

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СУМСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ЕЛЕКТРОНІКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК
СЕКЦІЯ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ПРОЕКТУВАННЯ

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

на тему: «Інформаційна технологія оцінювання ризику захворюваності на цервікальний рак за допомогою нейронних мереж»

за спеціальністю 122 «Комп'ютерні науки»,
освітньо-професійна програма «Інформаційні технології проектування»

Виконавець роботи: студентка групи ІТ.м.н-71 Єлісеєва Анна Романівна

**Кваліфікаційну роботу
захищено на засіданні ЕК
з оцінкою**

«___» травня 2019 р.

Науковий керівник

(підпис)

к.т.н., доц., Шендрик В.В.

Голова комісії

(підпис)

к.т.н. Шифрин Д.М.

Засвідчую, що у цій дипломній роботі немає
запозичень з праць інших авторів
без відповідних посилань.

Студент _____
(підпис)

Суми-2019

Сумський державний університет
Факультет електроніки та інформаційних технологій
Кафедра комп'ютерних наук
Секція інформаційних технологій проектування
Спеціальність 122 «Комп'ютерні науки»
Освітньо-наукова програма «Інформаційні технології проектування»

ЗАТВЕРДЖУЮ

Зав. секцією ІТП

_____ В. В. Шендрик
«__» _____ 2019 р.

ЗАВДАННЯ

на кваліфікаційну роботу магістра студентіві

Єлісєєва Анна Романівна

(прізвище, ім'я, по батькові)

1 Тема проекту Інформаційна технологія оцінювання ризику захворюваності на цервікальний рак за допомогою нейронних мереж

затверджена наказом по університету від «05» лютого 2019 р. №0237-III

2 Термін здачі студентом закінченого проекту « 18 » травня 2019 р.

3 Вхідні дані до проекту _____ Масив неопрацьованої вхідної інформації
_____ про індивідуальні показники організму пацієнтів з результатами
_____ захворюваності на цервікальний рак

4 Зміст розрахунково-пояснювальної записки (перелік питань, що їх належить розробити) _____ 1 Аналіз предметної області

_____ 2 Постановка задачі та методи дослідження

_____ 3 Моделювання інформаційної технології

_____ 4 Розробка інформаційної технології

5 Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень) _____ Презентація

6. Консультанти випускної роботи із зазначенням розділів, що їх стосуються:

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

Дата видачі завдання _____.

Керівник _____
(підпис)

Завдання прийняв до виконання _____
(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів випускної проекту	Термін виконання етапів проекту	Примітка
1	Аналіз предметної області	До 10.01.2019 р.	
2	Постановка задачі інформаційної технології	До 01.02.2019 р.	
3	Розробка нейронної мережі	До 01.03.2019 р.	
4	Створення бази даних	До 01.04.2019 р.	
5	Розробка інформаційної технології	До 28.04.2019 р.	
6	Тестування інформаційної технології	До 10.05.2019 р.	
7	Оформлення пояснювальної записки до дипломного проекту	До 17.05.2019 р.	
8	Підготовка презентації	До 18.05.2019 р.	
9	Захист дипломного проекту	21.05.2019 р.	

Магістрант _____

Слісєєва А.Р.

Керівник роботи _____

к.т.н., доц. Шендрік В.В.

РЕФЕРАТ

Тема кваліфікаційної роботи магістра «Інформаційна технологія оцінювання ризику захворюваності на цервікальний рак за допомогою нейронних мереж».

Пояснювальна записка складається зі вступу, 4 розділів, висновків, списку використаних джерел із 63 найменувань, додатків. Загальний обсяг роботи – 180 сторінок, у тому числі 111 сторінок основного тексту, 65 рисунків, 15 таблиць, 6 сторінок списку використаних джерел, 3 додатки, які складають 62 сторінки.

Кваліфікаційну роботу магістра присвячено розробці інформаційної технології оцінювання ризику захворюваності на цервікальний рак за допомогою нейронних мереж. У роботі виконано аналіз проведених досліджень, визначена постановка задачі та методи дослідження, а також змодельовано інформаційну технологію з описом реалізації її роботи. У роботі виконано розробку інформаційної технології у вигляді web-додатку, який у свою чергу підключений до бази даних PostgreSQL та містить у собі Python-скрипт, у якому зберігається навчена нейронна мережа. Інформаційна технологія дозволяє обробляти введені показники, запропоновані web-додатком, зберігати усі дані до спеціально розробленої бази даних та завдяки попередньо навченій нейронній мережі на тренувальних та тестових даних коректно виконувати оцінку ризику появи досліджуваного онкологічного захворювання. Результатом проведеної роботи є web-додаток, завдяки якому користувачі зареєструвавшись або авторизувавшись у ньому, можуть виконати запит для оцінювання ризику захворюваності на цервікальний рак на основі індивідуальних показників організму.

Ключові слова: ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ, ЦЕРВІКАЛЬНИЙ РАК, НЕЙРОННА МЕРЕЖА, ОЦІНКА РИЗИКУ, POSTGRESQI, PYTHON, WEB-ДОДАТОК.

ЗМІСТ

Вступ.....	8
1 Аналіз предметної області	9
1.1 Аналіз проведених досліджень	9
1.2 Аналіз наявних аналогів програмних продуктів	12
2 Постановка задачі та методи дослідження.....	19
2.1 Мета та задачі дослідження.....	19
2.2 Методи дослідження	20
2.3 Вибір засобів реалізації.....	23
3 Моделювання інформаційної технології.....	26
3.1 Побудова діаграм в нотації сімейства IDEF	26
3.1.1 Діаграма у нотації IDEF0.....	26
3.1.2 Діаграма у нотації DFD.....	28
3.1.3 Діаграма у нотації IDEF3.....	29
3.2 Архітектура інформаційної системи	30
3.3 Побудова діаграми варіантів використання	31
3.4 Побудова діаграми послідовності.....	33
3.5 Розробка алгоритму створення неронної мережі	34
3.5.1 Ціль класифікації.....	34
3.5.2 Підготовка початкових даних	34
3.5.3 Кодування вихідних значень.....	35
3.5.4 Вибір об'єму мережі	35

3.5.5	Вибір архітектури мережі.....	36
3.5.6	Алгоритм побудови класифікатора на основі нейронних мереж.....	36
3.6	Концептуальне моделювання.....	38
3.6.1	Опис сутностей.....	39
3.6.2	Зв'язки між об'єктами.....	51
3.7	Логічне проектування	52
3.7.1	Нормалізація таблиць.....	52
3.7.2	Функціональні залежності.....	53
3.7.3	Визначення первинних та зовнішніх ключів.....	56
3.8	Фізичне моделювання	56
3.8.1	Створення базових таблиць.....	56
3.8.2	Створення унікального індексу первинного ключа.....	57
3.8.3	Створення зовнішніх ключів.....	57
4	Розробка інформаційної технології.....	59
4.1	Обробка даних.....	59
4.1.1	Обробка вхідних даних.....	61
4.1.2	Розробка навчальної та тестової вибірок.....	69
4.1.3	Кореляція між параметрами вхідних даних	70
4.2	Створення бази даних	73
4.3	Імпорт даних до бази даних.....	75
4.4	Розробка нейронної мережі	77
4.5	Створення web-додатку.....	89
4.5.1	Створення моделей (Models).....	91
4.5.2	Створення представлення (Views).....	93

4.5.3	Створення контролерів (Controllers)	95
4.5.4	Підключення до бази даних	96
4.5.5	Підключення зовнішніх модулів для авторизації користувача	97
4.5.6	Валідація введених даних.....	98
4.5.7	Підключення Python-скрипту з нейронною мережею.....	99
4.5.8	Міграції.....	100
4.5.9	Модулі проекту.....	102
4.6	Реалізація роботи web-додатку	103
	Висновок	110
	Список літератури	112
	Додаток А.....	118
	Додаток Б.....	132
	Додаток В	139

ВСТУП

На даний час вчені всього світу приділяють все більшу увагу нейронним мережам у різних сферах людської діяльності. З розвитком інформаційних технологій у медицині почали використовувати комп'ютерні методи та експертні системи для діагностики різноманітних хвороб та постановки діагнозу.

