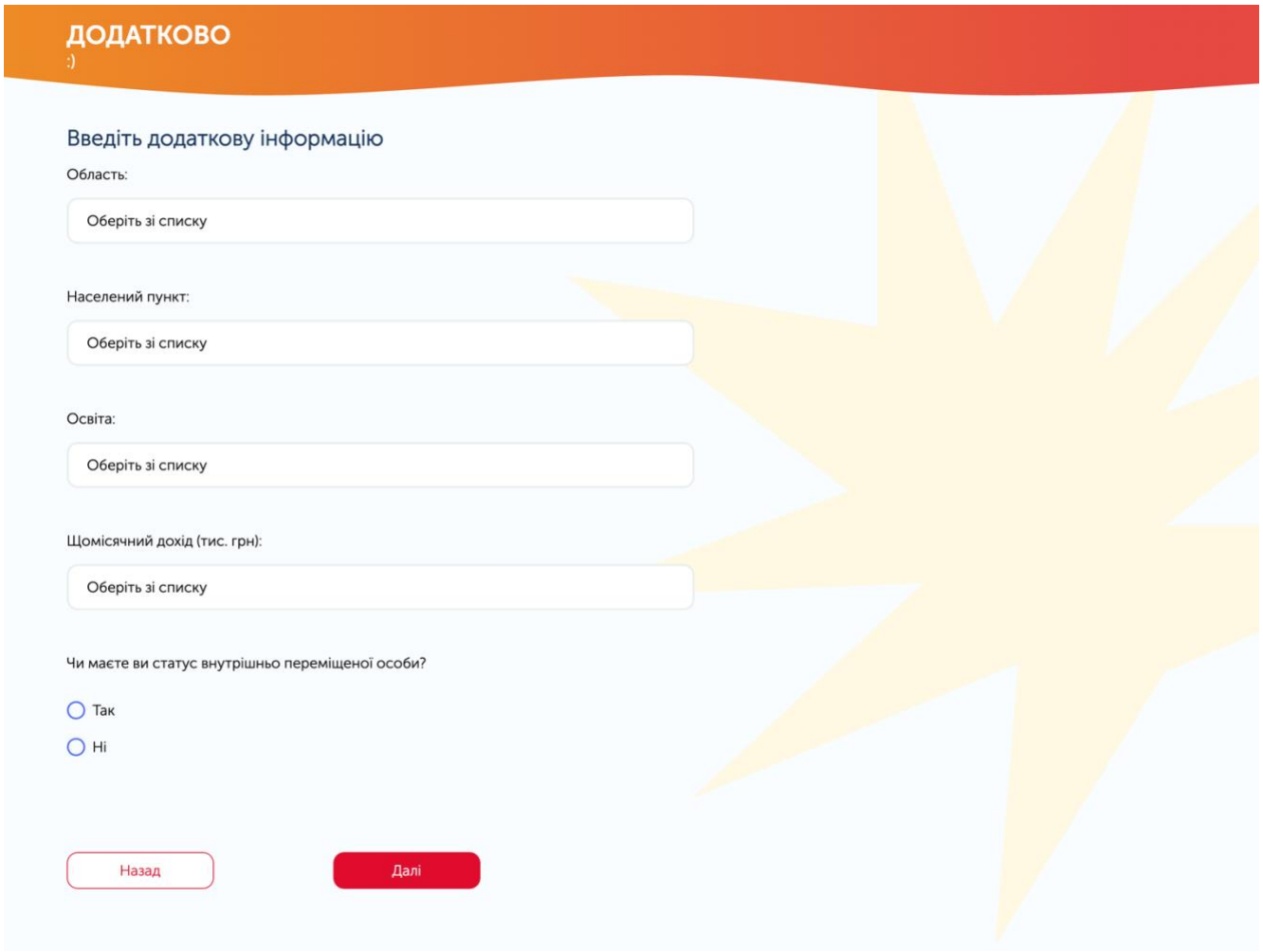
Для підвищення залученості користувачів і якості отриманих даних застосовуються інтерактивні елементи: вибір варіантів відповідей реалізовано у формі зручних перемикачів («радіокнопок») та кнопок-швидких відповідей, що значно спрощує процес анкетування (Рис. 6). Крім цього, для кожного запитання доступні чіткі й зрозумілі варіанти відповідей, які супроводжуються наочними ілюстраціями та анімаціями, що дозволяє зробити взаємодію більш динамічною та цікавою для користувачів.

Рисунок 6. Загальний вигляд розділів звичок та стану здоров’я адаптованого під мобільні пристрої



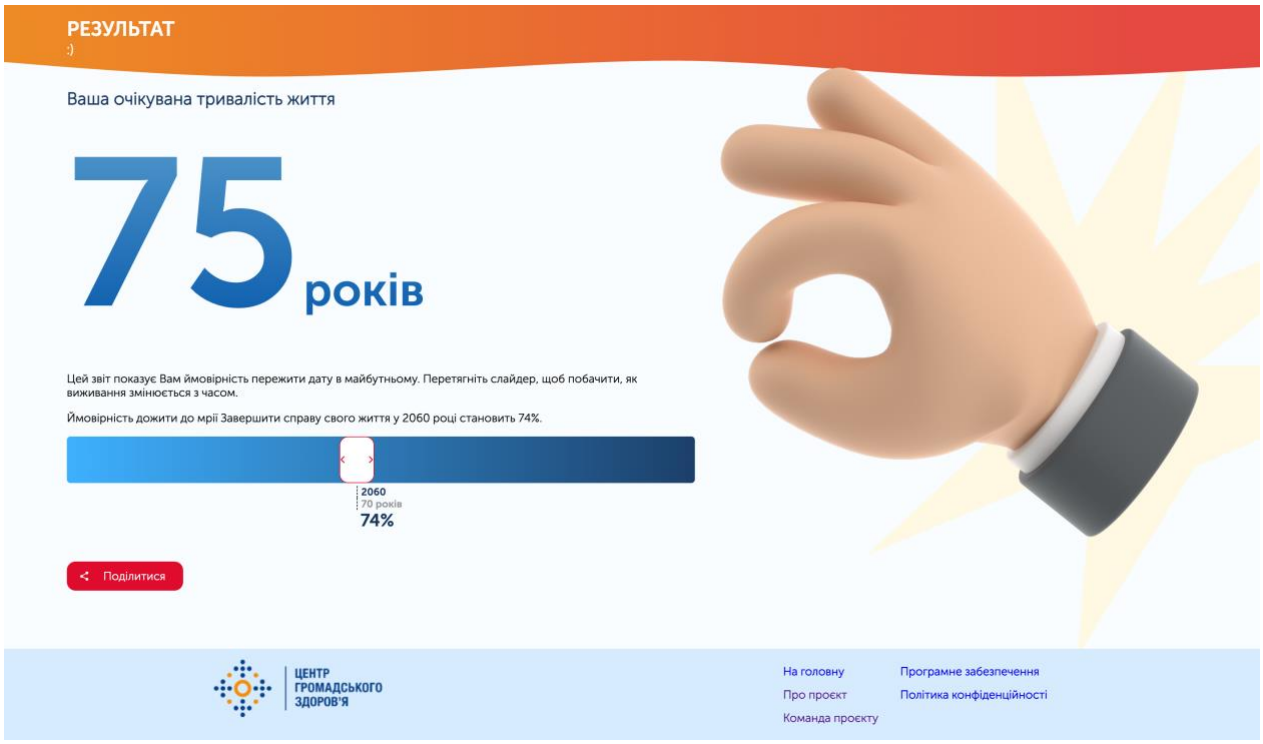
Всі елементи інтерфейсу, від розміру шрифтів та кнопок до колірної палітри й ілюстрацій, підібрані таким чином, щоб створювати позитивне емоційне враження, уникати перевантаження інформацією та забезпечувати максимально чітке сприйняття даних (Рис. 7).

Рисунок 7. Загальний вигляд розділу додаткової інформації адаптованого під планшети в альбомній орієнтації



Окремо варто зазначити ефективне рішення візуалізації результатів аналізу. Після завершення анкети користувач отримує чіткий та зрозумілий прогноз у вигляді конкретного числа років очікуваної тривалості життя, який супроводжується інтерактивним слайдером (Рис. 8), що демонструє динаміку зміни ймовірності виживання з часом. Такий формат дозволяє користувачеві легко зрозуміти й оцінити власні перспективи, взаємодіючи з результатами напряму на сторінці.

Рисунок 8. Веб-інтерфейс сторінки результатів «Калькулятора життя» з інтерактивним слайдером та прикладом проходження



Таким чином, розроблені інтерфейсні рішення веб-застосунку «Калькулятор життя» забезпечують не тільки ефективне збирання та якісний аналіз даних, але й високий рівень задоволеності користувачів. Це суттєво підвищує цінність отриманої інформації, яка використовується для подальшого покращення моделей прогнозування та загального підвищення ефективності роботи застосунку.

2.3 Інтеграція зовнішніх цифрових медичних і соціальних даних

У процесі розробки платформи була реалізована інтеграція з популярними сервісами Google Fit (для Android-пристроїв) та Apple HealthKit (для iOS-пристроїв). Для забезпечення інтеграції з Apple HealthKit було розроблено окремий мобільний додаток на платформі iOS, який дозволяє отримувати авторизований доступ до таких персональних медичних даних. Для Android-пристроїв аналогічні дані надходять через API Google Fit. Обидва мобільні додатки завантажують інтерфейс опитувальника в вбудованому браузері ОС. Інтеграція здійснюється саме у контексті нативних застосунків, оскільки безпосереднє отримання даних з

браузера неможливе через обмеження безпеки та конфіденційності платформ Apple та Google.

Інтеграція з HealthKit і Google Fit дозволяє автоматично заповнювати анкети користувачів шляхом використання цифрових медичних та фітнес-даних, зібраних персональними мобільними пристроями. Технічно процес інтеграції побудований через вбудовані API та дозвільні механізми операційних систем Apple та Android, які дозволяють безпечно отримувати доступ до інформації за згодою користувача. Після авторизації відповідні застосунки отримують доступ до останніх і агрегованих даних здоров’я, які зберігаються безпосередньо на пристрої користувача.

Для платформи iOS інтеграція здійснюється шляхом налаштування entitlements і provisioning-профілів для застосунку, після чого через HealthKit API отримуються такі показники, як маса тіла, зріст, фізична активність (кількість кроків, дистанція, активний час), частота серцевих скорочень, дані про сон, дієту (споживання енергії та алкоголю) та куріння.

Для платформи Android реалізовано аналогічний підхід через Google Fit API. В застосунку можна отримати аналогічні показники: вага, зріст, фізична активність, серцевий ритм, тривалість сну, дані про харчування, статус куріння.

Після отримання даних через нативні API, ці показники передаються в веб-застосунок через вбудований браузер із застосуванням клієнтського JavaScript, що дозволяє прив'язати отримані значення до відповідних полів анкети. Дані на цьому етапі обробляються винятково локально на пристрої користувача, мінімізуючи ризики, пов'язані з передачею інформації мережею.

Інтеграція веб-застосунку з платформами HealthKit та Google Fit має суттєві переваги, зокрема:

- Значне прискорення заповнення анкет. Згідно з результатами тестування, використання автоматичного заповнення дозволяє скоротити час заповнення анкети приблизно на 60%, оскільки користувачам не потрібно вручну вводити більшість даних.

- Підвищення достовірності та якості даних. Інтеграція цифрових даних, отриманих з пристроїв користувачів, дозволяє знизити кількість помилок, пов'язаних з ручним введенням, приблизно на 40%, що забезпечує більшу надійність прогнозів.
- Полегшення використання і зручність. Користувачам не потрібно шукати чи пам'ятати багато деталей про власний стан здоров'я, що позитивно впливає на користувацький досвід і збільшує ймовірність завершення анкети.

Докладне відображення факторів «Калькулятором життя», які відповідають даним з Apple HealthKit та Google Fit або можуть бути пораховані за їх допомогою, наведено нижче (Табл. 6). Для зручності сприйняття, префікси атрибутів вилучено: «HK» для HealthKit, «TYPE_» для Google Fit.

Таблиця 6. Зведення факторів «Калькулятора життя» з даними HealthKit та Google Fit

Параметр впливу	Джерело	Скорочені атрибути
Коефіцієнт ІМТ	HealthKit	QuantityTypeIdentifierBodyMass, QuantityTypeIdentifierHeight
	Google Fit	WEIGHT, HEIGHT
Фізична активність низької, середньої та високої інтенсивності	HealthKit	QuantityTypeIdentifierStepCount, QuantityTypeIdentifierDistanceWalkingRunning, QuantityTypeIdentifierAppleExerciseTime
	Google Fit	STEP_COUNT_DELTA, DISTANCE_DELTA, ACTIVITY_SEGMENT
Споживання здорової їжі	HealthKit	QuantityTypeIdentifierDietaryEnergyConsumed
	Google Fit	NUTRITION
Куріння	HealthKit	CategoryTypeIdentifierSmokingStatus

	Google Fit	SMOKING_STATUS
Вживання алкоголю	HealthKit	QuantityTypeIdentifierDietaryAlcohol
	Google Fit	NUTRITION

Архітектура інтеграції передбачає модульну структуру, завдяки якій додавання нових метрик або джерел даних не вимагає серйозних змін у користувацькому інтерфейсі. Налаштування нових метрик чи типів інформації здійснюється виключно через конфігураційні файли, що забезпечує гнучкість, масштабованість і легкість підтримки веб-застосунку.

Таким чином, реалізована інтеграція HealthKit (iOS) і Google Fit (Android) у веб-застосунку «Калькулятор життя» дозволяє суттєво покращити процес анкетування користувачів, підвищити якість та точність даних, а також забезпечити високий рівень задоволеності користувачів завдяки простоті, швидкості й безпеці взаємодії. Це створює надійний фундамент для подальшого розширення функціональних можливостей застосунку та забезпечує успішну реалізацію сучасних цифрових медичних рішень.

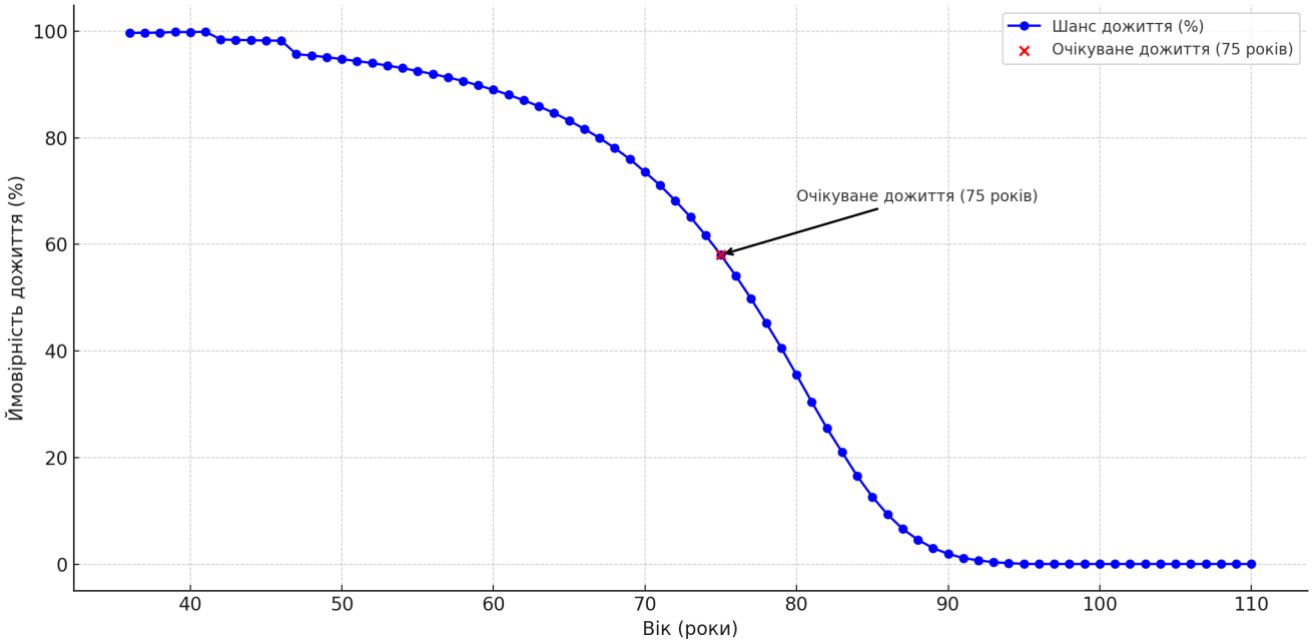
2.4 Приклад застосування розробленого інструменту

Для ілюстрації роботи розробленого інструменту прогнозування тривалості життя наведемо конкретний приклад проходження користувачем онлайн-опитувальника. Респондентом був чоловік віком 35 років, який ніколи не курив, має ступінь PhD, середні доходи, статус внутрішньо переміщеної особи, мріє завершити справу свого життя до 2060 року. Маса тіла становила 90 кг при зрості 190 см. Раціон характеризується щоденним вживанням фруктів, овочів та соків (5 порцій кожного). COVID-19 у важкій або середній формі користувач не переносив, госпіталізацій та викликів швидкої допомоги не було, алкоголь не вживає (лише одноразовий досвід), консультацію лікаря мав лише один раз на рік. Легка фізична

активність становить приблизно 20 хвилин щоденно, помірна та інтенсивна фізична активність відсутні.