Однією з найбільших переваг нейронних мереж у медицині є швидка та точна обробка даних в умовах недостатньої визначеності необхідних параметрів, які важко детермінувати лікареві на початкових стадіях розвитку онкологічного захворювання.

За даними Всесвітньої організації охорони здоров'я у світі рак шийки матки (РШМ) займає друге місце у структурі онкологічної захворюваності з-поміж жінок віком 15-44 роки. Згідно статистичних даних такий діагноз ставиться понад 600 тис. жінок кожного року. Виявлення раку шийки матки є складним завданням навіть у сучасному світі з прогресивною медициною, оскільки дане захворювання протікає без будь-яких симптомів. Проте більш складним завданням для лікарів є оцінка ризику виявлення даного захворювання у жінок, спираючись лише на клінічні аналізи без застосування діагностичних скринінгових систем.

Основним засобом для оцінки ризику виявлення ракових пухлин в області шийки матки будуть використовуватися нейронні мережі, так як вони мають перевагу над іншими системами для обробки і аналізу даних завдяки здатності до навчання та виправлення помилок.

На даний момент існує досить велика кількість наукових робіт автоматичної класифікації захворювання на цервікальний рак (ЦР), проте вони спираються на цитологічні зображення за допомогою згорткових нейронних мереж. Також приділяють увагу прогнозуванню результату виживання жінок з поставленим діагнозом за допомогою різних методів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Аналіз проведених досліджень

Медицина завжди використовує передові технології. Технологічні досягнення, такі як комп'ютери, лазери, ультразвукові зображення тощо, набагато збільшили рівень медицини. У даний час штучна нейронна мережа є наступною перспективною сферою інтересів для вчених. Вважається, що нейронні мережі матимуть широке застосування до біомедичних проблем у найближчі кілька років. Існують перші успіхи її застосування в таких областях медицини, як діагностичні системи, біохімічний аналіз, аналіз зображень і розробка лікарських засобів [1].

Нейронні мережі – нелінійні системи, які дозволяють більш точно класифікувати дані, ніж звичайно використовувані лінійні методи. У додатку до медичній діагностиці вони дають можливість значно підвищити специфічність методу, не знижуючи його чутливості [2].

Одними з перших, хто почав застосовувати нейронні мережі для діагностики раку шийки матки були вчені з Осацького університету. Вчені пояснили, що ракові клітини можуть відрізнятися не тільки у різних пацієнтів, але і у одного і того ж пацієнта. Визначення різновидів ракових клітин людини допомагає медикам вибрати найбільш ефективний метод лікування. На практиці завдання з ідентифікації ракових клітин займає у фахівця багато часу, ця робота ускладнюється через людські фактори ризику.

Команда вчених з Осацького університету навчила штучний інтелект виявляти серед ракових клітин шийки матки резистентні для радіотерапії. Нейронній мережі передаються для аналізу відскановані зображення клітин, отримані за допомогою мікроскопа. Штучний інтелект аналізує зображення і виконує класифікацію клітин [3]. Навчання нейронної мережі проводилося на 8000

зображень та перевірялося на 2000 зображень. Точність отриманих результатів становила 96-98%.

Малишевська Е.Н. виконувала аналіз використання нейромереж для діагностики раку шийки матки за мультиспектральними зображеннями. Нею була створена комп'ютерна система, яка може правильно класифікувати різні види тканин за мультиспектральним зображенням, застосовуючи нейронні мережі. У її роботі розглянута, в першу чергу діагностика станів дисплазії і неоплазії (CIN-I, II, III), а також наявність тканин з доброякісними змінами (MET) [4].

Смекалов В.П., Сепп А.В. займалися вивченням можливостей застосування штучних нейронних мереж для прогнозування результатів лікування передракових захворювань шийки матки. Клінічні і морфологічні параметри, що визначають результат лікування у їхньому дослідженні піддавалися факторному аналізу для виявлення найбільш значних чинників ризику недосягнення позитивного результату. У результаті були відібрані такі значення, як вік хворого, методи лікування, що застосовувалися раніше, результати після хірургічного лікування (терміни епітелізації), зміни морфологічного висновку при повторних зверненнях, їх кількість, факт інфікування вірусом папіломи людини (HPV-16, 18, 31 ...) і виставлений морфологічний діагноз. Усі процедури навчання проводилися з використанням комп'ютерної програми Statistica Neural Network (SNN) v.4.0 розробки StatSoft [5].

М. Anousouya Devi, S. Ravi займалися класифікацією раку шийки матки, використовуючи нейронні мережі. Вони займалися оцінкою точності результатів роботи нейронних мереж, використовуючи різні методи для їх навчання [6].

Stephen Pfohl, Oskar Triebe та Ben Marafino займалися розробкою керування лікуванням раку шийки матки з згортковими нейронними мережами. Вони описали розробку та впровадження згорткового алгоритму нейромережі для класифікації раку шийки матки типу TZ за допомогою зображень злоякісних пухлин [7].

Siti Noraini Sulaimana, Nor Ashidi Mat-Isab, Nor Hayati Othmanc, Fadzil Ahmada виконували удосконалення особливостей процесу екстракції та класифікації раку шийки матки для системи NeuralPap. Результати їх досліджень показують, що всі

використані ними алгоритми для розробки нейронних мереж здатні ідентифікувати клітини шийки матки на основі візуальної оцінки. Результати також підтримують кількісні оцінювання, при якому існує висока кореляційна величина у порівнянні з клінічним оглядом [8].

Haidar A. Almubaraka, R. Joe Stanleya, Rodney Longb, Sameer Antanib, George Thomab, Rosemary Zunac, Shelliane R. Frazierd проводили локалізовану класифікацію зображень на основі маткових цервікальних ракових захворювань, що базуються на нейромережевій мережі [9].

Priyanka K Malli, Dr. Suvarna Nandyal виконали розробку техніки машинного навчання для виявлення рак шийки матки, використовуючи k-NN і штучну нейронну мережу. Результатом дослідження є база машинного навчання в поєднанні з ефективною нечіткою логікою, яка у свою чергу включає точну сегментацію і класифікацію клітини в нормальну або ракову та визначає її типи. Середній результат класифікації був отриманий з точністю 88,04% для k-NN і 54% для штучної нейронної мережі [10].

Xiaoping Qiu, Ning Tao, Yun Tan, Xinxing Wu займалися побудовою моделі класифікації ризику раку шийки матки за допомогою штучних мереж. У своїх дослідженнях вони застосовували метод зворотного поширення помилки для побудови даної моделі з отриманими результатами її роботи для показників чутливості та специфічності, що становить 98% та 97% відповідно [11].

Koji Matsuo, SanjayPurushotham, BoJiang, Rachel S.Mandelbaum та інші проводили дослідження для прогнозування результату виживання при раку шийки матки, порівнюючи модель Кокса та модель глибокого навчання. У цілому було сформовано 40 ознак, включаючи демографічні показники пацієнтів, життєво важливі ознаки, результати лабораторних досліджень, характеристики пухлини та типи лікування, які були оцінені для аналізу та згруповані в три набори ознак. Модель поглибленого вивчення нейронної мережі порівнювали з моделлю пропорційної регресії ризику Кокса і трьома іншими моделями аналізу для виживання без прогресування і загального виживання [12].

Актуальність роботи полягає у тому, що наданий час у медицині відсутні аналоги для проведення аналізу та оцінювання ризику появи захворюваності раку шийки матки, спираючись на досить велику кількість невизначеностей та впливових факторів. Такими основними факторами є індивідуальні показники характеристики жінки (вік, кількість вагітностей, кількість статевих партнерів, використання гормональних контрацептивів, наявність захворювань, що передаються статевим шляхом тощо), результати клінічних аналізів (тест Хільсельмана, тест Шилера, цитологія, біопсія тощо), шкідливі звички (паління), які впливають на ризик появи та характер протікання даного онкологічного захворювання. Саме тому вирішення такої проблеми вимагає використання адаптивних автоматизованих систем, які використовують гнучкі нейронні мережі.

1.2 Аналіз наявних аналогів програмних продуктів

Навесні 1992 компанія Neuromedical Systems Inc. of Suffern, NY, випустила на ринок продукт під назвою Papnet. Система Papnet являє собою інструмент, що використовує принцип нейронних мереж, який допомагає фахівцям цитології діагностувати рак шийки матки більш точно і, що важливіше, з меншими затратами. Papnet використовує вдосконалену систему розпізнавання образів, побудовану на принципі нейронних мереж, і відбирає 128 найбільш «підозрілих» клітин досліджуваного мазка для подальшої оцінки фахівцем-цитологом.

Тест-система діагностики раку «ЕпіДжін» проектної компанії новосибірського наноцентру. Технологія, що використовується в тест-системі «ЕпіДжін» - епігенетичний аналіз молекули ДНК з використанням особливого ферменту, який з високою точністю оцінює статус генів, що пригнічують розвиток пухлин при нормальній роботі. Епігенетичний аналіз виявляє «відключення» цих генів на ранній стадії утворення пухлини, ще до прояву клінічних ознак. Розробка тест-системи для

колоректального раку (злоякісна пухлина товстого кишечника та прямої кишки у людини) завершена. Готуються тест-системи для діагностики раку молочної залози, раку шлунку і раку легенів.

Молекулярна Діагностична Система *in vivo* (IVMD) заснована на конфокальній Раманівській спектроскопії і включає запатентований конфокальний волоконно оптичний зонд і спеціалізоване програмне забезпечення, яке протягом декількох секунд обробляє біомолекулярні відбитки тканини і дозволяє поставити діагноз безпосередньо в процесі ендоскопічного дослідження. IVMD система використовувалася в Сінгапурі у більш, ніж 500 пацієнтів з раком шлунку, стравоходу, товстої кишки, прямої кишки, голови і шиї, шийки матки.

Короткий огляд показників якості, що характеризують альтернативи. Було визначено наступні критерії для вибору найкращої з альтернатив:

- вартість системи;
- додаткова перевірка цитологом;
- виявлення стадії раку;
- визначення лікування.

Обґрунтування методу вибору доцільного варіанту. Система підтримки прийняття рішень (СППР) "Вибір" - аналітична система, заснована на методі аналізу ієрархій (MAI), є простим і зручним засобом, яке допоможе:

- структурувати проблему;
- побудувати набір альтернатив;
- виділити характеризують їх чинники;
- задати значимість цих факторів;
- оцінити альтернативи по кожному з факторів;
- знайти неточності і протиріччя в судженнях особи приймає рішення;
- провести аналіз рішення і обґрунтувати отримані результати.

Система спирається на математично обґрунтований метод аналізу ієрархій Томаса Сааті. СППР "Вибір" на основі MAI може використовуватися при вирішенні наступних типових задач:

- оцінка якості організаційних, проектних і конструкторських рішень;
- визначення політики інвестицій в різних областях; завдання розміщення (вибір місця розташування шкідливих та небезпечних виробництв, пунктів обслуговування);
- розподіл ресурсів;
- проведення аналізу проблеми за методом "вартість-ефективність";
- стратегічне планування;
- проектування і вибір устаткування, товарів;
- вибір професії, місця роботи, підбір кадрів.

Реалізація процедури оцінювання альтернатив і вибору оптимальної альтернативи. Використовуючи програму ССПР «Вибір» виконаємо побудову матриць попарних порівнянь на рівні цілей та визначимо важливість критеріїв між собою (рис. 1.1).

Получение матрицы парных сравнений

Относительно фактора
Уровень1 Ціль
необходимо провести парное
сравнение следующих факторов
уровня
Уровень2

№	Фактор	Вес
1	Вартість	0,114
2	Додаткова переві...	0,077
3	Виявлення стадії ...	0,400
4	Визначення лікув...	0,409

Матрица парных сравнений:

	1	2	3	4
1	1	3	1/6	1/5
2	1/3	1	1/3	1/5
3	6	3	1	1
4	5	5	1	1

Какой из факторов предпочтительнее ?

Степень предпочтения:

Вартість
 Вартість
 Одинаково важны
 Не могу сказать

Абсолютно превосходит
 Промежуточное значение
 Значительно превосходит
 Промежуточное значение
 Существенно превосходит
 Промежуточное значение
 Умеренно превосходит
 Промежуточное значение
 Одинаково важны

Просмотр проекта $\lambda = 4,309$ ИС = 0,103 ОС = 0,115 OK Отмена

Рисунок 1.1 – Матриця попарних порівнянь на рівні цілей

Побудуємо матрицю порівнянь для критерія вартості (рис. 1.2).

Получение матрицы парных сравнений

Относительно фактора
Уровень 2. Вартість
необходимо провести парное
сравнение следующих факторов
уровня
Уровень 3

№	Фактор	Вес
1	Рарнет	0,185
2	ЕпіДжін	0,030
3	IVMD	0,085
4	Дипломний проект	0,700

Матрица парных сравнений:

	1	2	3	4
1	1	7	5	1/9
2	1/7	1	1/7	1/9
3	1/5	7	1	1/9
4	9	9	9	1

Какой из факторов предпочтительнее ?

IVMD
 Дипломний проект
 Одинаково важны
 Не могу сказать

Степень предпочтения:

Абсолютно превосходит
 Промежуточное значение
 Значительно превосходит
 Промежуточное значение
 Существенно превосходит
 Промежуточное значение
 Умеренно превосходит
 Промежуточное значение
 Одинаково важны

Просмотр проекта $\lambda = 4,854$ ИС = 0,285 ОС = 0,316

Рисунок 1.2 – Матрица попарных порівнянь для критерія вартості

Побудуємо матрицю порівнянь для критерія додаткова перевірка цитологом (рис. 1.3).

Получение матрицы парных сравнений

Относительно фактора
Уровень 2. Додаткова перевірка чит
необходимо провести парное
сравнение следующих факторов
уровня
Уровень 3

№	Фактор	Вес
1	Рарнет	0,119
2	ЕпіДжін	0,222
3	IVMD	0,119
4	Дипломний проект	0,539

Матрица парных сравнений:

	1	2	3	4
1	1	1	1	1/7
2	1	1	1	1
3	1	1	1	1/7
4	7	1	7	1

Какой из факторов предпочтительнее ?

IVMD
 Дипломний проект
 Одинаково важны
 Не могу сказать

Степень предпочтения:

Абсолютно превосходит
 Промежуточное значение
 Значительно превосходит
 Промежуточное значение
 Существенно превосходит
 Промежуточное значение
 Умеренно превосходит
 Промежуточное значение
 Одинаково важны

Просмотр проекта $\lambda = 4,505$ ИС = 0,168 ОС = 0,187

Рисунок 1.3 – Матрица попарных порівнянь для критерія додаткова перевірка
ЦИТОЛОГОМ

Побудуємо матрицю порівнянь для критерія виявлення стадії раку (рис. 1.4).

Получение матрицы парных сравнений

Относительно фактора
Уровень 2. Выявления стадии рака
необходимо провести парное
сравнение следующих факторов
уровня
Уровень 3

№	Фактор	Вес
1	Рарнет	0,125
2	Епіджин	0,125
3	IVMD	0,125
4	Дипломний проект	0,625

Матрица парных сравнений:

	1	2	3	4
1	1	1	1	1/5
2	1	1	1	1/5
3	1	1	1	1/5
4	5	5	5	1

Какой из факторов предпочтительнее ?

IVMD
 Дипломний проект
 Одинаково важны
 Не могу сказать

Степень предпочтения:

Абсолютно превосходит
 Промежуточное значение
 Значительно превосходит
 Промежуточное значение
 Существенно превосходит
 Промежуточное значение
 Умеренно превосходит
 Промежуточное значение
 Одинаково важны

Просмотр проекта $\lambda = 4,000$ ИС = 0,000 ОС = 0,000

Рисунок 1.4 – Матрица попарных порівнянь для критерія виявлення стадії раку

Побудуємо матрицю порівнянь для критерія визначення лікування (рис. 1.5).