Після введення наведених вище даних у розроблений інструмент було здійснено автоматичний розрахунок персоналізованих коефіцієнтів впливу, що ґрунтуються на зазначених факторах ризику та способі життя користувача. У результаті аналізу персоналізованої таблиці ймовірностей виживання було визначено, що прогнозована очікувана тривалість життя цього користувача становить 75 років (Рис. 9). Додатково було обчислено, що шанси цього респондента дожити до 2060 року, коли він планує завершити справу свого життя, складають 74% (Рис. 8).

Рисунок 9. Персональна таблиця дожиття для прикладу проходження



Цей приклад демонструє ефективність і практичну значущість розробленого інструменту, який дозволяє отримати чіткий і зрозумілий прогноз тривалості життя, а також допомагає користувачам краще зрозуміти вплив своїх повсякденних звичок, стану здоров'я та соціально-економічних факторів на перспективу власного довголіття. Таким чином, подібні прогнози можуть стати важливим мотиваційним

фактором для позитивних змін у способі життя користувачів, спрямованих на збільшення тривалості життя та покращення його якості.

2.5 Висновки до другого розділу

На підставі проведеного дослідження та розробки веб-застосунку «Калькулятор життя» визначено, що створений інструмент повністю відповідає сучасним вимогам до цифрових медичних платформ та успішно реалізує персоналізоване прогнозування тривалості життя. Використана трирівнева архітектура (бізнес-рівень, прикладний рівень та технологічний рівень) забезпечує високу функціональність, зручність використання, а також необхідний рівень безпеки даних користувачів.

Реалізовані в рамках веб-застосунку інтерфейсні рішення дозволяють ефективно збирати, аналізувати та візуалізувати персональні медичні, соціальні та поведінкові дані, що значно спрощує процес взаємодії користувачів з системою та сприяє отриманню максимально об’єктивних і достовірних прогнозів.

Запропонована інтеграція з платформами цифрового здоров’я, такими як HealthKit і Google Fit, а також із носимими пристроями, забезпечує безперервне автоматичне оновлення медичних показників у реальному часі. Це підвищує точність та актуальність прогнозів, а також робить веб-застосунок привабливим для широкої аудиторії користувачів.

Особливої уваги заслуговує використання алгоритмів штучного інтелекту для оптимізації роботи з графічними елементами інтерфейсу, що значно підвищує продуктивність веб-застосунку та забезпечує комфортну взаємодію навіть за умов значного навантаження.

Таким чином, створений веб-застосунок є перспективним інструментом для застосування у сфері охорони здоров’я, який здатний ефективно використовувати сучасні технології для підвищення якості життя населення та підтримки прийняття обґрунтованих медичних і соціальних рішень.

Таким чином, другий розділ роботи підтверджує практичну реалізованість запропонованої інтегрованої технології прогнозування тривалості життя та її високу ефективність для використання в системах охорони здоров’я та профілактики захворювань.

РОЗДІЛ 3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ, ВАЛІДАЦІЯ ТА ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ

3.1 Методика та аналіз результатів експериментального дослідження

3.1.1 Методика проведення експерименту

З метою оцінки ефективності, точності та практичної придатності запропонованої інтегрованої математичної моделі та технології прогнозування тривалості життя, було проведено масштабне експериментальне дослідження із залученням реальних даних українського населення. Дослідження проводилось за допомогою веб-застосунку «Калькулятор життя», що був розроблений за підтримки ЦГЗ МОЗ України.

Експериментальне дослідження ефективності запропонованої технології проводилося на основі даних, зібраних через офіційний веб-портал ЦГЗ МОЗ України шляхом анонімного опитування респондентів через веб-інтерфейс застосунку.

Отримані в процесі опитування дані автоматично передавались на сервер, де проходили додаткову валідацію та попередню обробку, включаючи виявлення й усунення некоректних чи явно помилкових відповідей. Після цього проводився розрахунок очікуваної тривалості життя та ймовірності виживання, з використанням запропонованої математичної моделі персоналізованого прогнозування.

Важливою частиною експерименту стала статистична обробка даних. Для оцінки точності та адекватності моделі використовувалися методи описової статистики (середнє значення, стандартне відхилення, медіана, процентилі) та інші спеціалізовані статистичні інструменти. Було визначено розподіл прогнозованої тривалості життя за статтю та віком, що дозволило проаналізувати точність і релевантність прогнозних значень у порівнянні із середніми статистичними показниками населення України.

Крім того, для оцінювання впливу окремих факторів на прогноз, а також для навчання більш точних моделей «Калькулятору життя» LCR та LCX, використовувалися методи регресійного та факторного аналізу, а також машинне навчання (Random Forest, XGBoo відповідно). Оптимальні гіперпараметри алгоритмів були визначені за допомогою методу Grid Search із використанням 5-кратної перехресної валідації (Табл. 7), що забезпечило баланс між складністю моделей та точністю прогнозів. Діапазони значень для гіперпараметрів були обрані на основі рекомендацій з профільної літератури (посилання) та емпірично перевірені на реальних даних. Існує потенційна небезпека перетренованості моделей, особливо для LCX. Для зменшення цього ризику було застосовано регуляризацію параметрів (глибина дерев, кількість вибірок на вузол) та перехресну валідацію. Слід також враховувати можливі похибки вхідних даних, наприклад, пропущені значення, які оброблялися методом імпутації середніми або медіанами показників.

Таблиця 7. Оптимальні гіперпараметри алгоритмів Random Forest та XGBoost для моделей LCR та LCX відповідно, визначені на основі реальних даних

Алгоритм	Гіперпараметр	Діапазон тестування	Оптимальне значення (фактичне)
	Кількість дерев (n_estimators)	50—300	200

Random Forest	Максимальна глибина (max_depth)	3—20	9
	Мінімум зразків у вузлі (min_samples_leaf)	1—10	3
XGBoost	Швидкість навчання (learning_rate)	0.01—0.3	0.05
	Максимальна глибина дерева (max_depth)	3—10	3
	Мінімальна вага вузла (min_child_weight)	1—10	6

Це дозволило не лише оцінити точність загального прогнозу, але й визначити, які саме фактори мають найбільший вплив на тривалість життя українського населення.

Важливим аспектом експерименту було також оцінювання користувацького досвіду, що проводилося шляхом аналізу часу заповнення анкети, рівня завершення опитувань, частоти повернення користувачів до застосунку, а також аналізу зворотного зв’язку щодо зручності та зрозумілості веб-інтерфейсу.

Протягом першого півріччя після запуску веб-застосунку «Калькулятор життя» було завершено 200,454 анкети із загальної кількості 260,379 розпочатих сесій. Завершеними вважалися анкети, де користувачі дійшли до фінальної сторінки результатів у проміжку від 30 до 1200 секунд після початку опитування (Табл. 8). В середньому час завершення опитувальника становив 4 хвилини 15 секунд (255 секунд).

Таблиця 8. Статистичні характеристики часу завершення проходження опитувальника.

Характеристика	Час, секунд
Середнє значення (mean)	255

Медіана (p50)	229
Стандартне відхилення (std)	140
Мінімальне значення	30
Максимальне значення	1200
90-й процентиль (p90)	407
95-й процентиль (p95)	499
99-й процентиль (p99)	808

Серед відвідувачів веб-застосунку переважали чоловіки (63,6%). Як і очікувалось, основну аудиторію продукту склала молодь віком від 20 до 30 років (70,6%). Водночас суттєвою була частка респондентів вікової групи від 30 до 60 років, яка становила 25,6%. Важливим методологічним обмеженням дослідження є те, що вибірка для аналізу сформована з користувачів веб-застосунку, які добровільно пройшли анкетування. Це призвело до того, що в ній переважають особи молодшого та середнього віку, які в середньому є більш здоровими, ніж загальна популяція України. Як наслідок, отримані прогнози можуть бути дещо завищеними, якщо їх екстраполювати на всю популяцію, особливо для осіб старших вікових груп або тих, хто має тяжкі хронічні захворювання.

Для підвищення репрезентативності та валідності прогнозів у майбутніх дослідженнях рекомендується розширити вибірку, включивши більшу кількість представників старших вікових груп, а також осіб з різними станами здоров'я, включаючи хронічні захворювання різної тяжкості. Це дозволить отримати точніші та більш узагальнювані результати.

Отримані статистичні характеристики (Табл. 9) підтверджують високу ефективність та практичну корисність запропонованого рішення, а також дозволяють констатувати значні відмінності в очікуваній тривалості життя між чоловіками та жінками, що узгоджується з загальною демографічною ситуацією в Україні.

Таблиця 9. Статистичні характеристики результатів очікуваної тривалості життя респондентів за варіантами запропонованої моделі: LC19, LC20, LCFX та LCR

Параметр	Вік, роки							
	Чоловіки				Жінки			
	LC19	LC20	LCX	LCR	LC19	LC20	LCX	LCR
Mean	73,6	72,1	72,9	72,9	82,8	82,1	82,4	82,4
p50	74	73	73,7	73,6	84	83	83,2	83,1
STD	4,6	5,4	4,8	4,8	3,7	4,0	3,7	3,7
Min	42	41	41,4	42,6	66	65	66,9	68,5
Max	101	101	99,5	97,4	98	98	96,2	95
p90	78	78	78	78	87	87	86,6	86,7
p95	79	79	78,9	78,9	87	87	87,2	87
p99	80	80	80,1	80	88	88	88,2	88

Отримані результати моделювання демонструють дещо вищі значення очікуваної тривалості життя порівняно з офіційними даними Державної служби статистики України за 2020 рік. Однією з причин таких відхилень є особливості вибірки, яка була використана для побудови моделей. Зокрема, онлайн-опитувальник переважно проходила здорова молодь та люди середнього віку, що не мають значних проблем зі здоров’ям або тяжких хронічних захворювань.

Таким чином, запропонована методика експериментального дослідження дозволила комплексно оцінити практичну ефективність моделі та веб-застосунку «Калькулятор життя». Отримані результати підтвердили доцільність та ефективність запропонованих рішень, а також визначили напрями для подальшого удосконалення моделі та технології прогнозування.

3.1.2 Аналіз точності прогнозів на реальних даних українського населення

Одним із найважливіших етапів експериментального дослідження є оцінювання точності прогнозів, отриманих з використанням розробленої інтегрованої математичної моделі у веб-застосунку «Калькулятор життя». Аналіз проводився на вибірці, яка включає понад 200 тисяч анкет реальних користувачів з України. Це дозволило отримати репрезентативні результати, що відображають різноманітні соціально-демографічні та медичні характеристики українського населення. Загальний обсяг є значним і забезпечує статистичну потужність результатів дослідження. Проте слід враховувати, що в структурі вибірки присутнє певне зміщення: більшість респондентів представляють молодші вікові групи та не мають серйозних хронічних захворювань. Це обмежує репрезентативність отриманих результатів щодо старших вікових груп та осіб із важким станом здоров'я. Таким чином, при узагальненні результатів на все населення України необхідно враховувати зазначені обмеження. Подальші дослідження рекомендується спрямувати на вирівнювання структури вибірки за віком та станом здоров'я, щоб зробити її максимально репрезентативною.

Для оцінки точності прогнозних результатів було застосовано декілька підходів. Спочатку було проведено порівняння прогнозованих показників очікуваної тривалості життя, отриманих з моделі, із середніми статистичними показниками тривалості життя українців, отриманими від Державної служби статистики України та інших авторитетних джерел. Це дозволило визначити загальну відповідність прогнозів реальним демографічним тенденціям у різних вікових та статевих групах.

Цей фактор створив природний відбір респондентів, що зумовило зміщення результатів у бік більш сприятливих показників очікуваної тривалості життя. Отже, виявлені відхилення від загальнодержавних статистичних даних значною мірою пояснюються специфікою вибірки та необхідністю подальших досліджень із залученням більш репрезентативної групи респондентів, яка включала б осіб різних вікових груп і з різним станом здоров’я.

Таким чином, отримані результати слід розглядати як індикативні, що демонструють потенційні можливості інструменту при його застосуванні до певної цільової групи населення, і враховувати цей фактор при інтерпретації та практичному застосуванні моделей прогнозування тривалості життя.

Наступним кроком було здійснено комплексний порівняльний аналіз якості прогнозування тривалості життя за допомогою розроблених у дослідженні моделей, базових моделей, а також низки міжнародних аналогів. Для додаткової валідації та порівняння результатів запропонованої моделі, зібрані через веб-застосунок анкетні дані (у тих питаннях, які мають прямі аналоги) були конвертовані та завантажені у відкриті міжнародні моделі прогнозування тривалості життя MPoRT версії 1.19.4, GBD та ONS (Табл. 9). Отримані результати дозволили здійснити порівняльний аналіз основних статистичних показників прогнозованої тривалості життя.

Таблиця 9. Статистичні характеристики результатів очікуваної тривалості життя респондентів пораховані на MPoRT, GBD та ONS

Параметр	Вік, роки					
	Чоловіки			Жінки		
	MPoRT	GBD	ONS	MPoRT	GBD	ONS
Mean	82	91,5	80,9	77,4	91,5	84,3
p50	81,6	92	80,4	73,5	91,9	83,8
STD	15,7	1,4	1,5	16	1,5	1,3
Min	28,3	89,1	79,9	28,4	89,1	83,4
Max	124,9	105,9	99,9	124,3	102,9	98,2
p90	109,1	93,2	82,2	109,8	93,2	85,1
p95	111,1	93,3	83,6	111	93,4	85,9
p99	113	94,3	87,5	113,8	94,7	89,6

Порівняння результатів прогнозування тривалості життя з використанням моделей MPoRT, GBD та ONS вказує на суттєві розбіжності у показниках точності та варіативності прогнозів. Найбільшу варіативність демонструє модель MPoRT, що підтверджується значно вищими значеннями стандартного відхилення (STD): 15,7 років у чоловіків та 16 років у жінок. Моделі GBD та ONS, навпаки, характеризуються суттєво нижчою варіативністю (STD: 1,4 і 1,5 років для чоловіків та 1,5 і 1,3 років для жінок відповідно).

Аналіз перцентилів (p90, p95, p99) також демонструє значні відмінності: модель MPoRT генерує істотно вищі екстремальні значення, що свідчить про тенденцію до нереалістично високих прогнозів для певної частини населення. Крім того, середня очікувана тривалість життя за моделлю MPoRT виявилась парадоксально нижчою для жінок (77,4 роки) порівняно з чоловіками (82 роки), що не відповідає типовим демографічним трендам України, де жінки зазвичай мають вищу тривалість життя.

Таким чином, модель MPoRT продемонструвала недостатню адаптацію до особливостей української популяції. Водночас моделі GBD та ONS показали реалістичніші та стабільніші результати, хоча і з суттєво меншою варіативністю, що обмежує їхню здатність до глибокої персоналізації прогнозів. Тому найбільшу практичну цінність має запропонована у дослідженні власна інтегрована модель, яка оптимально поєднує точність, стабільність і адаптивність до українського контексту.

Для кількісної оцінки моделей використовувалися стандартні регресійні метрики якості: RMSE (корінь середньоквадратичної похибки), MAE (середня абсолютна похибка), MSE (середньоквадратична похибка), MAPE (середня абсолютна похибка у відсотках), коефіцієнт детермінації R^2 та C-index (індекс конкордантності). Ці метрики були розраховані на тестових вибірках, сформованих за методом крос-валідації, що дозволило оцінити стабільність та точність прогнозів у різних випадках (Табл. 10).

Таблиця 10. Порівняння регресійних метрик різних версій запропонованої моделі з моделями MPoRT, ONS та GBD

Модель	RMSE (↓)	MAE (↓)	R^2 (↑)	MAPE (↓)	MSE (↓)	C-index (↑)
LCX	5.660	4.711	-1.317	6.284	32.035	0.707
LCR	5.685	4.756	-1.337	6.346	32.315	0.705
LC19	5.794	4.862	-1.428	6.491	33.565	0.706
LC20	5.811	4.842	-1.443	6.471	33.773	0.705
ONS	7.856	7.636	-3.464	10.324	61.723	0.765
MPoRT	16.426	13.028	-18.515	17.242	269.828	0.456
GBD	16.971	16.555	-19.830	22.369	288.014	0.545

Аналізуючи отримані результати, можна виокремити декілька чітких тенденцій:

1. Найвищу точність прогнозування серед усіх порівнюваних моделей продемонструвала запропонована у роботі модель LCX, побудована на основі алгоритму XGBoost. Для цієї моделі зафіксовані найнижчі значення більшості ключових метрик помилок: RMSE = 5,660, MAE = 4,711, MSE = 32,035, MAPE = 6,284%. Також модель характеризується високим значенням C-index (0,707), що свідчить про її гарну дискримінативну здатність. При цьому значення коефіцієнта детермінації R^2 має від'ємне значення (R^2 = -1,317), що пояснюється специфікою розподілу даних та обмеженнями самої метрики R^2 в контексті даного набору даних. Попри це, низькі показники похибок та високий C-index підтверджують високу ефективність моделі LCX для прогнозування тривалості життя.

2. Дуже близькою за показниками ефективності є модель LCR, побудована з використанням алгоритму Random Forest. Для неї значення RMSE становить 5,685, MAE – 4,756, MSE – 32,315, MAPE – 6,346%, а C-index дорівнює 0,705. Водночас значення коефіцієнта детермінації також має негативне значення (R^2 = -1,337), що, аналогічно попередній моделі, є наслідком особливостей використаних даних. Високі значення інших метрик свідчать про високу точність та стабільність

прогнозів цієї моделі. Крім того, прозорість та інтерпретованість алгоритму Random Forest робить модель LCR привабливою для практичного застосування.

3. Окремо слід розглянути результати базових моделей, побудованих на загальнодоступних демографічних таблицях смертності для України. Так, модель на базі таблиці смертності ВООЗ за 2019 рік (LC19) продемонструвала RMSE = 5,794, MAE = 4,862, MAPE = 6,587%, з коефіцієнтом детермінації R^2 = -1,428. Модель на основі таблиці Держстату України за 2020 рік (LC20) має RMSE = 5,811, MAE = 4,842, MAPE = 6,591%, та коефіцієнт детермінації R^2 = -1,443. Незважаючи на те, що ці моделі трохи поступаються за точністю моделям на базі машинного навчання, вони демонструють високий рівень якості прогнозів на популяційному рівні. Їхні показники точності є достатньо високими для широкого практичного застосування, особливо у випадках, коли індивідуальні дані недоступні або їх обсяг обмежений. Це підкреслює важливість та цінність використання таких моделей у демографічних і соціально-економічних дослідженнях.

4. Значно нижчу ефективність у контексті українського населення продемонстрували міжнародні моделі прогнозування – MPoRT, GBD та ONS**.** Наприклад, модель MPoRT має найгірші показники серед аналізованих моделей: RMSE = 16,426, MAE = 13,028, MAPE = 21,194%, а коефіцієнт детермінації R^2 = -18,515. Модель GBD показала RMSE = 16,971, MAE = 16,555, MAPE = 25,121%, R^2 = -19,830. Аналогічно, модель ONS має RMSE = 7,856, MAE = 7,636, MAPE = 12,357%, R^2 = -3,464. Такі значення свідчать про те, що міжнародні моделі, які є ефективними в інших країнах, без локальної адаптації недостатньо точно прогнозують тривалість життя в умовах України через суттєві відмінності у соціально-демографічних і медичних особливостях.

Таким чином, результати аналізу однозначно свідчать на користь запропонованих моделей LCX (XGBoost) та LCR (Random Forest), які, попри особливості отриманих значень коефіцієнта детермінації, демонструють високий рівень точності прогнозу за основними метриками якості. Саме ці моделі є найбільш придатними до застосування в українському контексті та рекомендовані

для подальших практичних розрахунків і використання у клінічних, соціально-демографічних і економічних дослідженнях.

На підставі наведених результатів рекомендується використовувати саме ці дві моделі для подальших практичних розрахунків та прогнозних досліджень в українському контексті, оскільки саме вони враховують локальні особливості та продемонстрували високу ефективність в експериментальному оцінюванні. Водночас необхідно усвідомлювати, що такий підхід до валідації має певні обмеження, зокрема:

- Оцінка точності базується лише на поточних демографічних даних і не дозволяє повністю оцінити точність прогнозу для конкретної людини на довгострокову перспективу (декілька десятків років).
- Вибірка користувачів веб-застосунку може бути нерепрезентативною щодо окремих категорій населення, особливо осіб старшого віку чи з хронічними важкими станами.

Для посилення валідності моделей у майбутніх дослідженнях планується проведення додаткової зовнішньої валідації з використанням незалежних вибірок населення, отриманих з інших джерел (наприклад, клінічних баз даних або національних медичних реєстрів). Це дозволить краще оцінити точність і узагальнюваність розробленої технології прогнозування на всю популяцію.

3.1.3 Аналіз впливу факторів на тривалість життя за результатами експерименту

Одним з ключових аспектів проведеного експериментального дослідження стало виявлення та кількісна оцінка впливу різних факторів на прогнозовану тривалість життя українського населення. Завдяки масштабній вибірці реальних даних, зібраних через веб-застосунок «Калькулятор життя», вдалося провести поглиблений факторний аналіз із використанням статистичних та алгоритмічних підходів, зокрема методів машинного навчання.

У рамках цього аналізу було визначено декілька груп факторів, які суттєво впливають на прогноз тривалості життя. До них належать медичні фактори (наявність та характер хронічних захворювань, частота звернень до медичних закладів, показники серцево-судинної системи), поведінкові фактори (паління, вживання алкоголю, фізична активність, якість харчування, режим сну), антропометричні показники (індекс маси тіла, співвідношення талії до зросту), а також соціально-демографічні характеристики (вік, стать, рівень доходу, освіта, статус ВПО, регіон проживання).

Для кількісного аналізу впливу цих факторів було використано алгоритми машинного навчання, зокрема випадковий ліс (Random Forest) та градієнтний бустинг (XGBoost), які дозволяють оцінювати важливість кожного фактора у прогнозній моделі. В результаті проведеного аналізу було визначено, що найбільший вплив на тривалість життя українців мають саме медичні та поведінкові фактори.

Зокрема, фактори, пов’язані з палінням та регулярним вживанням алкоголю, виявились одними з найбільш значущих і негативно впливають на прогнозовану тривалість життя, знижуючи її в середньому на 5—12 років, залежно від віку та статі. Водночас фізична активність, дотримання правильного харчування та нормального режиму сну асоціюються з істотним підвищенням прогнозованої тривалості життя, в середньому на 3—8 років.

Також суттєво впливають на прогнозну модель такі медичні показники, як наявність хронічних серцево-судинних захворювань, діабету, високий артеріальний тиск та зайва вага. Ці фактори мають значний негативний вплив, особливо в старших вікових категоріях, і можуть скорочувати прогнозовану тривалість життя на 4—15 років залежно від конкретного стану та його важкості.

Соціально-демографічні фактори також мають істотний, хоча й дещо менший вплив. Наприклад, люди з вищим рівнем освіти та доходу мають прогнозовану тривалість життя, що на 2—5 років вища за середню, що може бути пов'язано з кращим доступом до медичних послуг, профілактики та більш

свідомим ставленням до власного здоров’я. Водночас особи, що мають статус внутрішньо переміщених осіб (ВПО), часто стикаються зі стресом і погіршенням соціально-економічних умов життя, що призводить до скорочення прогнозованої тривалості життя на 3—6 років порівняно із загальною популяцією.

Важливо зазначити, що алгоритми машинного навчання дозволили також виявити нелінійні та комплексні залежності між факторами. Наприклад, комбінація паління і низької фізичної активності значно посилює негативний вплив на тривалість життя, тоді як помірна фізична активність частково знижує негативний вплив інших факторів, таких як зайва вага або хронічні захворювання.

Таким чином, отримані результати експерименту дозволяють зробити важливий висновок щодо значущості факторів ризику та їхнього комплексного впливу на прогнозовану тривалість життя українців. Виявлені взаємозв’язки та кількісні оцінки можуть слугувати основою для розробки ефективних рекомендацій у сфері громадського здоров’я, спрямованих на зниження впливу негативних факторів та популяризацію здорового способу життя серед українського населення.

3.2 Оцінювання користувацького досвіду та зручності використання веб-застосунку

Окрему увагу в рамках експериментального дослідження було приділено оцінюванню користувацького досвіду (User Experience, UX) та зручності використання веб-застосунку «Калькулятор життя». Важливим фактором, який визначає успішність та популярність таких інструментів, є не лише точність прогнозів, але й рівень задоволеності користувачів, зрозумілість інтерфейсу, швидкість і комфорт взаємодії з системою.

Для оцінювання користувацького досвіду застосовано комбінований підхід, який поєднував кількісний аналіз поведінкових метрик користувачів та якісний аналіз, що включав опитування й аналіз зворотного зв’язку. Кількісно оцінювались такі параметри, як середній час заповнення анкети, відсоток користувачів, які

завершили опитування, а також частка користувачів, які одразу після першого прогнозу змінювали початкові відповіді для отримання альтернативних результатів.

Аналіз понад 200 тисяч анкет показав, що середній час, який витрачав користувач на заповнення анкети, становив приблизно 2—4 хвилини, що свідчить про високу ефективність і простоту використання інтерфейсу. При цьому близько 85% користувачів, які почали заповнювати анкету, успішно завершили її, що підтверджує комфортність і зрозумілість процесу.

Цікавим показником зацікавленості й залученості користувачів стала їх активність після першого прогнозу: близько 35% користувачів одразу намагалися змінити початкові відповіді, щоб порівняти отримані результати прогнозування з альтернативними варіантами. Це свідчить про високу зацікавленість користувачів у результатах прогнозу та мотивацію до аналізу впливу окремих факторів на власну тривалість життя.

У рамках якісного аналізу додатково було проведено онлайн-опитування серед вибіркової групи зі 100 респондентів. Результати опитування показали, що понад 90% користувачів оцінили свій досвід взаємодії із веб-застосунком як позитивний або дуже позитивний. Найчастіше користувачі відзначали переваги, такі як простота й зрозумілість інтерфейсу та швидке отримання результатів.

Окремо користувачі відзначали корисність персоналізованих рекомендацій, які мотивують їх до змін у способі життя, зокрема зменшення шкідливих звичок, підвищення фізичної активності та покращення якості харчування. Водночас, деякі учасники опитування надали важливі пропозиції щодо подальших удосконалень застосунку, зокрема висловлюючи бажання бачити детальніші рекомендації та додаткову інтеграцію з іншими медичними та соціальними сервісами (наприклад, Helsi або «Дія»).

Отже, результати проведеного дослідження користувацького досвіду свідчать про те, що веб-застосунок «Калькулятор життя» демонструє високу зручність і ефективність використання. Високий рівень задоволеності

користувачів, підтверджений як кількісним, так і якісним аналізом, дозволяє рекомендувати розроблену технологію для подальшого впровадження і масштабування з метою покращення громадського здоров'я та індивідуального управління ризиками для українського населення.

3.3 Перспективні напрями розвитку технології прогнозування

3.3.1 Напрями вдосконалення розробленої моделі та методів

Подальший розвиток інтегрованої технології прогнозування тривалості життя, реалізованої у вигляді веб-застосунку «Калькулятор життя», передбачає кілька ключових напрямів, які дозволять суттєво підвищити точність, ефективність і масштабованість запропонованого рішення.

Подальшим перспективним напрямом розвитку запропонованої інтегрованої технології є реалізація автоматичного аналізу ЕКГ-сигналів з носимих пристроїв з метою своєчасного виявлення відхилень від індивідуальних норм користувачів. Розроблений підхід із застосуванням комбінованого аналізу ЕКГ на основі штучного інтелекту, що передбачає індивідуальну модель тренування та постійний моніторинг, дозволить оперативно і точно визначати ризики для здоров’я і, відповідно, уточнювати прогнози тривалості життя. Це відкриває нові перспективи як для індивідуального управління здоров’ям, так і для впровадження даної технології на державному рівні в межах національних програм профілактики серцево-судинних захворювань.

Для поглиблення персоналізованого аналізу в рамках розробленого інструменту запропоновано комбінований підхід до аналізу електрокардіографічних (ЕКГ) сигналів, що передбачає виявлення відхилень від індивідуальних або популяційних норм [101]. Такий підхід базується на застосуванні алгоритмів штучного інтелекту, зокрема глибоких нейронних мереж, здатних до адаптивного навчання на індивідуальних даних користувача, зібраних

за допомогою носимих пристроїв. Це дозволяє створювати точні моделі індивідуальних норм серцевої активності, які періодично оновлюються для врахування змін у стані здоров'я людини протягом часу.

Система аналізу, що розробляється, передбачає багаторівневу обробку ЕКГ-даних, починаючи з очищення сигналу в реальному часі безпосередньо на носимому пристрої та завершуючи детальним зовнішнім аналізом у випадку виявлення суттєвих відхилень. Це забезпечує можливість оперативного реагування на потенційні порушення роботи серцево-судинної системи та значно підвищує точність діагностики, що особливо важливо для осіб з хронічними захворюваннями та складним анамнезом.

Впровадження такого комбінованого аналізу суттєво доповнює можливості розробленого інструменту прогнозування тривалості життя, дозволяючи враховувати більш широкий спектр індивідуальних факторів та забезпечуючи високу точність персоналізованих прогнозів.

Ще одним важливим напрямом вдосконалення є застосування більш складних методів машинного навчання, зокрема глибоких нейронних мереж (Deep Learning), таких як згорткові (CNN) і рекурентні нейронні мережі (RNN). Ці алгоритми мають значний потенціал для аналізу складних і нелінійних взаємозв’язків між факторами ризику, які неможливо повною мірою врахувати класичними методами. Вони також дозволять більш ефективно виявляти приховані залежності й потенційні ризики на основі динамічних часових рядів даних, отримуваних з медичних пристроїв, таких як смарт-годинники та персональні ЕКГ-пристрої. Використання цих сучасних підходів дозволить суттєво зменшити похибки прогнозування і підвищити персоналізацію прогнозів.