Получение матрицы парных сравнений

Относительно фактора
Уровень 2. Визначення лікування
необходимо провести парное
сравнение следующих факторов
уровня
Уровень 3

№	Фактор	Вес
1	Рарнет	0,158
2	Епіджин	0,070
3	IVMD	0,070
4	Дипломний проект	0,702

Матрица парных сравнений:

	1	2	3	4
1	1	3	3	1/8
2	1/3	1	1	1/7
3	1/3	1	1	1/7
4	8	7	7	1

Какой из факторов предпочтительнее ?

IVMD
 Дипломний проект
 Одинаково важны
 Не могу сказать

Степень предпочтения:

Абсолютно превосходит
 Промежуточное значение
 Значительно превосходит
 Промежуточное значение
 Существенно превосходит
 Промежуточное значение
 Умеренно превосходит
 Промежуточное значение
 Одинаково важны

Просмотр проекта $\lambda = 4,204$ ИС = 0,068 ОС = 0,076

Рисунок 1.5 – Матрица попарних порівнянь для критерія визначення лікування

Як бачимо з діаграм дипломний проект має найбільшу вагу за визначеними критеріями для діагностики раку шийки матки та визначення лікування даного онкологічного захворювання (рис. 1.6 – 1.7).

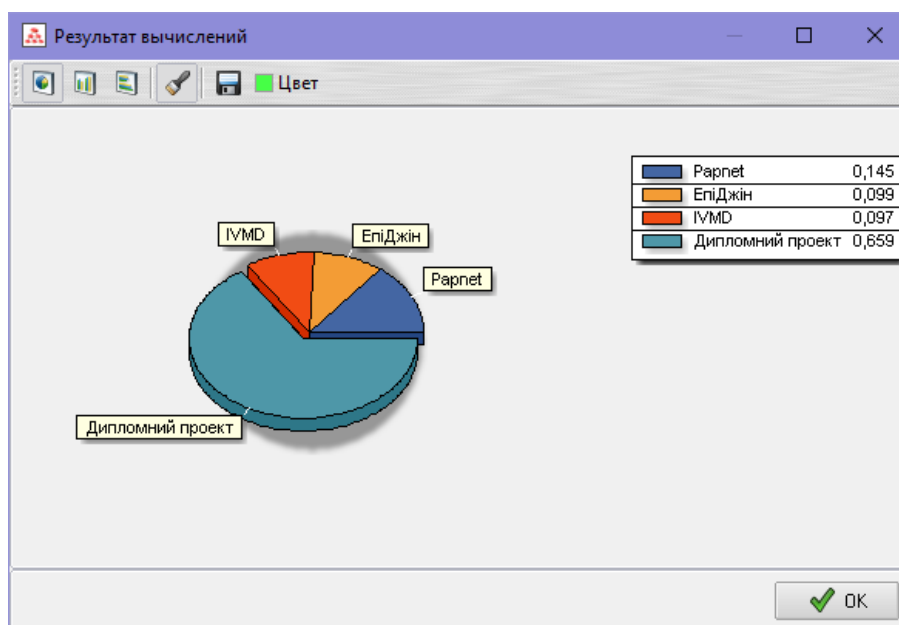


Рисунок 1.6 – Результат обчислень у вигляді кругової діаграми

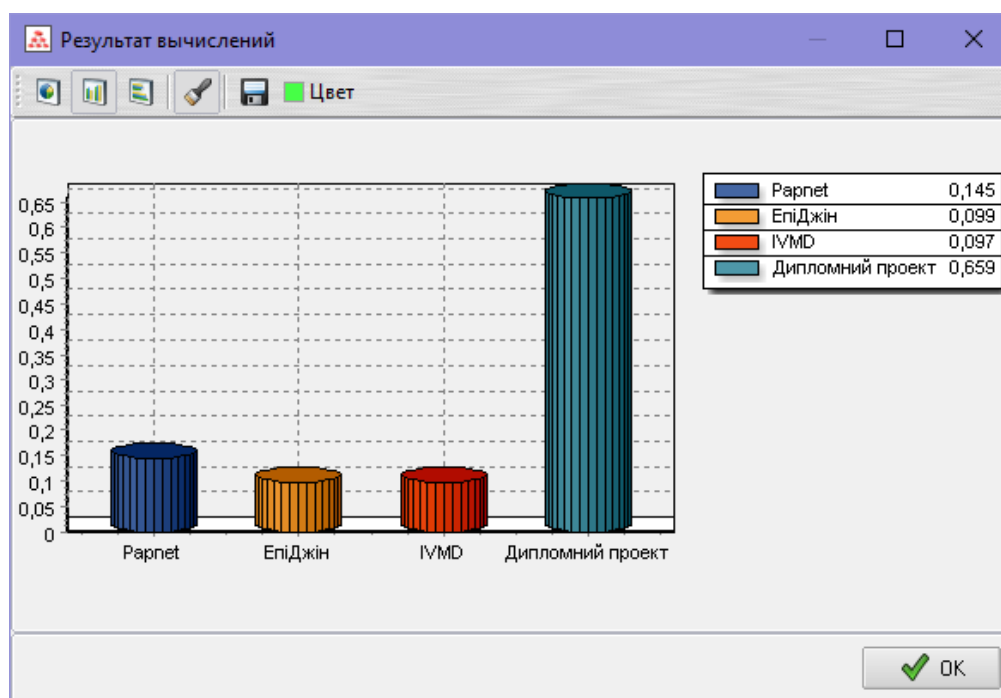


Рисунок 1.7 – Результат обчислень у вигляді стовпчикової діаграми

Результатом даного етапу є проведений аналіз предметної області, а саме досліджень, які проводилися для вивчення питання прогнозування виникнення раку шийки матки за допомогою нейронних мереж, а також було здійснено огляд наявних аналогів програмних продуктів, які можна використовувати для отримання результатів оцінки ризику виявлення захворюваності на рак.

2 ПОСТАНОВКА ЗАДАЧІ ТА МЕТОДИ ДОСЛІДЖЕННЯ

2.1 Мета та задачі дослідження

Метою даної роботи є розробка інформаційної технології для оцінки ризику захворюваності цервікального раку на основі індивідуальних показників жінки за допомогою нейронних мереж.

Об'єктом дослідження є алгоритм функціонування інформаційної технології оцінювання ризику захворюваності на РШМ за допомогою нейронних мереж.

Предметом дослідження є моделі нейронних мереж, за допомогою яких буде проводитися оцінка ризику появи цервікального раку у жінки.

Задачі, які слід вирішити для досягнення поставленої мети:

- огляд сучасних методів автоматизованого аналізу і обробки статистичних даних;
- вибір методів для оцінювання ризику розвитку цервікальної ракової пухлини, заснованих на нейронних мережах;
- побудова нейронної мережі та створення її оптимальної архітектури з відповідними параметрами та її навчання;
- виконання оцінки якості нейронної мережі;
- виконання перевірки ступеню впливу різних факторів на результати, отримані нейромережею;
- розробка інформаційної технології з інтуїтивним інтерфейсом для зручного та швидкого доступу для виконання аналізу.

Об'єктом спостереження є алгоритм прогнозування та оцінки ризику на цервікальний рак. Вхідним матеріалом для проведення дослідження становлять результати аналізів опитаних 859 жінок. Статистичні дані були вилучені на офіційному сайті збору статистичних даних <https://www.kaggle.com>. Всі дані будуть

поділені для навчальної (80%) та тестової вибірок (20%). Вхідним значенням буде результат чи ризикує (позитивний) жінка захворіти даним онкологічним захворюванням чи ні (негативний).

Спираючись на проведений аналіз предметної області можна сказати, що більшість досліджень щодо діагностики раку шийки матки виконувалися на основі аналізу відповідних зображень злоякісних пухлин. Також було розроблено та покращено досить велика кількість методів для навчання нейронної мережі з тестуванням її функціоналу в спеціалізованих платних додатках. Проте кожне з досліджень висвітлює лише один аспект у діагностиці або лікуванні цервікального раку, та не розглядають дану проблему комплексно. Відсутність відповідного інструменту, який включає обробку статистичних даних на основі проведених клінічних аналізів, технології виявлення онкологічного захворювання, проведення аналізу для встановлення стадії його розвитку, прогнозування відповідного лікування з мінімальними затратами часу та прогнозування появи рецидивів, обумовила на розробку універсальної діагностичної системи з використанням нейронних мереж.