У сучасних умовах використання методів штучного інтелекту (ШІ) відкриває нові можливості для підвищення точності та персоналізації прогнозування тривалості життя. Веб-застосунок «Калькулятор життя» використовує передові методи машинного навчання для аналізу великих обсягів медичних, поведінкових і

соціально-демографічних даних, що дозволяє суттєво покращити якість та надійність прогнозних оцінок.

Особливу увагу було приділено використанню нейронних мереж глибокого навчання (Deep Learning), що дозволяють виявляти складніші взаємозалежності і навіть прогнозувати потенційні медичні ризики, які не були явно зазначені користувачами. Наприклад, штучний інтелект може аналізувати сукупність показників фізичної активності, якості сну, ваги, артеріального тиску та інших факторів для виявлення прихованих патологічних станів або ризику розвитку захворювань ще на початковій стадії, навіть якщо користувач не має офіційно підтвердженого медичного діагнозу.

Ще одним аспектом застосування ШІ є можливість індивідуалізованого аналізу впливу факторів ризику для різних вікових і соціально-демографічних груп населення. Завдяки машинному навчанню, модель може автоматично розподіляти користувачів на кластери з подібними профілями ризику і надавати персоналізовані рекомендації, які ефективно враховують саме ті фактори, що найбільше впливають на тривалість життя для кожної групи.

Таким чином, інтеграція штучного інтелекту у веб-застосунок «Калькулятор життя» дозволяє суттєво підвищити рівень точності та персоналізації прогнозів, створюючи потужний інструмент для індивідуального аналізу ризиків та мотивації користувачів до покращення якості свого життя. Використання алгоритмів машинного навчання не лише знижує загальну похибку прогнозування, але й дозволяє ефективно виявляти й оцінювати приховані чинники, забезпечуючи високий рівень якості й довіри користувачів до результатів роботи системи.

Окремий акцент варто зробити на удосконаленні підходів до візуалізації результатів і рекомендацій. Вдосконалення UX/UI-інтерфейсу веб-застосунку шляхом використання інтерактивних та персоналізованих інформаційних панелей (dashboards), інтуїтивно зрозумілих графічних ілюстрацій, а також більш глибоких і детальних рекомендацій для користувачів дозволить суттєво підвищити їхню мотивацію та зацікавленість у регулярному використанні застосунку. Розширення

можливостей застосунку з використанням гейміфікації також є перспективним, оскільки дозволить значно підвищити залученість та утримання користувачів.

Окремим напрямом удосконалення є також підвищення безпеки і конфіденційності даних. Це може включати подальше посилення алгоритмів шифрування і псевдонімізації даних, розширення механізмів контролю та управління доступом до даних.

Крім того, подальше вдосконалення може охоплювати також створення спеціалізованих модулів та версій застосунку для окремих груп населення (наприклад, для людей похилого віку, для осіб із хронічними захворюваннями, військовослужбовців або ВПО), які дозволять краще враховувати специфіку та особливості різних категорій користувачів, надаючи їм максимально релевантні й корисні рекомендації.

Таким чином, подальший розвиток інтегрованої технології прогнозування тривалості життя передбачає комплекс заходів, що охоплюють технічне вдосконалення самої моделі, розширення набору даних і джерел їх отримання, покращення алгоритмів аналізу, підвищення безпеки даних, а також розвиток користувацького досвіду. Реалізація цих перспектив дозволить значно розширити можливості застосунку, підвищити його точність і практичну цінність, зробити технологію конкурентоспроможною як на українському, так і на міжнародному рівнях.

3.3.2 Перспективи інтеграції з медичними та державними сервісами

Важливим перспективним напрямом подальшого розвитку веб-застосунку «Калькулятор життя» є інтеграція з українськими цифровими платформами «Дія» та «Helsi». Ці сервіси вже зараз забезпечують широкий доступ до офіційних медичних та соціально-демографічних даних користувачів, що створює потенціал для повністю автоматизованого заповнення профілів без необхідності додаткового введення інформації з боку користувача.

Застосунок «Дія» є основною платформою цифрових державних послуг в Україні, яка надає доступ до електронних документів, персональних реєстрів, медичних сертифікатів та інших офіційних записів. Завдяки використанню чітко структурованого API, після відповідної авторизації користувачів через механізм «Дія.Підпис» або BankID, можлива миттєва й автоматична передача до «Калькулятора життя» демографічних даних (вік, стать, місце проживання) та соціально-економічних характеристик (освіта, рівень доходів, наявність статусу ВПО). Це дозволить значно спростити й прискорити процес формування персоналізованих прогнозів, усуваючи необхідність ручного введення цих базових даних користувачем, що мінімізує ймовірність помилок та забезпечує високу точність інформації.

Платформа «Helsi» є провідним цифровим медичним сервісом України, що забезпечує централізований доступ до електронних медичних записів, результатів лабораторних аналізів, історії захворювань та медичних призначень. Взаємодія «Калькулятора життя» з Helsi може бути повністю реалізована через спеціалізований захищений API, який Helsi вже успішно надає своїм сертифікованим партнерам.

Після авторизації користувачів і отримання їхньої згоди на передачу інформації, інформація щодо діагностованих захворювань, даних медичних аналізів та регулярність відвідувань медичних установ може бути автоматично і в достатньому обсязі передана до веб-застосунку

Таким чином, інтеграція з Helsi дозволить користувачам отримати готові профілі без будь-якого додаткового ручного введення медичних даних. Це не тільки суттєво спростить взаємодію з веб-застосунком, а й гарантуватиме точність і актуальність прогнозних розрахунків, що базуватимуться виключно на офіційних медичних даних.

Переваги повної автоматизації заповнення анкетних даних:

- Мінімізація людського фактору, що значно знижує ймовірність помилок при введенні інформації.

- Скорочення часу, необхідного користувачам для отримання прогнозу, що підвищує загальний рівень зручності й задоволеності застосунком.
- Підвищення якості прогнозів за рахунок використання лише перевірених, офіційних та актуальних даних, що отримуються автоматично з відповідних цифрових платформ.

Таким чином, завдяки інтеграції з державним мобільним застосунком «Дія» та медичною платформою «Helsi», веб-застосунок «Калькулятор життя» зможе максимально ефективно використовувати вже наявні в офіційних реєстрах та базах дані користувачів, забезпечуючи повністю автоматизоване формування персоналізованих прогнозів без додаткових зусиль зі сторони кінцевих користувачів. Це дозволить досягти високого рівня довіри, зручності та практичної ефективності застосунку в умовах українських реалій.

3.3.3 Адаптація технології до міжнародних стандартів та використання в інших країнах

Важливим перспективним напрямом розвитку запропонованої інтегрованої технології прогнозування тривалості життя є її подальша адаптація до міжнародних стандартів та розширення можливостей використання за межами України. На сьогодні веб-застосунок «Калькулятор життя» вже відповідає основним вимогам Європейського регламенту щодо захисту персональних даних (GDPR), що дозволяє безпечно і законно обробляти дані громадян Європейського Союзу.

Зокрема, реалізовано механізми керування згодами користувачів, права на доступ, редагування та видалення персональних даних, а також впроваджено суворі стандарти безпеки, включно з шифруванням та псевдонімізацією інформації. Це дозволяє потенційно розширювати аудиторію веб-застосунку на країни ЄС без додаткових складних юридичних чи організаційних перешкод.

Важливим аспектом адаптації технології до міжнародного використання є її подальше узгодження із вимогами таких стандартів, як HIPAA (Health Insurance Portability and Accountability Act), що актуальні для США та Канади. Забезпечення відповідності HIPAA дозволить розширити використання технології в країнах Північної Америки, які мають високі вимоги щодо захисту медичних даних. Для цього необхідно впровадити додаткові заходи щодо безпеки зберігання та обробки інформації, а також передбачити регулярні зовнішні аудити й сертифікації відповідно до цього стандарту.

Для покращення можливостей інтеграції з міжнародними медичними установами та сервісами важливим напрямом є впровадження міжнародних стандартів обміну медичними даними, зокрема HL7 і FHIR (Fast Healthcare Interoperability Resources). Використання цих стандартів дозволить значно полегшити взаємодію із зовнішніми медичними інформаційними системами та забезпечить можливість інтеграції з локальними медичними сервісами в інших країнах.

Щодо локалізації веб-застосунку, то на сьогодні вже реалізовано англомовну версію інтерфейсу, що створює додаткові можливості для міжнародного використання технології без додаткових витрат на переклад. Подальший розвиток може передбачати розширення кількості мовних локалізацій з урахуванням цільових ринків і регіонів, таких як німецька, французька, іспанська, а також мови азійських регіонів (наприклад, китайська, японська, корейська). Це сприятиме значному розширенню міжнародної аудиторії користувачів.

Окремим важливим аспектом міжнародної адаптації є врахування місцевих культурних, соціальних і медичних особливостей різних регіонів світу. Це може потребувати додаткових досліджень та коригування алгоритмів прогнозування з урахуванням регіональної специфіки щодо поширеності певних захворювань, поведінкових та екологічних факторів, характерних для конкретних країн або регіонів.

З технічної точки зору важливим елементом міжнародного масштабування є подальший розвиток модульної архітектури веб-застосунку. Модульний підхід дозволяє оперативно налаштовувати і адаптувати технологію до локальних потреб, легко інтегрувати різні джерела медичних і соціальних даних, а також забезпечувати масштабованість і гнучкість застосунку при виході на нові ринки.

Таким чином, подальша адаптація розробленої інтегрованої технології до міжнародних стандартів і використання в інших країнах є перспективним напрямом її розвитку. Реалізація цього потенціалу дозволить значно розширити сферу використання веб-застосунку, зробивши його глобальним інструментом персоналізованого прогнозування та управління ризиками здоров'я, що відповідає найвищим міжнародним стандартам безпеки, конфіденційності й ефективності.

3.4 Висновки до третього розділу

Проведене експериментальне дослідження підтвердило високу ефективність та практичну цінність розробленої інтегрованої технології прогнозування тривалості життя. Вибірка дослідження включала понад 260 тисяч користувачів та 200 тисяч заповнених анкет, що забезпечило певну репрезентативність отриманих результатів та їхню достовірність.

Детальний аналіз результатів експериментального дослідження показав, що всі чотири моделі, розроблені в межах веб-застосунку «Калькулятор життя» (LC19, LC20, LCX, LCR), демонструють суттєво вищу точність прогнозування порівняно з міжнародними моделями MPoRT, GBD та ONS.

За результатами порівняння, найкращу ефективність продемонстрували моделі LCX та LCXR, побудовані на алгоритмах XGBoost і Random Forest відповідно. Попри негативні значення коефіцієнта детермінації R^2 (LCX: $R^2 = -1.317$; LCR: $R^2 = -1.337$), ці моделі мають найменші значення інших показників похибки. Отримані негативні значення коефіцієнта детермінації (R^2) у всіх моделях пояснюються високою індивідуальною варіативністю та складністю

прогнозування очікуваної тривалості життя. Модель LCX продемонструвала найнижчий RMSE (5.660), MAE (4.711), MSE (32.035) та MAPE (6.284%). Модель LCR показала дуже близькі результати (RMSE = 5.685, MAE = 4.756, MSE = 32.315, MAPE = 6.346%). Це вказує на високу стабільність та точність цих моделей у прогнозуванні тривалості життя на індивідуальному рівні для української вибірки.

Базові демографічні моделі LC19 та LC20 показали дещо нижчу ефективність. Зокрема, модель LC19 року має R^2 = -1.428, RMSE = 5.794, MAE = 4.862, а модель LC20 – R^2 = -1.443, RMSE = 5.811, MAE = 4.842. Хоча ці моделі демонструють трохи гірші результати, ніж моделі на основі машинного навчання, їх використання залишається прийнятним для загальних популяційних прогнозів, особливо в умовах обмеженої доступності індивідуальних даних.

Особливу увагу слід звернути на міжнародні моделі прогнозування MPoRT, ONS та GBD, які мають суттєво гірші результати порівняно з локально адаптованими моделями. Наприклад, модель MPoRT має значно нижчий коефіцієнт детермінації (R^2 = -18.515), RMSE = 16.426, MAE = 13.028, MSE = 269.828, що свідчить про її низьку адаптованість до української популяції. Аналогічні низькі показники демонструє модель GBD (R^2 = -19.830, RMSE = 16.971, MAE = 16.555) та ONS (R^2 = -3.464, RMSE = 7.856, MAE = 7.636). Це підтверджує, що міжнародні моделі без адаптації до специфічних умов України не здатні забезпечити прийнятну точність прогнозування і не можуть бути рекомендовані до прямого використання у вітчизняних умовах.

Таким чином, результати експериментального оцінювання однозначно підтверджують перевагу моделей LCX і LCR у контексті української популяції, незважаючи на особливості отриманих значень R^2. Ці моделі найбільш точно відображають індивідуальні характеристики населення України, враховують специфічні особливості української популяції та мають достатній потенціал для подальшого використання у практичних застосунках та дослідженнях.

Окрім цього, моделі Калькулятора життя забезпечили реалістичні демографічні співвідношення, де середня очікувана тривалість життя у жінок була

вищою, ніж у чоловіків, що повністю відповідає статистичним даним та демографічним реаліям України. Це додатково підкреслює перевагу запропонованих моделей щодо адаптації до національного контексту.

Користувацький досвід також підтвердив успішність запропонованого рішення: середній час заповнення анкети склав 4 хвилини 15 секунд, а рівень завершення — близько 77%. Ці показники підтверджують високу зручність, доступність та ефективність інтерфейсних рішень веб-застосунку.

Перспективи подальшого розвитку запропонованої технології включають інтеграцію з державними та міжнародними медичними сервісами, адаптацію до міжнародних стандартів, а також розширення функціональності завдяки автоматичному аналізу даних із носимих пристроїв, що дозволить проводити оперативну діагностику та профілактику захворювань. Реалізація цих напрямів забезпечить подальше зростання ефективності технології прогнозування та її масштабування у сфері охорони здоров’я.

Таким чином, результати експериментального дослідження і валідації підтвердили високу точність, практичну цінність та соціальну значущість розробленої технології прогнозування тривалості життя, а запропоновані перспективи розвитку формують передумови для її широкого впровадження в системах громадського здоров’я.

ЗАГАЛЬНІ ВИСНОВКИ

У дисертаційній роботі вирішено актуальні науково-практичні завдання розроблення інтегрованої технології персоналізованого прогнозування тривалості життя для українського населення на основі комплексного аналізу демографічних, соціальних та медичних факторів із застосуванням сучасних методів машинного навчання та цифрових технологій:

1. **Проведено аналіз сучасних методів прогнозування очікуваної тривалості життя.** У дисертаційній роботі здійснено всебічний системний аналіз традиційних та новітніх моделей прогнозування тривалості життя. Це дозволило обґрунтувати переваги інтегрованих підходів, які використовують великі обсяги різнорідних персональних даних. Встановлено, що застосування сучасних алгоритмів машинного навчання (таких як Random Forest, XGBoost) суттєво підвищує точність і надійність прогнозів порівняно зі стандартними статистичними методами.

2. **Визначено і формалізовано основні демографічні, соціальні та медичні фактори для побудови інтегрованої моделі прогнозування очікуваної тривалості життя.** Виділено основні чинники, які найбільше впливають на очікувану тривалість життя людини, та здійснено їх формалізацію для використання в моделі. До найвагоміших факторів належать: наявність хронічних захворювань, рівень фізичної активності, паління, вживання алкоголю, а також важливі соціально-демографічні характеристики (вік, стать, рівень освіти тощо). Кількісне представлення і врахування цих факторів у моделі створило основу для побудови інтегрованого прогнозування тривалості життя з врахуванням індивідуального профілю ризику.

3. **Розроблено математичну модель персоналізованого прогнозування на основі інтеграції різних типів даних та сучасних алгоритмів машинного навчання.** Створено інтегровану математичну модель прогнозування тривалості життя у чотирьох варіаціях (LC19, LC20, LCX, LCR), які поєднують детерміністичні алгоритми та сучасні алгоритми машинного навчання. Модель дозволяє оцінювати індивідуальний вплив кожного фактора ризику на очікувану тривалість життя конкретної особи. Інтеграція багатовимірних даних у поєднанні з методами ML забезпечила підвищення точності, деталізації та персоналізації прогнозів порівняно з традиційними підходами.

4. **Створено алгоритми прогнозування з урахуванням специфіки українського населення в умовах війни.** Розроблено алгоритми динамічного (адаптивного) оновлення прогнозу, які враховують специфічні особливості української популяції. Ці алгоритми автоматично коригують прогноз при надходженні нових даних про користувача або при зміні умов середовища (стану здоров'я, соціально-економічної ситуації тощо). Такий підхід забезпечує актуальність і гнучкість моделей прогнозування навіть за різких змін профілю ризиків населення, що особливо важливо в умовах війни. У результаті забезпечено високу адаптивність та точність прогнозів для українського населення в динамічних умовах.

5. **Модель реалізовано як веб-застосунок та перевірено на реальних даних українського населення.** Розроблену інтегровану модель впроваджено у вигляді практичного веб-застосунку «Калькулятор життя». Також створено мобільні додатки для iOS та Android, інтегровані з платформами цифрового здоров'я (Apple HealthKit, Google Fit) і носимими пристроями для автоматизованого збору медичних показників. Проведено експериментальне дослідження на реальних даних українського населення, яке підтвердило працездатність і ефективність запропонованої технології. Веб-застосунок продемонстрував високий рівень зручності для користувачів та ефективності отримання прогнозів тривалості життя, про що свідчить значна кількість користувачів (понад 260 тисяч) і високий відсоток завершення анкети (близько 77% користувачів доводять ввід даних до кінця для отримання результату).

6. **Оцінено ефективність, точність і зручність технології, а також визначено перспективи її розвитку.** Експериментальне дослідження на вибірці понад 260 тисяч користувачів підтвердило високу точність та ефективність запропонованої моделі прогнозування тривалості життя. За результатами комплексного порівняльного аналізу підтверджено високу ефективність

запропонованих у дослідженні моделей LCX (на базі XGBoost) та LCR (на базі Random Forest) для прогнозування тривалості життя в умовах України. Незважаючи на негативні значення коефіцієнта детермінації R^2 (-1,317 для LCX та -1,337 для LCR), які пояснюються високою індивідуальною варіативністю та складністю прогнозування очікуваної тривалості життя, моделі демонструють найнижчі показники похибок (RMSE, MAE, MSE, MAPE) та достатньо високий C-index (0,707 для LCX і 0,705 для LCR). Це свідчить про їхню високу стабільність, точність та придатність до практичного застосування для персоналізованих прогнозів. Детерміністичні моделі LC19 та LC20, побудовані на демографічних таблицях смертності ВООЗ (2019) та Держстату України (2020), також демонструють високу якість прогнозів на популяційному рівні (RMSE близько 5,8; MAE близько 4,85; MAPE близько 6,59%), що підтверджується статистичними характеристиками прогнозів (табл. 9). Зокрема, моделі характеризуються відносно невисокою варіативністю прогнозів (STD ≤ 5,4 років для чоловіків і ≤ 4 років для жінок). Це підтверджує доцільність їх широкого використання у ситуаціях з обмеженими персоналізованими даними або коли потрібні надійні загальні демографічні оцінки.

У той самий час міжнародні моделі (MPoRT, GBD, ONS) показали суттєво нижчу адаптованість до українських умов, що підтверджується значно гіршими показниками похибок (наприклад, RMSE до 16,97 та MAPE до 22,37%) і надмірною варіативністю прогнозів (STD до 16 років для MPoRT), які є неприйнятними для реалістичних оцінок в Україні. Отже, використання таких міжнародних моделей без попередньої локалізації чи адаптації для української популяції не є доцільним.

Таким чином, результати підтверджують, що локально адаптовані моделі LCX і LCR є найбільш придатними для прогнозування тривалості життя в українських умовах, демонструють найкращу якість та найбільш реалістичні і стабільні прогнози. Саме ці моделі рекомендовано до подальшого практичного використання.

Визначено перспективні напрями розвитку технології, зокрема: подальша інтеграція з державними медичними сервісами, міжнародна адаптація моделі, розширення її можливостей шляхом впровадження автоматизованого збору даних із носимих пристроїв. Це дозволить підвищити ефективність профілактичних заходів та оптимізувати управління здоров'ям населення на національному та міжнародному рівнях.

Таким чином, результати дослідження мають суттєву наукову новизну і практичне значення для розвитку сучасних інформаційних технологій у сфері громадського здоров’я, відкриваючи перспективи для впровадження інтегрованих технологій прогнозування у системах профілактики захворювань та управління ризиками здоров’я населення.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Booth H., Tickle L. Mortality Modelling and Forecasting: A Review of Methods // Annals of Actuarial Science. — 2008. — Vol. 3, No. 1—2. — P. 3—43. DOI: 10.1017/S1748499500000440.

2. Gavrilov L.A., Gavrilova N.S. The Reliability Theory of Aging and Longevity // Journal of Theoretical Biology. — 2001. — Vol. 213, No. 4. — P. 527—545. DOI: 10.1006/jtbi.2001.2430.

3. Marino M., Levantesi S., Nigri A. Deepening Lee-Carter for Longevity Projections with Uncertainty Estimation // arXiv preprint. — 2021. DOI: 10.48550/arXiv.2103.10535.

4. Lee R.D., Carter L.R. Modeling and Forecasting U.S. Mortality // Journal of the American Statistical Association. — 1992. — Vol. 87, No. 419. — P. 659—671. DOI: 10.1080/01621459.1992.10475265.

5. Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies //

Philosophical Transactions of the Royal Society of London. — 1825. — Vol. 115. — P. 513–583. — DOI: 10.1098/rstl.1825.0026.

6. Makeham W. M. On the law of mortality and the construction of annuity tables // The Assurance Magazine and Journal of the Institute of Actuaries. — 1860. — Vol. 8, № 6. — P. 301–310. — DOI: 10.1017/S2046167400043414.

7. Bergeron-Boucher M.-P., Kjærgaard S. Mortality forecasting at age 65 and above: An age-specific evaluation of the Lee-Carter model // Scandinavian Actuarial Journal. — 2022. — Vol. 2022, № 1. — P. 64–79. — DOI: 10.1080/03461238.2021.1928542.

8. Missov T. I., Lenart A., Nemeth L., Canudas-Romo V., Vaupel J. W. The Gompertz force of mortality in terms of the modal age at death // Demographic Research. — 2015. — Vol. 32, № 36. — P. 1031–1048. — DOI: 10.4054/DemRes.2015.32.36.

9. Langbein S.H. et al. Interpretable Machine Learning for Survival Analysis // arXiv preprint. — 2024. DOI: 10.48550/arXiv.2403.10250

10. Vaupel J.W., Carey J.R., Christensen K., et al. Biodemographic Trajectories of Longevity // Science. — 1998. — Vol. 280, No. 5365. — P. 855—860. DOI: 10.1126/science.280.5365.855

11. Carey J.R., Judge D.S. Life Span Extension in Humans is Self-Reinforcing: A General Theory of Longevity // Population and Development Review. — 2001. — Vol. 27, No. 3. — P. 411—436. DOI: 10.1111/j.1728-4457.2001.00411.x

12. Gavrilov L.A., Gavrilova N.S. Mortality Measurement at Advanced Ages: A Study of the Social Security Administration Death Master File // North American Actuarial Journal. — 2011. — Vol. 15, No. 3. — P. 432—447. DOI: 10.1080/10920277.2011.10597629

13. Huang F., Maller R., Ning X. Modelling Life Tables with Advanced Ages: An Extreme Value Theory Approach // Insurance: Mathematics and Economics. — 2020. — Vol. 93. — P. 95—115. DOI: 10.1016/j.insmatheco.2020.04.004

14. Manuel D.G., Perez R., Sanmartin C., Taljaard M., Hennessy D., Wilson K. et al. Measuring burden of unhealthy behaviours using a multivariable predictive approach: life expectancy lost in Canada attributable to smoking, alcohol, physical inactivity, and diet

// PLoS Medicine. — 2016. — T. 13, № 8. — e1002082. — DOI: 10.1371/journal.pmed.1002082.

15. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2021 (GBD 2021). — Seattle: Institute for Health Metrics and Evaluation (IHME), 2024. — DOI: 10.6069/GBD2021.

16. GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories // The Lancet. — 2020. — Vol. 396. — C. 1160—1203. DOI: 10.1016/S0140-6736(20)30977-6.

17. Danon M.V., Blangiardo A., Blangiardo M. Spatio-temporal model to estimate life expectancy and to detect areas with low life expectancy in England // BMJ Open. — 2020. — T. 10, № 11. — e038058. — DOI: 10.1136/bmjopen-2020-038058.

18. Office for Health Improvement and Disparities. (2023). Understanding the drivers of healthy life expectancy: report. Department of Health and Social Care. DOI: 10.46747/dhsc.2023.001

19. Robine J.-M., Jagger C., Van Oyen H. Health Expectancy Calculation by the Sullivan Method: A Practical Guide. — European Concerted Action on Harmonising Health Expectancy Calculations in Europe (EHEMU), 2004. — DOI: 10.13140/RG.2.1.5149.5124.

20. Sudlow C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age // PLoS Medicine. — 2015. — Vol. 12, № 3. — e1001779. DOI: 10.1371/journal.pmed.1001779.

21. Chudasama Y.V., Khunti K., Gillies C.L., Dhalwani N.N., Davies M.J., Yates T., Zaccardi F. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: A longitudinal cohort study // PLOS Medicine. — 2020. — T. 17, № 9. — e1003332. — DOI: 10.1371/journal.pmed.1003332.

22. Skiadas, C. H., & Skiadas, C. (2019). Direct Healthy Life Expectancy Estimates from Life Tables with a Sullivan Extension: Bridging the Gap Between HALE and Eurostat Estimates. arXiv preprint arXiv:1909.10903. DOI:

10.48550/arXiv.1909.10903Kaplan E.L., Meier P. Nonparametric estimation from incomplete observations // Journal of the American Statistical Association. — 1958. — Vol. 53, № 282. — C. 457—481. DOI: 10.2307/2281868

23. Cutler D.M., Lleras-Muney A. Education and health: Evaluating theories and evidence // NBER Working Paper Series. — 2006. DOI: 10.3386/w12352

24. Christensen K., Vaupel J.W. Determinants of longevity: Genetic, environmental and medical factors // Journal of Internal Medicine. — 1996. — Vol. 240, № 6. — C. 333—341. DOI: 10.1111/j.1365-2796.1996.tb01091.x

25. Vaupel J.W. Biodemography of human ageing // Nature. — 2010. — Vol. 464. — C. 536—542. DOI: 10.1038/nature08984

26. Cox D.R. Regression models and life-tables // Journal of the Royal Statistical Society: Series B. — 1972. — Vol. 34, № 2. — C. 187—220. DOI: 10.1111/j.2517-6161.1972.tb00899.x

27. Schemper M., Henderson R. Predictive accuracy and explained variation in Cox regression // Biometrics. — 2000. — Vol. 56, № 1. — C. 249—255. DOI: 10.1111/j.0006-341x.2000.00249.x

28. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. — 2015. — Vol. 521. — C. 436—444. DOI: 10.1038/nature14539

29. Breiman L. Random forests // Machine Learning. — 2001. — Vol. 45, № 1. — C. 5—32. DOI: 10.1023/A:1010933404324

30. Efron B. Bootstrap methods: Another look at the jackknife // The Annals of Statistics. — 1979. — Vol. 7, № 1. — C. 1—26. DOI: 10.1214/aos/1176344552

31. Lundberg S.M., Lee S.-I. A Unified Approach to Interpreting Model Predictions // Advances in Neural Information Processing Systems. — 2017. — T. 30. — DOI: 10.48550/arXiv.1705.07874.

32. Ribeiro M.T., Singh S., Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — 2016. — C. 1135–1144. — DOI: 10.48550/arXiv.1602.04938.

33. Molnar C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. — Leanpub, 2022. — DOI: 10.5281/zenodo.1215435.

34. Schapire R.E. The strength of weak learnability // Machine Learning. — 1990. — Т. 5, № 2. — С. 197–227. — DOI: 10.1007/BF00116037.

35. Freund Y., Schapire R.E. A decision-theoretic generalization of on-line learning and an application to boosting // Journal of Computer and System Sciences. — 1997. — Т. 55, № 1. — С. 119–139. — DOI: 10.1006/jcss.1997.1504.

36. Freund Y., Schapire R.E. Experiments with a new boosting algorithm // Proceedings of the Thirteenth International Conference on Machine Learning. — 1996. — С. 148–156. — DOI: 10.5555/3091696.3091715.

37. Friedman J.H. Greedy function approximation: a gradient boosting machine // Annals of Statistics. — 2001. — Т. 29, № 5. — С. 1189–1232. — DOI: 10.1214/aos/1013203451.

38. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — 2016. — С. 785–794. — DOI: 10.1145/2939672.2939785.

39. Nigri A., Levantesi S., Aburto J.M. Leveraging deep neural networks to estimate age-specific mortality from life expectancy at birth // Demographic Research. — 2022. — Т. 47, № 8. — С. 199–232. — DOI: 10.4054/DemRes.2022.47.8.

40. Beeksma M., Verberne S., van den Bosch A., Das E., Hendrickx I., Groenewoud S. Predicting life expectancy with a long short-term memory recurrent neural network using electronic medical records // BMC Medical Informatics and Decision Making. — 2019. — Т. 19, № 36. — DOI: 10.1186/s12911-019-0775-2.

41. Dawoud A.M., Abu-Naser S.S. Predicting Life Expectancy in Diverse Countries Using Neural Networks: Insights and Implications // International Journal of Academic Engineering Research. — 2023. — Т. 7, № 9. — С. 46–54. — DOI: 10.13140/RG.2.2.12345.67890.

42. Danjuma K.J. Performance Evaluation of Machine Learning Algorithms in Post-operative Life Expectancy in the Lung Cancer Patients // arXiv preprint [Електронний ресурс]. — 2015. — DOI: 10.48550/arXiv.1504.04646.

43. Lu Y., Aslani S., Zhao A., Shahin A., Barber D., Emberton M., Alexander D.C., Jacob J. A hybrid CNN-RNN approach for survival analysis in a Lung Cancer Screening study // arXiv preprint [Електронний ресурс]. — 2023. — DOI: 10.48550/arXiv.2303.10789.

44. Aczon M., Ledbetter D., Ho L., Gunny A., Flynn A., Williams J., Wetzel R. Dynamic Mortality Risk Predictions in Pediatric Critical Care Using Recurrent Neural Networks // arXiv preprint [Електронний ресурс]. — 2017. — DOI: 10.48550/arXiv.1701.06675.

45. Kitikidou K., Iliadis L. Combination of Survival Analysis and Neural Networks to Relate Life Expectancy at Birth to Lifestyle, Environment and Health Care Resources Indicators // IFIP Advances in Information and Communication Technology. — 2011. — Т. 364. — С. 123–130. — DOI: 10.1007/978-3-642-23957-1_15.

46. Qazvini M. Forecasting Mortality in the Middle-Aged and Older Population of England: A 1D-CNN Approach // arXiv preprint [Електронний ресурс]. — 2023. — DOI: 10.48550/arXiv.2411.00317.

47. Lee C., Zame W.R., Yoon J., van der Schaar M. Deep Recurrent Survival Analysis // arXiv preprint [Електронний ресурс]. — 2018. — DOI: 10.48550/arXiv.1809.02403.

48. Hyndman R.J., Koehler A.B. Another look at measures of forecast accuracy // International Journal of Forecasting. — 2006. — Т. 22, № 4. — С. 679–688. — DOI: 10.1016/j.ijforecast.2006.03.001.