Тому новизна результатів полягає у тому, що було одержано подальший розвиток способу використовувати нейронні мережі для діагностики РШМ на основі обробки статистичних даних за результатами аналізів жінок. Також буде удосконалено математичну модель функціональної нейромережі, шляхом оцінки якості її навчання.

2.2 Методи дослідження

Науковою проблемою є складність оцінювання виникнення цервікального раку за допомогою нейронних мереж на основі визначених факторів ризику, які по

різному впливають на організм жінки з можливістю подальшої появи та прогресування даної злоякісної пухлини.

Для прогнозування результатів лікування РШМ потрібно розробити нейронну мережу результату його наявності в залежності від епідеміологічних і клінічних факторів.

Основними факторами ризику розвитку даного виду онкологічного захворювання є:

- вірус папіломи людини;
- вік;
- велика кількість сексуальних партнерів;
- рання вагітність;
- велика кількість вагітностей;
- інші захворювання, що передаються статевим шляхом;
- куріння;
- слабкий імунітет;
- вживання протизаплідних ліків;
- соціоекономічний статус [13, 14, 15, 16].

Інформаційна технологія повинна вирішувати основне завдання – проводити оцінку ризику появи захворюваності даного онкологічного захворювання в залежності від наступних параметрів:

- вік жінки;
- кількість статевих партнерів;
- у якому віці відбувся перший статевий акт;
- кількість вагітностей;
- чи палить жінка;
- скільки років палить жінка;
- кількість пачок випалених цигарок на рік;
- чи вживала жінка гормональні контрацептиви;
- яка тривалість у роках вживання гормональних контрацептивів;

- чи використовувала жінка внутрішньоматкову спіраль;
- яку кількість років жінка використовувала внутрішньоматкову спіраль;
- чи хворіла жінка захворюваннями, що передаються статевим шляхом;
- кількість захворювань, якими хворіла жінка що передаються статевим шляхом;
- чи хворіла жінка на кондиломатоз;
- чи хворіла жінка на цервікальний кондиломатоз;
- чи хворіла жінка на вагінальний кондиломатоз;
- чи хворіла жінка на вульво-перинеальний кондиломатоз;
- чи хворіла жінка сифілісом;
- чи були у жінки запальні захворювання тазових органів;
- чи хворіла жінка на генітальний герпес;
- чи була у жінки хвороба: контагіозний молюск;
- чи хворіла жінка СНІДом;
- чи хворіла жінка ВІЛом;
- чи хворіла жінка гепатитом Б;
- чи хворіла жінка вірусом папіломи людини;
- загальна кількість діагнозів;
- час з моменту першого діагнозу, у роках;
- час, який минув від останнього діагнозу, у роках;
- чи був вже раніше продіагностований цервікальний рак;
- чи була у жінки дисплазія шийки матки;
- чи був продіагностований вірус папіломи людини;
- чи був раніше продіагностований рак;
- результати кольпоскопії Хінсельмана;
- результати тесту Шилера;
- результати цитології.

Цільовими змінними є:

- результати біопсії.

Дані параметри пацієнтів будуть використовуватися в якості вхідного шару нейронів.

Для вирішення задачі прогнозування діагностування раку шийки матки було обрано метод класифікації за допомогою штучних нейронних мереж типу багат шарового персептрону з методом навчання бінарної крос-ентропії. Багат шаровий персептрон є найбільш поширеним методом вирішення задач класифікації [17, 18, 19]. Так як результатом проведеного дослідження оцінки ризику виникнення раку шийки матки є «так» (1) або «ні» (0), то можемо зробити висновок, що вирішувана задача має тип бінарної класифікації. Крос-ентропія – це функція втрати за замовчуванням, яка використовується саме для таких задач, де цільові значення знаходяться в наборі $\{0, 1\}$. Бінарна крос-ентропія вимірює як далеко від істинного значення (0 або 1) прогнозування для кожного з класів, а потім складає в середньому ці класові помилки, щоб отримати фінальне значення втрати [20, 21, 22, 23].

2.3 Вибір засобів реалізації

Для зручного представлення необхідної інформації було створено веб-додаток, реалізованого за основі кросплатформного середовища ASP.NET Core з відкритим вихідним кодом для створення сучасних хмарних додатків, підключених до мережі Інтернет. Його головними перевагами є:

- можливість створювати веб-додатки та служби, серверні частини для мобільних додатків;
- можливість використовувати вибрані засоби розробки в Windows, macOS і Linux;

- можливість виконувати розгортання в хмарі або локальному середовищі;
- можливість працювати в .NET Core або .NET Framework [24].

У порівнянні з іншими платформами для розробки веб-додатків ASP.NET Core містить зручний метод для створення користувацького веб-інтерфейсу – Razor Pages, який робить створення кодів сценаріїв для сторінок простіше та ефективніше, підтримує версійність на одному локальному комп'ютері, можливість створення веб-додатку на різних програмних продуктах, таких як Visual Studio, Visual Studio для Mac або Visual Studio Code з використанням C# або F#. ASP.NET Core має більш високу продуктивність, ніж, наприклад, ASP.NET 4.x або .NET Framework [25].

Усі отримані дані зберігаються у реляційній базі даних – об'єктно-реляційній системі управління базами даних PostgreSQL. Це надійна гнучка система, яка містить велику кількість функцій, що у свою чергу дозволяє розробникам створювати додатки практично на всіх UNIX-платформах, системі Windows, а також FreeBSD и Linux. Вона є безкоштовною та має відкритий вихідний код, який забезпечує можливість створювати власні типи даних та функції, а також писати код на різних мовах програмування без перекомпіляції баз даних. PostgreSQL може застосовуватися в клієнт-серверному середовищі, завдяки чому зручно створювати розподілену систему. Клієнтські програми не можуть отримати доступ до даних самостійно, навіть якщо вони працюють на тому ж комп'ютері, на якому виконується серверний процес [26, 27].

Робота з нейронною мережею здійснювалася за допомогою високорівневої мови програмування Python. Його основними перевагами є простота та потужний функціонал, завдяки стандартним та великій кількості додаткових бібліотек, які можна використовувати в залежності від типу виконуваних задач [28, 29].

Протягом виконання програмного продукту використовувалася система контролю версій GIT. Система контролю версій (Version Control System, VCS) являє собою програмне забезпечення, яке дозволяє відстежувати зміни в документах, при необхідності проводити їх відкат, визначати, хто і коли виконав виправлення і т.д. [30]. Головна відмінність Git'a від будь-яких інших СКВ (наприклад, Subversion

і їй подібних) – це те, як Git розглядає дані. Більшість інших систем зберігає інформацію як список змін для файлів. Ці системи (CVS, Subversion, Perforce, Bazaar і інші) відносяться до збережених даних як до набору файлів і змін, зроблених для кожного з цих файлів у часі. Git не зберігає свої дані в такому вигляді. Замість цього він вважає збережені дані набором зліпків невеликої файлової системи. Кожен раз, коли відбувається фіксація поточної версії проекту, Git, зберігає зліпок того, як виглядають всі файли проекту на поточний момент. Заради ефективності, якщо файл не змінювався, він не зберігається файл знову, а робить посилання на раніше збережений файл [31].

У додатку А міститься розділ з плануваннями робіт.

У результаті було детально описано постановку задачі та методи дослідження. Також було аргументовано вибір засобів реалізації інформаційної технології та виконано планування робіт.

3 МОДЕЛЮВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

3.1 Побудова діаграм в нотації сімейства IDEF

3.1.1 Діаграма у нотації IDEF0

Для формалізації функцій, які повинні виконуватися при обробці інформації у рамках інформаційної технології виконаємо структурно-функціональне моделювання в нотації IDEF0. Особливістю такої діаграми є її акцент на ієрархічне представлення об'єктів, що значно полегшує розуміння предметної області. В IDEF0 розглядаються логічні зв'язки між роботами, а не послідовність їх виконання в часі [32]. На рисунках 3.1 – 3.2 зображені контекстна діаграма у нотації IDEF0 та декомпозиція 1-го рівня моделювання процесу оцінювання ризику захворюваності на цервікальний рак.