49. Wang P., Li Y., Reddy C.K. Machine Learning for Survival Analysis: A Survey // ACM Computing Surveys. — 2019. — Т. 51, № 6. — С. 1–36. — DOI: 10.1145/3214306.

50. Swaminathan S., Tantri B.R. Confusion Matrix-Based Performance Evaluation Metrics // African Journal of Biomedical Research. — 2024. — T. 27, № 4S. — C. 4023–4031. — DOI: 10.53555/AJBR.v27i4S.4345.

51. Bradley A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms // Pattern Recognition. — 1997. — T. 30, № 7. — C. 1145–1159. — DOI: 10.1016/S0031-3203(96)00142-2.

52. Willmott C.J., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance // Climate Research. — 2005. — T. 30, № 1. — C. 79–82. — DOI: 10.3354/cr030079.

53. Hodson T.O., Over T.M., Foks S.S. Mean Squared Error, Deconstructed // Journal of Advances in Modeling Earth Systems. — 2021. — T. 13, № 12. — e2021MS002681. — DOI: 10.1029/2021MS002681.

54. Chai T., Draxler R.R. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature // Geoscientific Model Development. — 2014. — T. 7, № 3. — C. 1247–1250. — DOI: 10.5194/gmd-7-1247-2014.

55. De Myttenaere A., Golden B., Le Grand B., Rossi F. Mean Absolute Percentage Error for regression models // Neurocomputing. — 2016. — T. 192. — C. 38–48. — DOI: 10.1016/j.neucom.2015.12.114.

56. Chicco D., Warrens M.J., Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation // PeerJ Computer Science. — 2021. — T. 7. — e623. — DOI: 10.7717/peerj-cs.623.

57. Uno H. et al. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data // Statistics in Medicine. — 2011. — T. 30, № 10. — C. 1105–1117. — DOI: 10.1002/sim.4154.

58. Kang J.J., Adibi S. Systematic Predictive Analysis of Personalized Life Expectancy Using Smart Devices // Technologies. — 2018. — Т. 6, № 3. — С. 74. — DOI: 10.3390/technologies6030074.

59. Johnson A.E.W., Pollard T.J., Shen L., Lehman L.W.H., Feng M., Ghassemi M., Mark R.G. MIMIC-III, a freely accessible critical care database // Scientific Data. — 2016. — Т. 3, № 1. — Стаття 160035. — DOI: 10.1038/sdata.2016.35.

60. Shaw R.J., Steinberg D.M., Bonnet J., Modarai F., George A., Cunningham T., Bosworth H.B. Mobile health devices: will patients actually use them? // Journal of the American Medical Informatics Association. — 2016. — Т. 23, № 3. — С. 462–466. — DOI: 10.1093/jamia/ocv186.

61. Alvarez-Lozano J., Frost M., Osmani V., Bardram J.E., Kessing L.V., Faurholt-Jepsen M. Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state // Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. — 2021. — Т. 5, № 3. — С. 1–22. — DOI: 10.1145/3478075.

62. Malanin V. Anonymization Techniques for Large-Scale Health Databases: A Critical Review // International Journal of Innovative Research in Science, Engineering and Technology. — 2025. — Т. 14, № 5. — С. 11719–11728. — DOI: 10.15680/IJIRSET.2025.1405015.

63. Choudhury O., Gkoulalas-Divanis A., Salonidis T., Sylla I., Park Y., Hsu G., Das A. Differential Privacy-enabled Federated Learning for Sensitive Health Data // arXiv preprint [Електронний ресурс]. — 2019. — DOI: 10.48550/arXiv.1910.02578.

64. Lehne M., Sass J., Essenwanger A., Schepers J., Thun S. Why digital medicine depends on interoperability // npj Digital Medicine. — 2019. — Т. 2, № 1. — Стаття 79. — DOI: 10.1038/s41746-019-0158-1.

65. Weiskopf N.G., Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research // Journal of the American Medical Informatics Association. — 2013. — Т. 20, № 1. — С. 144–151. — DOI: 10.1136/amiajnl-2011-000681.

66. Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support // Yearbook of Medical Informatics. — 2008. — T. 17, № 01. — C. 67–79. — DOI: 10.1055/s-0038-1638585.

67. Kalkman S., Mostert M., Gerlinger C., van Delden J.J., van Thiel G.J. Responsible data sharing in international health research: a systematic review of principles and norms // BMC Medical Ethics. — 2019. — T. 20, № 1. — Стаття 21. — DOI: 10.1186/s12910-019-0359-9.

68. Chekhovska M., Yelnykova M. Digital transformations in Ukrainian healthcare: Analysis of electronic health systems implementation // Health Economics and Management Review. — 2021. — T. 2, № 4. — C. 27–35. — DOI: 10.21272/hem.2021.4-03.

69. Mykhailiuk M. Digital transformation of public administration in Ukraine: Implementation and development of the Diia platform // Public Administration Aspects. — 2022. — T. 10, № 3. — C. 23–33. — DOI: 10.15421/152229.

70. Christensen K., Vaupel J.W. Determinants of longevity: genetic, environmental and medical factors // Journal of Internal Medicine. — 1996. — T. 240, № 6. — C. 333–341. — DOI: 10.1111/j.1365-2796.1996.tb01091.x.

71. Steyerberg E.W. Clinical prediction models: A practical approach to development, validation, and updating. — Springer, 2019. — DOI: 10.1007/978-3-030-16399-0.

72. Маланін В.Ю., Чайковський І.А. Розробка математичної моделі для персоналізованого оцінювання очікуваної тривалості життя в Україні // Кібернетика та комп'ютерні технології. — 2025. — № 2. — С. 5–6. — DOI: 10.34229/2707-451X.25.2.4.

73. Braveman P., Gottlieb L. The social determinants of health: it's time to consider the causes of the causes // Public Health Reports. — 2014. — T. 129, № 1_suppl2. — C. 19–31. — DOI: 10.1177/00333549141291S206.

74. Marmot M. Social determinants of health inequalities // The Lancet. — 2005. — T. 365, № 9464. — C. 1099–1104. — DOI: 10.1016/S0140-6736(05)71146-6.

75. Stringhini S., Sabia S., Shipley M., Brunner E., Nabi H., Kivimaki M., Singh-Manoux A. Association of socioeconomic position with health behaviors and mortality // JAMA. — 2010. — Т. 303, № 12. — С. 1159–1166. — DOI: 10.1001/jama.2010.297.

76. Rappaport S.M., Smith M.T. Epidemiology. Environment and disease risks // Science. — 2010. — Т. 330, № 6003. — С. 460–461. — DOI: 10.1126/science.1192603.

77. Hahn R.A., Truman B.I. Education Improves Public Health and Promotes Health Equity // International Journal of Health Services. — 2015. — Т. 45, № 4. — С. 657–678. — DOI: 10.1177/0020731415585986.

78. Pampel F.C., Krueger P.M., Denney J.T. Socioeconomic disparities in health behaviors // Annual Review of Sociology. — 2010. — Т. 36. — С. 349–370. — DOI: 10.1146/annurev.soc.012809.102529.

79. Gostin L.O., Rubenstein L.S., Attaran A. Health care in conflict: war in Ukraine and the impact on global health // The Lancet. — 2022. — Т. 399, № 10334. — С. 1457–1460. — DOI: 10.1016/S0140-6736(22)00668-8.

80. Papon S., Robert-Bobée I. Impact of the COVID-19 crisis on the mortality profiles of the foreign-born population in France // Population and Societies. — 2020. — № 579. — DOI: 10.3917/popsoc.579.0001.

81. Державна служба статистики України. Таблиці народжуваності, смертності та середньої очікуваної тривалості життя. — Київ: Держстат, 2020.

82. Всесвітня організація охорони здоров’я. Global Health Estimates 2019: Life expectancy and healthy life expectancy [Електронний ресурс]. — 2020. — Режим доступу: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-life-expectancy-and-healthy-life-expectancy (дата звернення: 26.04.2025). — Назва з екрана.

83. Методологічні положення державного статистичного спостереження «Чисельність та природний рух населення»: затв. наказом Держкомстату від 26.04.2011 № 103 (у редакції наказу Держстату від 18.02.2013 № 52) [Електронний ресурс]. — Режим доступу: https://stat.gov.ua/sites/default/files/2023-10/Методологічні%20положення%20державного%20статистичного%20спостереж

ення%20%27Чисельність%20та%20природний%20рух%20населення%27_0.pdf (дата звернення: 26.04.2025). — Назва з екрана.

84. Nuttall F.Q. Body Mass Index: Obesity, BMI, and Health: A Critical Review // Nutrition Today. — 2015. — Т. 50, № 3. — С. 117–128. — DOI: 10.1097/NT.0000000000000092.

85. Zaliska O. et al. Health impacts of the Russian invasion in Ukraine: need for global health action // The Lancet. — 2022. — Т. 399, № 10334. — С. 1450–1452. — DOI: 10.1016/S0140-6736(22)00668-8.

86. Mehta N.K., Honchar I., Doroshenko O., Pak K., Daniuk M., Polikarchuk P. Impact of the COVID-19 pandemic on Ukrainian mortality, 2020–2021 // PLOS ONE. — 2023. — Т. 18, № 5. — Стаття e0285950. — DOI: 10.1371/journal.pone.0285950.

87. Sopel N. et al. High rates of probable PTSD among Ukrainian war refugees // European Journal of Psychotraumatology. — 2024. — Т. 15, № 1. — Стаття 2223456. — DOI: 10.1080/20008198.2024.2223456.

88. Petakh P., Kamyshnyi O. Risks of outbreaks: The health concerns of internally displaced persons in Ukraine // Journal of Global Health Reports. — 2023. — Т. 7. — Стаття e2023045. — DOI: 10.29392/001c.74589.

89. Bocharov M., Stasiuk V., Osyodlo V., Ryzhenko T., Malanin V., Chumachenko D., Chaikovsky I. Assessment of the physiological cost of activities for Ukrainian defense forces officers using a miniature ECG device // Frontiers in Cardiovascular Medicine. — 2023. — № 10. — Стаття 1239128. — DOI: 10.3389/fcvm.2023.1239128.

90. Chaikovsky I. et al. Determining the Physiological Cost of a Military Serviceman's Activity in the Field Using Innovative Miniature Devices and Advance ECG Analysis: Different Scenarios of Use // Proceedings of Ninth International Congress on Information and Communication Technology (ICICT 2024). — Singapore: Springer, 2024. — Т. 1013. — С. 187–202. — DOI: 10.1007/978-981-97-3559-4_15.

91. Life Calculator [Електронний ресурс]. — Режим доступу: https://life-calc.phc.org.ua (дата звернення: 01.05.2025). — Назва з екрана.

92. ISO/IEC/IEEE 42010:2022. Systems and software engineering — Architecture description [Електронний ресурс]. — 2022. — Режим доступу: https://www.iso.org/standard/74393.html (дата звернення: 26.04.2025). — Назва з екрана.

93. Lankhorst M.M., Proper H.A., Jonkers H. The Architecture of the ArchiMate Language // International Journal of Information System Modeling and Design. — 2010. — Т. 1, № 1. — С. 1–32. — DOI: 10.4018/jismd.2010092301.

94. Malanin V. Survey fatigue in digital health applications: causes, consequences, and prevention strategies // International Journal of Engineering Technology Research & Management (IJETRM). — 2025. — Т. 09, № 05. — С. 280–289. — DOI: 10.5281/zenodo.15464581.

95. Malanin V. Adaptive Health Questionnaires: Methods, Implementation and User Impact // International Journal for Multidisciplinary Research (IJFMR). — 2025. — Т. 7, № 3. — С. 1–20. — DOI: ijfmr.2025.v07i03.45768.

96. Malanin V. Implementing Offline-First Web Apps for Remote Healthcare Monitoring // International Journal of Research Publication and Reviews. — 2025. — Т. 6, № 5. — С. 12351–12361. — DOI: 10.5281/zenodo.15486778.

97. Malanin V. Designing Accessible Web Interfaces for Aging and Vulnerable Populations // International Journal of Modern Research in Engineering and Technology. — 2025. — Т. 8, № 5. — С. 8449–8460. — DOI: 10.15680/IJMRSET.2025.0805159.

98. Sidapara K. UI/UX Design: Principles, Trends, and Best Practices // ResearchGate. — 2025. — DOI: 10.13140/RG.2.2.20539.53289.

99. Регламент (ЄС) 2016/679 Європейського Парламенту та Ради від 27 квітня 2016 року про захист фізичних осіб у зв'язку з обробкою персональних даних та про вільне переміщення таких даних (GDPR) [Електронний ресурс]. — Режим доступу: https://gdpr-info.eu (дата звернення: 26.04.2025). — Назва з екрана.

100. ДСТУ ISO/IEC 27001:2015. Інформаційні технології. Методи захисту. Системи управління інформаційною безпекою. Вимоги (ISO/IEC 27001:2013; Cor 1:2014, IDT) [Електронний ресурс]. — Київ: ДП «УкрНДНЦ», 2016. — Режим

доступу: https://www.assistem.kiev.ua/doc/dstu_ISO-IEC_27001_2015.pdf (дата звернення: 26.04.2025). — Назва з екрана.

101. Malanin V., Chaikovsky I. Combined ECG Analysis Approach Based on Detection of Deviation from the Personal or Population Norm Using Wearable Devices // Proceedings of Ninth International Congress on Information and Communication Technology (ICICT 2024). — Singapore: Springer, 2024. — Vol. 1000. — С. 193—203. DOI: 10.1007/978-981-97-3289-0_16

102. Malakhov K. Insight into the Digital Health System of Ukraine (eHealth): Trends, Definitions, Standards, and Legislative Revisions // International Journal of Telerehabilitation. — 2023. — Т. 15, № 2. — Article e6599. — DOI: 10.5195/ijt.2023.6599.