Даний процес моделювання складається з п'яти основних робіт:

- завантаження web-додатку;
- реєстрація користувача;
- створення запиту для оцінювання ризику;
- обробка запиту для оцінювання ризику;
- виведення даних на інтерфейс web-додатку.

У якості управління описаними роботами є:

- інструкція користувача;
- нормативна документація.

Блок «Механізм» включає в себе:

- програмне забезпечення;
- апаратне забезпечення;
- користувача;

- базу даних.

Вхідними даними є індивідуальні показники жінки щодо стану її організму.
 Вихідними даними є результат оцінювання виникнення ризику захворюваності на цервікальний рак.

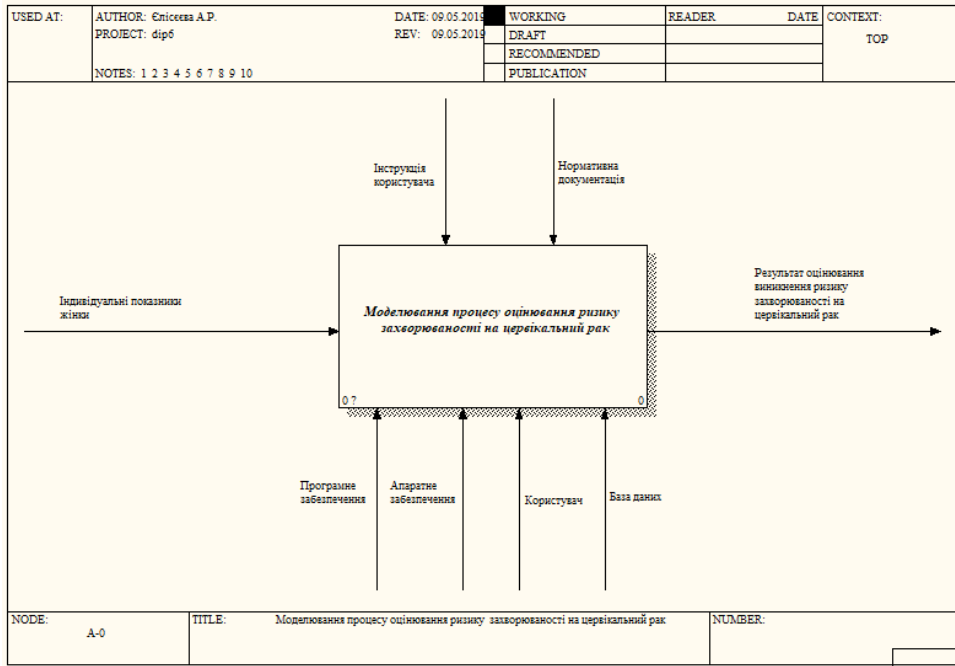


Рисунок 3.1 – Контекстна діаграма у нотації IDEF0

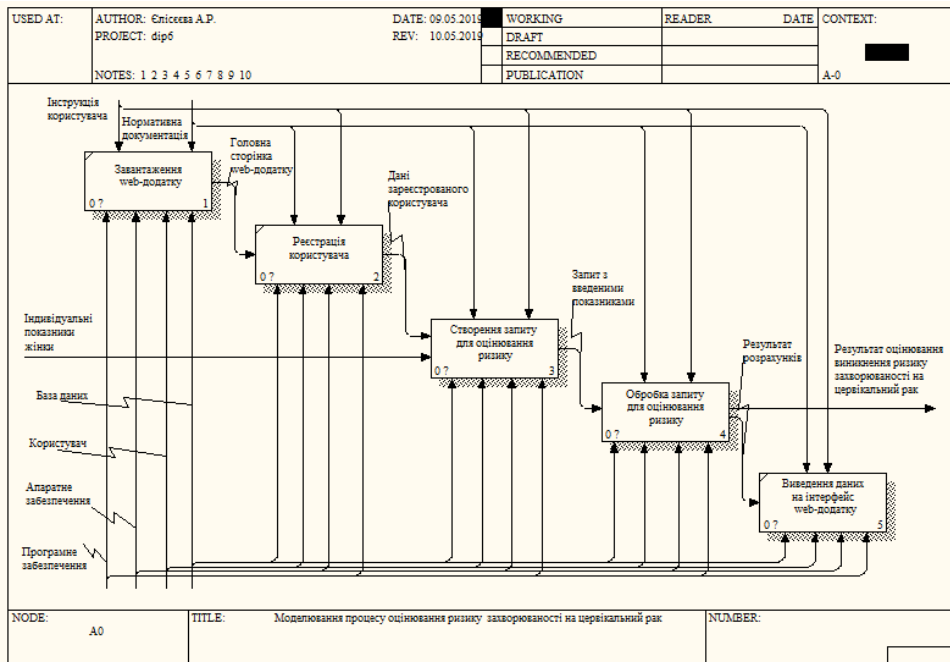


Рисунок 3.2 – Діаграма декомпозиції 1-го рівня у нотації IDEF0

3.1.2 Діаграма у нотації DFD

DFD (data flow diagrams) – діаграма потоків даних. Вона відображає структури системи, у якій зв'язками є інформаційні потоки. Діаграми потоків даних (DFD) є основним засобом моделювання функціональних вимог проектованої системи. З їх допомогою ці вимоги розбиваються на функціональні компоненти (процеси) і представляються у вигляді мережі, пов'язаної потоками даних. Головна мета таких засобів - продемонструвати, як кожен процес перетворить свої вхідні дані у вихідні, а також виявити відносини між цими процесами [33].

Основними елементами діаграми потоків даних DFD є:

- функції обробки інформації (роботи);
- документи (стрілки, arrow), об'єкти, співробітників або відділи, які беруть участь в обробці інформації;
- зовнішні посилання (external reference), які забезпечують інтерфейс з зовнішніми об'єктами, що знаходяться за межами модельованої системи;
- таблиці для зберігання документів (сховища даних, data store) [34].

На рисунку 3.3 зображена контекстна діаграма у нотації DFD для процесу створення запиту для оцінювання ризику. Вона включає наступні роботи:

- введення даних користувачем;
- перевірка введених даних на валідність;
- внесення даних у базу даних;
- запуск обробки запиту.

Сховища даних:

- ID користувача;
- індивідуальні показники користувача;
- масив індивідуальних показників користувача у базі даних.

Зовнішніми сутностями є:

- запит користувача.