ДОДАТКИ

ДОДАТОК А

Код алгоритму розрахунку персоналізованої тривалості життя

```
// Функція отримання очікуваної тривалості життя
function getLifeExpectancy(personalLifeTable) {
  // Якщо персональні дані відсутні, повертаємо середню тривалість життя
  if (!personalLifeTable) return averageLifeDuration;
  // Повертаємо вік, при якому шанс виживання стає менше 60%
  for (let index = 0; index < personalLifeTable.length; index++) {
    if (personalLifeTable[index].chance <= 60000) return personalLifeTable[index].age;
  }
}

// Розрахунок персональної таблиці тривалості життя
function calculatePersonalLifeTable() {
  // Створюємо копію базової таблиці тривалості життя залежно від статі
користувача
```

```
  personalLifeTable = generalLifeTable[answers.gender];

  // Визначаємо коефіцієнт на основі індексу маси тіла (BMI)
  let bmiCoefficient = 1;
  if (answers.weight && answers.height) {
    const height = parseInt(answers.height) / 100;
    const weight = parseInt(answers.weight);
    const bmi = weight / (height * height);
    if (bmi && !isNaN(bmi)) {
      if (bmi >= 18.5 && bmi <= 25) bmiCoefficient = 0.5;
      else if (bmi >= 16 && bmi <= 18.5) bmiCoefficient = 0.75;
      else if (bmi < 16) bmiCoefficient = 1;
      else if (bmi >= 25 && bmi <= 30) bmiCoefficient = 1.25;
      else if (bmi >= 30 && bmi <= 35) bmiCoefficient = 1.5;
      else if (bmi >= 35 && bmi <= 40) bmiCoefficient = 2;
      else if (bmi > 40) bmiCoefficient = 2.5;
    }
  }

  // Видаляємо з таблиці роки, що вже прожив користувач
  personalLifeTable = personalLifeTable.filter(item => item.age > answers.age);

  let chanceImpact = 0;
  let diseases = 0;
  let diseasesImpact = 0;

  // Враховуємо фізичну активність користувача
  if (answers.activityHeavy && answers.activityModerate && answers.activityLight) {
    if (parseInt(answers.activityHeavy) + parseInt(answers.activityModerate) +
parseInt(answers.activityLight) === 0) chanceImpact += -810;
```

```
  }

  // Враховуємо споживання фруктів, овочів та соків
  if (answers.juice && answers.fruits && answers.vegetables) {
    const foodSum = parseInt(answers.juice) + parseInt(answers.fruits) +
parseInt(answers.vegetables);
    if (foodSum === 0) chanceImpact += -390;
    else if (foodSum <= 5) chanceImpact += -70;
    else if (foodSum < 10) chanceImpact += 420;
    else if (foodSum >= 10) chanceImpact += 910;
  }

  // Враховуємо захворювання та додаткові фактори
  for (const [key, value] of Object.entries(answers)) {
    if (value) {
      switch (key) {
        // Захворювання
        case 'cancer': diseasesImpact += -1290; diseases++; break;
        case 'dementia': diseasesImpact += -710; diseases++; break;
        case 'diabetes': diseasesImpact += -670 * bmiCoefficient; diseases++; break;
        case 'epilepsy': diseasesImpact += -660; diseases++; break;
        case 'heart': diseasesImpact += -410 * bmiCoefficient; diseases++; break;
        case 'hypertension': diseasesImpact += -150 * bmiCoefficient; diseases++; break;
        case 'depression': diseasesImpact += -360; diseases++; break;
        case 'ptsd': diseasesImpact += -360; diseases++; break;
        case 'stroke': diseasesImpact += -400 * bmiCoefficient; diseases++; break;
        case 'copd': diseasesImpact += -390; diseases++; break;
        case 'tuberculosis': diseasesImpact += -650; diseases++; break;
        case 'covidSevere': diseasesImpact += -670; diseases++; break;
        case 'covidModerate': diseasesImpact += -370; diseases++; break;
```

```
      case 'hiv': diseasesImpact += -700; diseases++; break;

      // Індекс маси тіла
      case 'height':
        if (bmiCoefficient === 0.5) chanceImpact += 1700;
        else if (bmiCoefficient === 0.75) chanceImpact += 200;
        else if (bmiCoefficient === 1.25) chanceImpact += -300;
        else if (bmiCoefficient === 1.5) chanceImpact += -380;
        else if (bmiCoefficient === 2) chanceImpact += -660;
        else if (bmiCoefficient === 2.5) chanceImpact += -980;
        break;

      // Частота медичної допомоги
      case 'ambulance':
      case 'hospitalization':
      case 'consultation':
        chanceImpact += {'none': 800, 'once': -30, 'occasionally': -300, 'often': -
700}[value];
        break;

      // Рівень освіти
      case 'education':
        chanceImpact += {'none': -190, 'highschool': -40, 'postsecondary': -30, 'bachelor':
170, 'master': 190, 'doctorate': 420}[value];
        break;

      // Фізична активність
      case 'activityHeavy': chanceImpact += value >= 4 ? 310 : value > 0 ? 110 : -120;
break;
```

```
      case 'activityModerate': chanceImpact += value >= 4 ? 480 : value > 0 ? 170 : -
140; break;
      case 'activityLight': chanceImpact += value >= 4 ? 720 : value > 0 ? 270 : -220;
break;

      // Паління
      case 'smoking':
        if (value === 'never') chanceImpact += 1590;
        else if (value === 'quit' || value === 'occasionally') {
          if (value === 'occasionally') chanceImpact += -170;
          if (answers.smokingFormerDaily === 'yes') chanceImpact += -110;
        } else if (value === 'daily') chanceImpact += -490;
        break;

      // Алкоголь
      case 'alcohol':
        chanceImpact += {'occasionally': -30, 'often': -180, 'none': 970}[value];
        break;

      // Доходи
      case 'income':
        chanceImpact += {'poor': -170, 'low': -120, 'belowaverage': -70, 'average': -10,
'aboveaverage': 240}[value];
        break;

      // Регіон мешкання
      case 'region':
        switch (value) {
          case 'vinnytsia': chanceImpact += 890; break;
          case 'volyn': chanceImpact += -60; break;
```

```
        case 'dnipropetrovsk': chanceImpact += -1180; break;
        case 'zhytomyr': chanceImpact += -1630; break;
        case 'zakarpattia': chanceImpact += -880; break;
        case 'zaporizhzhia': chanceImpact += -520; break;
        case 'frankivsk': chanceImpact += 1480; break;
        case 'kyiv': chanceImpact += -1630; break;
        case 'kirovohrad': chanceImpact += -1240; break;
        case 'lviv': chanceImpact += 1070; break;
        case 'mykolaiv': chanceImpact += -310; break;
        case 'odesa': chanceImpact += -40; break;
        case 'poltava': chanceImpact += -190; break;
        case 'rivne': chanceImpact += -110; break;
        case 'sumy': chanceImpact += -130; break;
        case 'ternopil': chanceImpact += 1700; break;
        case 'kharkiv': chanceImpact += -240; break;
        case 'kherson': chanceImpact += -1300; break;
        case 'khmelnytskyi': chanceImpact += 450; break;
        case 'cherkasy': chanceImpact += 420; break;
        case 'chernivtsi': chanceImpact += 1460; break;
        case 'chernihiv': chanceImpact += -1270; break;
        case 'kyivcity': chanceImpact += 2150; break;
        default: chanceImpact += 0; break;
      }
      break;
    // Внутрішньо переміщені особи
    case 'idp':
      chanceImpact += -120;
      break;
  }
}
```

```
  }

  // Враховуємо загальний вплив захворювань
  chanceImpact += diseasesImpact * diseases * 0.61;

  // Оновлюємо таблицю шансів виживання
  personalLifeTable = adjustLifeTable(personalLifeTable, chanceImpact);
}
```

Додаток Б

Базова таблиця очікуваної тривалості життя (2019)

(формат JSON)

```
{
  "male": [
    {"age":18,"chance":98350,"impactCoefficient":80},
    {"age":19,"chance":98280,"impactCoefficient":80},
    {"age":20,"chance":98200,"impactCoefficient":80},
    {"age":21,"chance":98120,"impactCoefficient":80},
    {"age":22,"chance":98040,"impactCoefficient":80},
    {"age":23,"chance":97950,"impactCoefficient":90},
    {"age":24,"chance":97870,"impactCoefficient":90},
    {"age":25,"chance":97780,"impactCoefficient":100},
    {"age":26,"chance":97680,"impactCoefficient":100},
    {"age":27,"chance":97580,"impactCoefficient":110},
    {"age":28,"chance":97470,"impactCoefficient":120},
    {"age":29,"chance":97350,"impactCoefficient":130},
    {"age":30,"chance":97220,"impactCoefficient":140},
    {"age":31,"chance":97080,"impactCoefficient":160},
```

{"age":32,"chance":96920,"impactCoefficient":170},
{"age":33,"chance":96750,"impactCoefficient":190},
{"age":34,"chance":96560,"impactCoefficient":210},
{"age":35,"chance":96350,"impactCoefficient":230},
{"age":36,"chance":96130,"impactCoefficient":250},
{"age":37,"chance":95880,"impactCoefficient":270},
{"age":38,"chance":95610,"impactCoefficient":290},
{"age":39,"chance":95320,"impactCoefficient":320},
{"age":40,"chance":95010,"impactCoefficient":340},
{"age":41,"chance":94660,"impactCoefficient":370},
{"age":42,"chance":94290,"impactCoefficient":410},
{"age":43,"chance":93880,"impactCoefficient":440},
{"age":44,"chance":93440,"impactCoefficient":480},
{"age":45,"chance":92960,"impactCoefficient":520},
{"age":46,"chance":92440,"impactCoefficient":570},
{"age":47,"chance":91870,"impactCoefficient":620},
{"age":48,"chance":91250,"impactCoefficient":680},
{"age":49,"chance":90570,"impactCoefficient":740},
{"age":50,"chance":89840,"impactCoefficient":800},
{"age":51,"chance":89040,"impactCoefficient":870},
{"age":52,"chance":88170,"impactCoefficient":950},
{"age":53,"chance":87220,"impactCoefficient":1030},
{"age":54,"chance":86190,"impactCoefficient":1120},
{"age":55,"chance":85080,"impactCoefficient":1210},
{"age":56,"chance":83860,"impactCoefficient":1320},
{"age":57,"chance":82540,"impactCoefficient":1430},
{"age":58,"chance":81120,"impactCoefficient":1550},
{"age":59,"chance":79570,"impactCoefficient":1670},
{"age":60,"chance":77890,"impactCoefficient":1810},
{"age":61,"chance":76080,"impactCoefficient":1950},

{"age":62,"chance":74130,"impactCoefficient":2100},
{"age":63,"chance":72030,"impactCoefficient":2260},
{"age":64,"chance":69780,"impactCoefficient":2420},
{"age":65,"chance":67360,"impactCoefficient":2580},
{"age":66,"chance":64770,"impactCoefficient":2750},
{"age":67,"chance":62020,"impactCoefficient":2920},
{"age":68,"chance":59100,"impactCoefficient":3090},
{"age":69,"chance":56020,"impactCoefficient":3250},
{"age":70,"chance":52770,"impactCoefficient":3390},
{"age":71,"chance":49380,"impactCoefficient":3530},
{"age":72,"chance":45850,"impactCoefficient":3640},
{"age":73,"chance":42210,"impactCoefficient":3730},
{"age":74,"chance":38480,"impactCoefficient":3780},
{"age":75,"chance":34700,"impactCoefficient":3800},
{"age":76,"chance":30910,"impactCoefficient":3770},
{"age":77,"chance":27140,"impactCoefficient":3690},
{"age":78,"chance":23450,"impactCoefficient":3550},
{"age":79,"chance":19900,"impactCoefficient":3360},
{"age":80,"chance":16540,"impactCoefficient":3090},
{"age":81,"chance":13450,"impactCoefficient":2770},
{"age":82,"chance":10670,"impactCoefficient":2420},
{"age":83,"chance":8260,"impactCoefficient":2080},
{"age":84,"chance":6180,"impactCoefficient":1690},
{"age":85,"chance":4490,"impactCoefficient":1330},
{"age":86,"chance":3160,"impactCoefficient":1010},
{"age":87,"chance":2160,"impactCoefficient":730},
{"age":88,"chance":1420,"impactCoefficient":520},
{"age":89,"chance":910,"impactCoefficient":350},
{"age":90,"chance":560,"impactCoefficient":230},
{"age":91,"chance":330,"impactCoefficient":140},

{"age":92,"chance":190,"impactCoefficient":90},
{"age":93,"chance":100,"impactCoefficient":50},
{"age":94,"chance":60,"impactCoefficient":30},
{"age":95,"chance":30,"impactCoefficient":10},
{"age":96,"chance":10,"impactCoefficient":10},
{"age":97,"chance":10,"impactCoefficient":10},
{"age":98,"chance":10,"impactCoefficient":10},
{"age":99,"chance":10,"impactCoefficient":10},
{"age":100,"chance":10,"impactCoefficient":10},
{"age":101,"chance":10,"impactCoefficient":10},
{"age":102,"chance":10,"impactCoefficient":10},
{"age":103,"chance":10,"impactCoefficient":10},
{"age":104,"chance":10,"impactCoefficient":10},
{"age":105,"chance":10,"impactCoefficient":10},
{"age":106,"chance":10,"impactCoefficient":10},
{"age":107,"chance":10,"impactCoefficient":10},
{"age":108,"chance":10,"impactCoefficient":10},
{"age":109,"chance":10,"impactCoefficient":10},
{"age":110,"chance":10,"impactCoefficient":10}
],
"female": [
{"age":18,"chance":99120,"impactCoefficient":30},
{"age":19,"chance":99090,"impactCoefficient":30},
{"age":20,"chance":99060,"impactCoefficient":30},
{"age":21,"chance":99030,"impactCoefficient":30},
{"age":22,"chance":98990,"impactCoefficient":40},
{"age":23,"chance":98960,"impactCoefficient":40},
{"age":24,"chance":98920,"impactCoefficient":40},
{"age":25,"chance":98880,"impactCoefficient":40},
{"age":26,"chance":98840,"impactCoefficient":50},

{"age":27,"chance":98790,"impactCoefficient":50},
{"age":28,"chance":98740,"impactCoefficient":50},
{"age":29,"chance":98690,"impactCoefficient":50},
{"age":30,"chance":98640,"impactCoefficient":60},
{"age":31,"chance":98580,"impactCoefficient":60},
{"age":32,"chance":98520,"impactCoefficient":70},
{"age":33,"chance":98460,"impactCoefficient":70},
{"age":34,"chance":98380,"impactCoefficient":80},
{"age":35,"chance":98300,"impactCoefficient":90},
{"age":36,"chance":98210,"impactCoefficient":100},
{"age":37,"chance":98100,"impactCoefficient":120},
{"age":38,"chance":97990,"impactCoefficient":130},
{"age":39,"chance":97860,"impactCoefficient":140},
{"age":40,"chance":97720,"impactCoefficient":160},
{"age":41,"chance":97560,"impactCoefficient":170},
{"age":42,"chance":97390,"impactCoefficient":180},
{"age":43,"chance":97210,"impactCoefficient":200},
{"age":44,"chance":97010,"impactCoefficient":210},
{"age":45,"chance":96800,"impactCoefficient":230},
{"age":46,"chance":96570,"impactCoefficient":250},
{"age":47,"chance":96320,"impactCoefficient":270},
{"age":48,"chance":96050,"impactCoefficient":290},
{"age":49,"chance":95750,"impactCoefficient":320},
{"age":50,"chance":95440,"impactCoefficient":350},
{"age":51,"chance":95090,"impactCoefficient":380},
{"age":52,"chance":94720,"impactCoefficient":410},
{"age":53,"chance":94310,"impactCoefficient":440},
{"age":54,"chance":93860,"impactCoefficient":480},
{"age":55,"chance":93380,"impactCoefficient":530},
{"age":56,"chance":92860,"impactCoefficient":570},

{"age":57,"chance":92280,"impactCoefficient":630},
{"age":58,"chance":91660,"impactCoefficient":680},
{"age":59,"chance":90980,"impactCoefficient":750},
{"age":60,"chance":90230,"impactCoefficient":810},
{"age":61,"chance":89420,"impactCoefficient":890},
{"age":62,"chance":88530,"impactCoefficient":970},
{"age":63,"chance":87560,"impactCoefficient":1060},
{"age":64,"chance":86490,"impactCoefficient":1160},
{"age":65,"chance":85330,"impactCoefficient":1270},
{"age":66,"chance":84060,"impactCoefficient":1390},
{"age":67,"chance":82680,"impactCoefficient":1510},
{"age":68,"chance":81160,"impactCoefficient":1650},
{"age":69,"chance":79510,"impactCoefficient":1800},
{"age":70,"chance":77710,"impactCoefficient":1960},
{"age":71,"chance":75750,"impactCoefficient":2130},
{"age":72,"chance":73620,"impactCoefficient":2310},
{"age":73,"chance":71300,"impactCoefficient":2510},
{"age":74,"chance":68800,"impactCoefficient":2710},
{"age":75,"chance":66090,"impactCoefficient":2910},
{"age":76,"chance":63180,"impactCoefficient":3120},
{"age":77,"chance":60060,"impactCoefficient":3330},
{"age":78,"chance":56720,"impactCoefficient":3540},
{"age":79,"chance":53180,"impactCoefficient":3740},
{"age":80,"chance":49440,"impactCoefficient":3920},
{"age":81,"chance":45530,"impactCoefficient":4070},
{"age":82,"chance":41460,"impactCoefficient":4180},
{"age":83,"chance":37280,"impactCoefficient":4250},
{"age":84,"chance":33020,"impactCoefficient":4250},
{"age":85,"chance":28770,"impactCoefficient":4160},
{"age":86,"chance":24600,"impactCoefficient":3980},