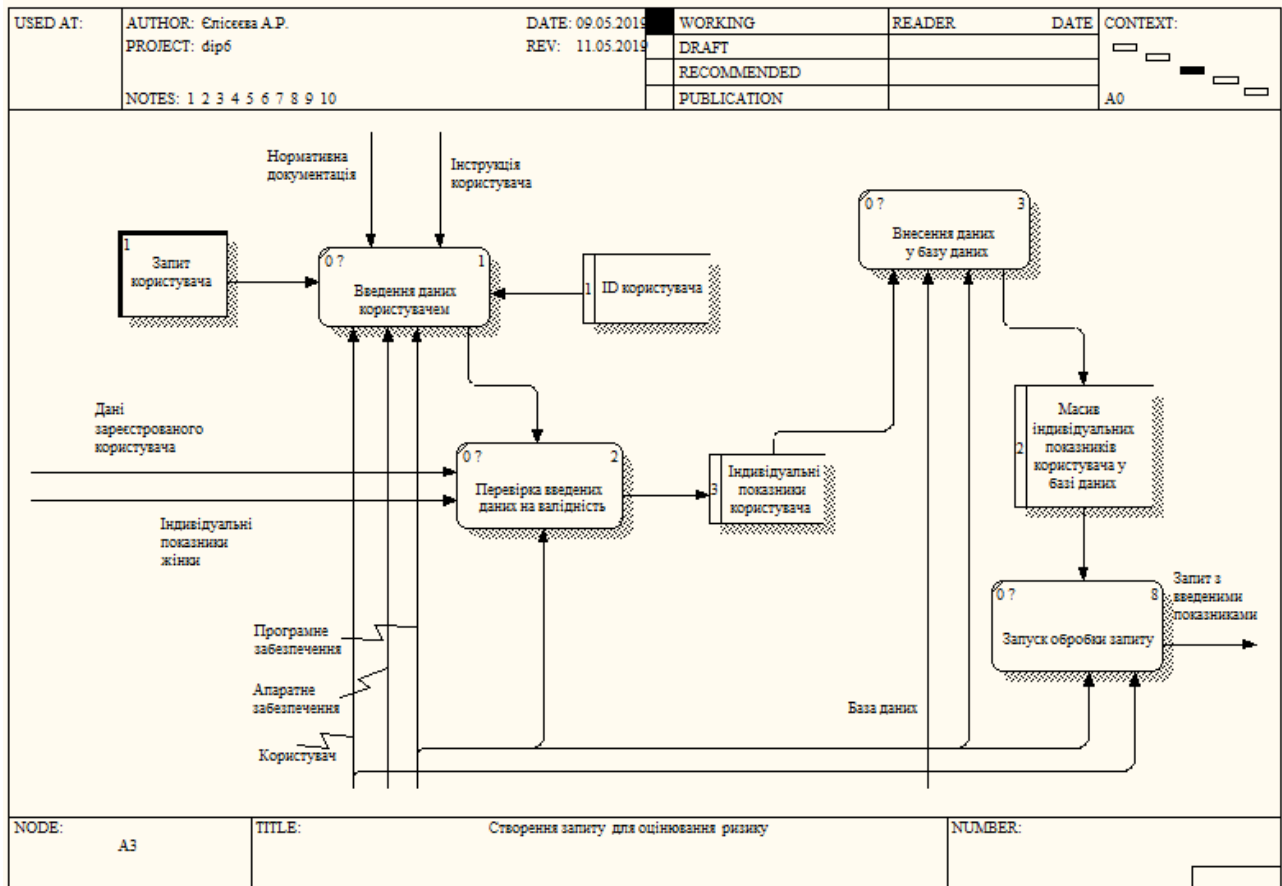


Рисунок 3.3 – Контекстна діаграма у нотації DFD

3.1.3 Діаграма у нотації IDEF3

Детально опишемо процеси, які відбуваються при обробці запиту для оцінювання ризику. Для цього виконаємо декомпозицію блоку «Обробка запиту для оцінювання ризику», побудувавши діаграму IDEF3 (рис. 3.4).

IDEF3 (англ. Integrated DEFinition for Process Description Capture Method) – методологія моделювання та стандарт документування процесів, що відбуваються в системі. Метод документування технологічних процесів є механізмом документування та збору інформації про процеси. IDEF3 показує причинно-наслідкові зв'язки між ситуаціями та подіями в зрозумілій експерту формі, використовуючи структурний метод вираження знань про те, як функціонує система, процес або підприємство [35].

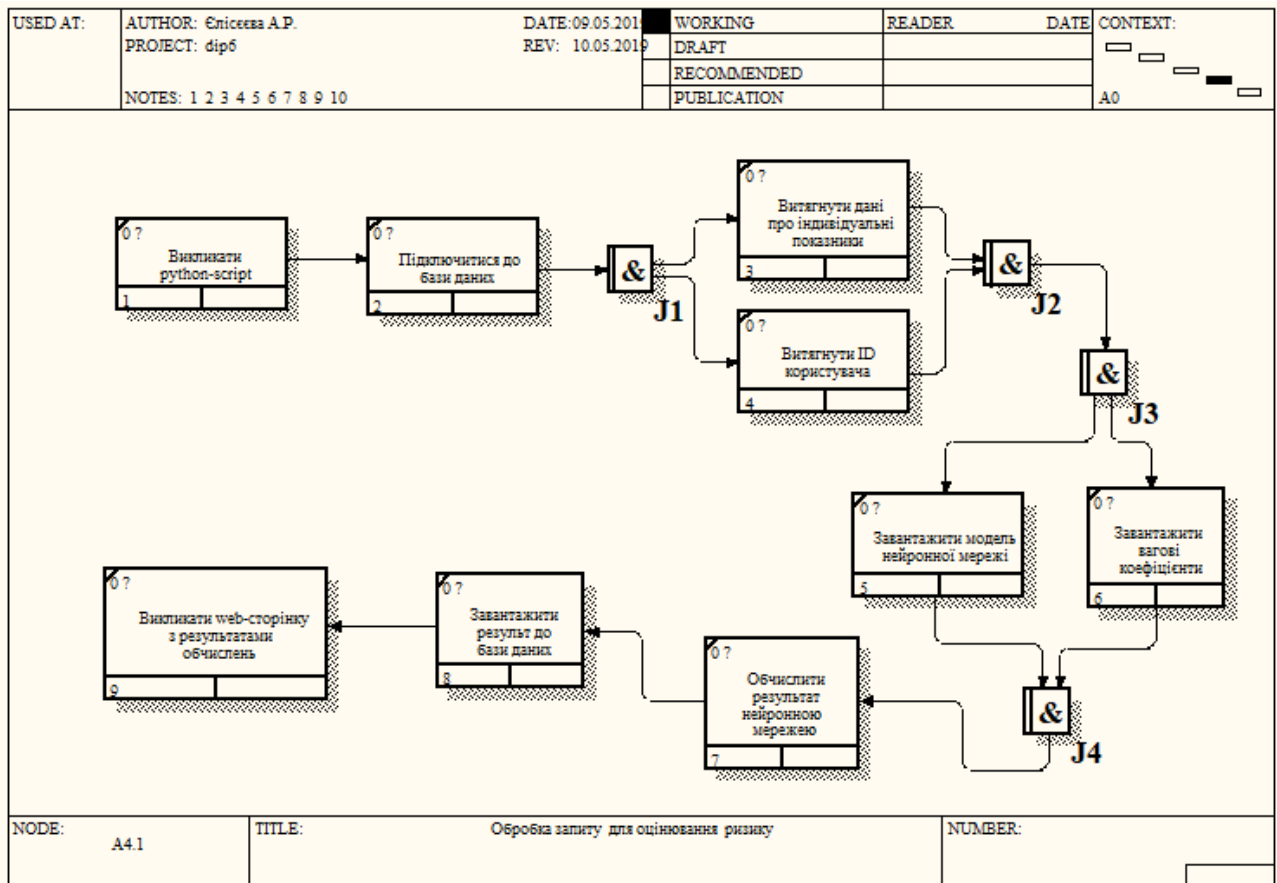


Рисунок 3.4 – Діаграма декомпозиції у нотації IDEF3

3.2 Архітектура інформаційної системи

Архітектура системи – принципова організація системи, втілена в її елементах, їх взаємозв'язків один з одним і з середовищем, а також принципи, що направляють її проектування і еволюцію [36].

На рис. 3.5 представлена архітектура розроблюваної інформаційної системи.

Архітектура інформаційної системи включає в себе web-додаток, створений за допомогою технології ASP NET Core; підключену до нього базу даних PostgreSQL та зовнішні Python скрипти, за допомогою яких відбувається обробка вхідних та вихідних масивів даних, а також створена нейронна мережа.



Рисунок 3.5 – Архітектура інформаційної системи

3.3 Побудова діаграми варіантів використання

Для того, щоб описати яким чином відбувається взаємодія об'єктів з інформаційною технологією, спроектуємо діаграму варіантів використання.

Діаграма прецедентів (діаграма варіантів використання) — діаграма, на якій зображено відношення між акторами та прецедентами в системі (рис. 3.6) [37].

Діаграма прецедентів є графом, що складається з множини акторів, прецедентів (варіантів використання) обмежених границею системи (прямокутник), асоціацій між акторами та прецедентами, відношень серед прецедентів, та відношень узагальнення між акторами. Діаграми прецедентів відображають елементи моделі варіантів використання [37].

Суть даної діаграми полягає в наступному: проєктована система представляється у вигляді безлічі сутностей чи акторів, що взаємодіють із системою за допомогою так званих варіантів використання. Варіант використання (англ. use case) використовують для описання послуг, які система надає актору. Іншими словами, кожен варіант використання визначає деякий набір дій, який виконує система при діалозі з актором. При цьому нічого не говориться про те, яким чином буде реалізована взаємодія акторів із системою [37].