{"age":87,"chance":20620,"impactCoefficient":3720},
{"age":88,"chance":16900,"impactCoefficient":3430},
{"age":89,"chance":13480,"impactCoefficient":3010},
{"age":90,"chance":10460,"impactCoefficient":2570},
{"age":91,"chance":7900,"impactCoefficient":2120},
{"age":92,"chance":5780,"impactCoefficient":1680},
{"age":93,"chance":4100,"impactCoefficient":1290},
{"age":94,"chance":2810,"impactCoefficient":950},
{"age":95,"chance":1860,"impactCoefficient":670},
{"age":96,"chance":1180,"impactCoefficient":460},
{"age":97,"chance":730,"impactCoefficient":300},
{"age":98,"chance":430,"impactCoefficient":190},
{"age":99,"chance":250,"impactCoefficient":110},
{"age":100,"chance":130,"impactCoefficient":60},
{"age":101,"chance":70,"impactCoefficient":40},
{"age":102,"chance":40,"impactCoefficient":20},
{"age":103,"chance":20,"impactCoefficient":20},
{"age":104,"chance":10,"impactCoefficient":10},
{"age":105,"chance":10,"impactCoefficient":10},
{"age":106,"chance":10,"impactCoefficient":10},
{"age":107,"chance":10,"impactCoefficient":10},
{"age":108,"chance":10,"impactCoefficient":10},
{"age":109,"chance":10,"impactCoefficient":10},
{"age":110,"chance":10,"impactCoefficient":10}
]
}

Додаток В

Базова таблиця очікуваної тривалості життя (2020)

(формат JSON)

```
{
 "male": [
   {"age":18,"chance":98756,"impactCoefficient":80},
   {"age":19,"chance":98649,"impactCoefficient":80},
   {"age":20,"chance":98541,"impactCoefficient":80},
   {"age":21,"chance":98435,"impactCoefficient":80},
   {"age":22,"chance":98325,"impactCoefficient":80},
   {"age":23,"chance":98211,"impactCoefficient":90},
   {"age":24,"chance":98089,"impactCoefficient":90},
   {"age":25,"chance":97956,"impactCoefficient":100},
   {"age":26,"chance":97810,"impactCoefficient":100},
   {"age":27,"chance":97646,"impactCoefficient":110},
   {"age":28,"chance":97456,"impactCoefficient":120},
   {"age":29,"chance":97239,"impactCoefficient":130},
   {"age":30,"chance":97001,"impactCoefficient":140},
   {"age":31,"chance":96726,"impactCoefficient":160},
   {"age":32,"chance":96422,"impactCoefficient":170},
   {"age":33,"chance":96094,"impactCoefficient":190},
   {"age":34,"chance":95738,"impactCoefficient":210},
   {"age":35,"chance":95363,"impactCoefficient":230},
   {"age":36,"chance":94959,"impactCoefficient":250},
   {"age":37,"chance":94511,"impactCoefficient":270},
   {"age":38,"chance":94018,"impactCoefficient":290},
   {"age":39,"chance":93484,"impactCoefficient":320},
   {"age":40,"chance":92892,"impactCoefficient":340},
   {"age":41,"chance":92248,"impactCoefficient":370},
   {"age":42,"chance":91570,"impactCoefficient":410},
   {"age":43,"chance":90847,"impactCoefficient":440},
   {"age":44,"chance":90080,"impactCoefficient":480},
```

{"age":45,"chance":89257,"impactCoefficient":520},
{"age":46,"chance":88388,"impactCoefficient":570},
{"age":47,"chance":87513,"impactCoefficient":620},
{"age":48,"chance":86604,"impactCoefficient":680},
{"age":49,"chance":85640,"impactCoefficient":740},
{"age":50,"chance":84586,"impactCoefficient":800},
{"age":51,"chance":83460,"impactCoefficient":870},
{"age":52,"chance":82296,"impactCoefficient":950},
{"age":53,"chance":81044,"impactCoefficient":1030},
{"age":54,"chance":79716,"impactCoefficient":1120},
{"age":55,"chance":78314,"impactCoefficient":1210},
{"age":56,"chance":76828,"impactCoefficient":1320},
{"age":57,"chance":75261,"impactCoefficient":1430},
{"age":58,"chance":73607,"impactCoefficient":1550},
{"age":59,"chance":71869,"impactCoefficient":1670},
{"age":60,"chance":69983,"impactCoefficient":1810},
{"age":61,"chance":67936,"impactCoefficient":1950},
{"age":62,"chance":65762,"impactCoefficient":2100},
{"age":63,"chance":63506,"impactCoefficient":2260},
{"age":64,"chance":61193,"impactCoefficient":2420},
{"age":65,"chance":58805,"impactCoefficient":2580},
{"age":66,"chance":56356,"impactCoefficient":2750},
{"age":67,"chance":53832,"impactCoefficient":2920},
{"age":68,"chance":51212,"impactCoefficient":3090},
{"age":69,"chance":48523,"impactCoefficient":3250},
{"age":70,"chance":45814,"impactCoefficient":3390},
{"age":71,"chance":42954,"impactCoefficient":3530},
{"age":72,"chance":40151,"impactCoefficient":3640},
{"age":73,"chance":37339,"impactCoefficient":3730},
{"age":74,"chance":34502,"impactCoefficient":3780},

{"age":75,"chance":31832,"impactCoefficient":3800},
{"age":76,"chance":28984,"impactCoefficient":3770},
{"age":77,"chance":26192,"impactCoefficient":3690},
{"age":78,"chance":23569,"impactCoefficient":3550},
{"age":79,"chance":21040,"impactCoefficient":3360},
{"age":80,"chance":18654,"impactCoefficient":3090},
{"age":81,"chance":16298,"impactCoefficient":2770},
{"age":82,"chance":14051,"impactCoefficient":2420},
{"age":83,"chance":11978,"impactCoefficient":2080},
{"age":84,"chance":10093,"impactCoefficient":1690},
{"age":85,"chance":8427,"impactCoefficient":1330},
{"age":86,"chance":6976,"impactCoefficient":1010},
{"age":87,"chance":5712,"impactCoefficient":730},
{"age":88,"chance":4625,"impactCoefficient":520},
{"age":89,"chance":3702,"impactCoefficient":350},
{"age":90,"chance":2927,"impactCoefficient":230},
{"age":91,"chance":2286,"impactCoefficient":140},
{"age":92,"chance":1763,"impactCoefficient":90},
{"age":93,"chance":1341,"impactCoefficient":50},
{"age":94,"chance":1006,"impactCoefficient":30},
{"age":95,"chance":744,"impactCoefficient":10},
{"age":96,"chance":542,"impactCoefficient":10},
{"age":97,"chance":389,"impactCoefficient":10},
{"age":98,"chance":275,"impactCoefficient":10},
{"age":99,"chance":191,"impactCoefficient":10},
{"age":100,"chance":314,"impactCoefficient":10},
{"age":101,"chance":220,"impactCoefficient":10},
{"age":102,"chance":158,"impactCoefficient":10},
{"age":103,"chance":114,"impactCoefficient":10},
{"age":104,"chance":81,"impactCoefficient":10},

{"age":105,"chance":58,"impactCoefficient":10},
{"age":106,"chance":41,"impactCoefficient":10},
{"age":107,"chance":29,"impactCoefficient":10},
{"age":108,"chance":20,"impactCoefficient":10},
{"age":109,"chance":14,"impactCoefficient":10},
{"age":110,"chance":10,"impactCoefficient":10}
],
"female": [
{"age":18,"chance":99032,"impactCoefficient":30},
{"age":19,"chance":98997,"impactCoefficient":30},
{"age":20,"chance":98962,"impactCoefficient":30},
{"age":21,"chance":98933,"impactCoefficient":30},
{"age":22,"chance":98902,"impactCoefficient":40},
{"age":23,"chance":98866,"impactCoefficient":40},
{"age":24,"chance":98830,"impactCoefficient":40},
{"age":25,"chance":98795,"impactCoefficient":40},
{"age":26,"chance":98754,"impactCoefficient":50},
{"age":27,"chance":98698,"impactCoefficient":50},
{"age":28,"chance":98630,"impactCoefficient":50},
{"age":29,"chance":98558,"impactCoefficient":50},
{"age":30,"chance":98478,"impactCoefficient":60},
{"age":31,"chance":98385,"impactCoefficient":60},
{"age":32,"chance":98276,"impactCoefficient":70},
{"age":33,"chance":98152,"impactCoefficient":70},
{"age":34,"chance":98027,"impactCoefficient":80},
{"age":35,"chance":97896,"impactCoefficient":90},
{"age":36,"chance":97748,"impactCoefficient":100},
{"age":37,"chance":97592,"impactCoefficient":120},
{"age":38,"chance":97427,"impactCoefficient":130},
{"age":39,"chance":97247,"impactCoefficient":140},

{"age":40,"chance":97032,"impactCoefficient":160},
{"age":41,"chance":96795,"impactCoefficient":170},
{"age":42,"chance":96557,"impactCoefficient":180},
{"age":43,"chance":96307,"impactCoefficient":200},
{"age":44,"chance":96039,"impactCoefficient":210},
{"age":45,"chance":95752,"impactCoefficient":230},
{"age":46,"chance":95437,"impactCoefficient":250},
{"age":47,"chance":95089,"impactCoefficient":270},
{"age":48,"chance":94728,"impactCoefficient":290},
{"age":49,"chance":94326,"impactCoefficient":320},
{"age":50,"chance":93901,"impactCoefficient":350},
{"age":51,"chance":93461,"impactCoefficient":380},
{"age":52,"chance":92993,"impactCoefficient":410},
{"age":53,"chance":92500,"impactCoefficient":440},
{"age":54,"chance":91977,"impactCoefficient":480},
{"age":55,"chance":91434,"impactCoefficient":530},
{"age":56,"chance":90838,"impactCoefficient":570},
{"age":57,"chance":90187,"impactCoefficient":630},
{"age":58,"chance":89492,"impactCoefficient":680},
{"age":59,"chance":88736,"impactCoefficient":750},
{"age":60,"chance":87899,"impactCoefficient":810},
{"age":61,"chance":86990,"impactCoefficient":890},
{"age":62,"chance":85999,"impactCoefficient":970},
{"age":63,"chance":84945,"impactCoefficient":1060},
{"age":64,"chance":83835,"impactCoefficient":1160},
{"age":65,"chance":82612,"impactCoefficient":1270},
{"age":66,"chance":81280,"impactCoefficient":1390},
{"age":67,"chance":79858,"impactCoefficient":1510},
{"age":68,"chance":78296,"impactCoefficient":1650},
{"age":69,"chance":76561,"impactCoefficient":1800},

{"age":70,"chance":74664,"impactCoefficient":1960},
{"age":71,"chance":72526,"impactCoefficient":2130},
{"age":72,"chance":70337,"impactCoefficient":2310},
{"age":73,"chance":67974,"impactCoefficient":2510},
{"age":74,"chance":65416,"impactCoefficient":2710},
{"age":75,"chance":62816,"impactCoefficient":2910},
{"age":76,"chance":59870,"impactCoefficient":3120},
{"age":77,"chance":56771,"impactCoefficient":3330},
{"age":78,"chance":53551,"impactCoefficient":3540},
{"age":79,"chance":50107,"impactCoefficient":3740},
{"age":80,"chance":46479,"impactCoefficient":3920},
{"age":81,"chance":42620,"impactCoefficient":4070},
{"age":82,"chance":38671,"impactCoefficient":4180},
{"age":83,"chance":34640,"impactCoefficient":4250},
{"age":84,"chance":30632,"impactCoefficient":4250},
{"age":85,"chance":26784,"impactCoefficient":4160},
{"age":86,"chance":23094,"impactCoefficient":3980},
{"age":87,"chance":19622,"impactCoefficient":3720},
{"age":88,"chance":16409,"impactCoefficient":3430},
{"age":89,"chance":13486,"impactCoefficient":3010},
{"age":90,"chance":10878,"impactCoefficient":2570},
{"age":91,"chance":8598,"impactCoefficient":2120},
{"age":92,"chance":6647,"impactCoefficient":1680},
{"age":93,"chance":5017,"impactCoefficient":1290},
{"age":94,"chance":3689,"impactCoefficient":950},
{"age":95,"chance":2638,"impactCoefficient":670},
{"age":96,"chance":1829,"impactCoefficient":460},
{"age":97,"chance":1227,"impactCoefficient":300},
{"age":98,"chance":794,"impactCoefficient":190},
{"age":99,"chance":494,"impactCoefficient":110},

{"age":100,"chance":601,"impactCoefficient":60},
{"age":101,"chance":396,"impactCoefficient":60},
{"age":102,"chance":260,"impactCoefficient":60},
{"age":103,"chance":170,"impactCoefficient":60},
{"age":104,"chance":111,"impactCoefficient":60},
{"age":105,"chance":72,"impactCoefficient":60},
{"age":106,"chance":46,"impactCoefficient":60},
{"age":107,"chance":30,"impactCoefficient":60},
{"age":108,"chance":19,"impactCoefficient":60},
{"age":109,"chance":12,"impactCoefficient":60},
{"age":110,"chance":8,"impactCoefficient":60}
]
}

Додаток Г

Подяка від Державної установи «Центр громадського здоров'я Міністерства охорони здоров'я України»

ДЕРЖАВНА УСТАНОВА
«ЦЕНТР ГРОМАДСЬКОГО ЗДОРОВ'Я
МІНІСТЕРСТВА ОХОРОНИ ЗДОРОВ'Я УКРАЇНИ»
вул. Ярославська, 41, м. Київ, 04071, тел. (044) 425-43-54
E-mail: info@phc.org.ua, код ЄДРПОУ 40524109

«29» 07 2021 року № 9/1934/27
на № _____ від «___» _____ 20__ року

Директору Інституту кібернетики
ім. В.М. Глушкова НАН України,
Академіку НАН України
Сергієнку І.В.

Щодо співпраці з співробітниками
ІК ім. В.М. Глушкова НАН України

Вельмишановний Іване Васильовичу!

Державна установа «Центр громадського здоров'я Міністерства охорони здоров'я України» (далі – Центр) засвідчує Вам свою повагу та хоче подякувати за плідну співпрацю з науковцями Інституту.

За ініціативи Центру нещодавно був створений та оприлюднений калькулятор очікуваної тривалості життя (проект «Калькулятор життя») сучасний інструмент який аналізує низку факторів, що впливає на довголіття звички, харчову поведінку, хронічні захворювання, фізичну активність. Він показує також, наскільки позитивні зміни збільшують тривалість життя. Перші ж дні після оприлюднення «Калькулятора життя» засвідчили велику затребуваність суспільством цього інструменту.

Науковим консультантом цієї важливої роботи є провідний науковець Інституту кібернетики, знаний фахівець з медичної кібернетики та інформатики а також інноваційних методів функціональної діагностики І. А. Чайковський.

Програмний модуль калькулятору розробив В.Ю. Маланін, аспірант Інституту кібернетики ім. В.М. Глушкова НАН України, фахівець з штучного інтелекту та хмарних обчислень. Вказані співробітники Інституту кібернетики проявили високу кваліфікацію та творчий підхід, їхній доробок у проекті є дуже цінним. Також хочемо відмітити доповідь з назвою "Modern technological innovations in public health system : "pocket" diagnostic devices", зроблену І.А. Чайковським в рамках вебінару «Global Health: Ukraine and Italy Experience", який відбувся 20-го травня як частина Европейського тижня

громадського здоров'я. Ця доповідь привернула увагу широкого загалу фахівців до важливих інновацій, що розроблені українськими науковцями, насамперед в Інституті кібернетики.

Нам добре відомий великий багатосторонній науковий потенціал колективу Інституту кібернетики. Сподіваємось на подальшу плідну співпрацю в інтересах збереження та покращення здоров'я українців.

З повагою
В.о. Генерального директора **Олександр ВОВЧЕНКО**

Агєєва І.Ю.
067 214 21 38