Рисунок 3.6 – Діаграма варіантів використання

3.4 Побудова діаграми послідовності

Встановимо як взаємодіють між собою об'єкти в деякий момент часу. Для цього побудуємо діаграму послідовності, яка буде включати в себе об'єкти та основні прецеденти діаграми використання (рис. 3.7).

Діаграма послідовності (англ. Sequence diagram) – діаграма, на якій для деякого набору об'єктів на єдиній тимчасовій осі показаний життєвий цикл будь-якого певного об'єкта та взаємодія акторів (дійових осіб) інформаційної системи в рамках певного прецеденту [38, 39].

Основними елементами діаграми послідовності є позначення об'єктів (прямокутники з назвами об'єктів), вертикальні «лінії життя», що відображають плин часу, прямокутники, що відображають діяльність об'єкта або виконання ним певної функції (прямокутники на пунктирній «лінії життя»), і стрілки, що показують обмін сигналами або повідомленнями між об'єктами. На даній діаграмі об'єкти розташовуються зліва направо [40].

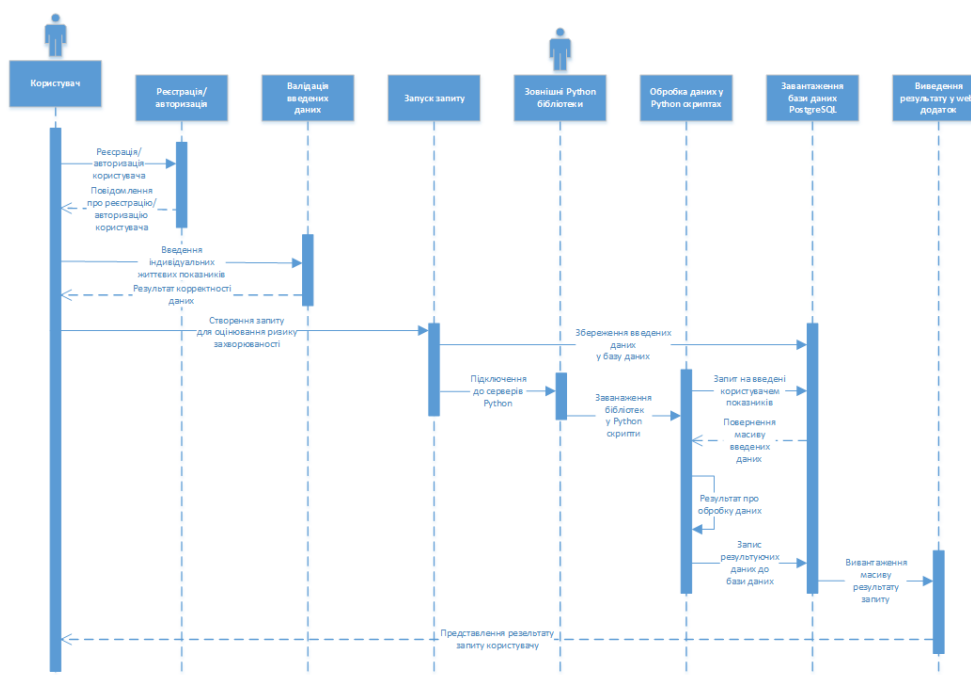


Рисунок 3.7 – Діаграма послідовності

3.5 Розробка алгоритму створення неронної мережі

3.5.1 Ціль класифікації

Задача класифікації представляє собою завдання віднесення зразка до одного з декількох множин, які попарно не перетинаються. Одним з прикладів таких завдань можуть бути медичні завдання, в яких необхідно визначити результат захворювання. При вирішенні задач класифікації необхідно віднести наявні статичні зразки до певних класів. Для медичних завдань в якості компонентів цього вектора можуть бути дані з медичної карти хворого. Таким чином, на підставі деякої інформації про зразок, необхідно визначити, до якого класу його можна віднести.

Мережі з прямим зв'язком є універсальним засобом апроксимації функцій, що дозволяє їх використовувати в рішенні задач класифікації. Як правило, нейронні мережі є найбільш ефективним способом класифікації, тому що генерують фактично велику кількість регресійних моделей (які використовуються в рішенні задач класифікації статистичними методами) [17].

3.5.2 Підготовка початкових даних

Для побудови класифікатора необхідно визначити, які параметри впливають на прийняття рішення до якого класу належить зразок. При цьому можуть виникнути дві проблеми. По-перше, якщо кількість параметрів недостатня, то може виникнути ситуація, при якій один і той же набір вхідних даних відповідає зразкам, які перебувають у різних класах. У цьому випадку навчити нейронну мережу неможливо і система не буде коректно працювати. Вхідні дані не повинні суперечити один одному. Для вирішення цієї проблеми необхідно збільшити розмірність простору ознак (кількість компонентів вхідного вектора, відповідного зразку). Але при збільшенні розмірності простору ознак може виникнути ситуація, коли число зразків може стати недостатнім для навчання мережі і вона замість узагальнення лише запам'ятає ці зразки з навчальної вибірки, в результаті чого не

зможе правильно функціонувати. Таким чином, при визначенні ознак необхідно знайти їх оптимальну кількість [17].

Далі необхідно визначити спосіб представлення вхідних даних для нейронної мережі, тобто визначити спосіб нормування. Нормування даних необхідне, оскільки нейронні мережі працюють з даними, представленими числами в діапазоні 0..1, а вхідні дані можуть мати довільний діапазон або взагалі бути нечисловими даними. При цьому можливі різні способи їх перетворення, починаючи від простого лінійного перетворення в необхідний діапазон і закінчуючи багатовимірним аналізом параметрів з нелінійним нормуванням в залежності від впливу параметрів один на одного [17].

3.5.3 Кодування вихідних значень

Завдання класифікації при наявності двох класів може бути вирішена за допомогою мережі з одним нейроном у вихідному шарі, який може приймати одне з двох значень 0 або 1, залежно від того, до якого класу належить зразок. При наявності декількох класів виникає проблема, пов'язана з поданням цих даних у вихідному шарі мережі. Найбільш простим способом представлення вихідних даних в такому випадку є вектор, компоненти якого відповідають різним номерам класів [17].

3.5.4 Вибір об'єму мережі

Існують два основні підходи до побудови мережі – конструктивний і деструктивний. При першому з них спочатку будується мережа мінімального розміру, поступово збільшуючи її для досягнення необхідної точності. При цьому на кожному кроці її перенавчають. Також існує так званий метод каскадної кореляції, при якому після закінчення епохи відбувається коригування архітектури мережі з метою мінімізації помилки. При деструктивному підході спочатку будується мережа великого розміру, потім з неї видаляються вузли та зв'язки, які майже не впливають на рішення. При цьому число зразків в навчальній множині має бути більшою за

кількість налаштованих ваг. Інакше замість узагальнення мережа тільки запам'ятає дані і втратить здатність до класифікації – результат буде невизначений для зразків, які не ввійшли в навчальну вибірку [17].

3.5.5 Вибір архітектури мережі

При виборі архітектури мережі зазвичай випробовується кілька конфігурацій з різною кількістю елементів. При цьому основним показником є обсяг навчальної множини та узагальнююча здатність мережі. Зазвичай використовується алгоритм навчання Back Propagation (зворотного поширення) з підтверджуючою безліччю [17].

3.5.6 Алгоритм побудови класифікатора на основі нейронних мереж

Основною задачею, яку потрібно вирішити за допомогою нейронної мережі є задача класифікації, тобто результатом оцінювання ризику виникнення захворюваності цервікальним раком є $\{0\}$ (негативний), або - $\{1\}$ (позитивний) результат. Для коректної роботи нейромережі потрібно побудувати алгоритм побудови класифікатора. Він буде включати в себе чотири основні процеси:

1. Робота з даними:

- a. створити базу даних із прикладів, характерних для даного завдання;
- b. розбити всю сукупність даних на дві вибірки: навчальну та тестову.

2. Попередня обробка:

- a. вибрати систему ознак, характерних для даного завдання, і перетворити дані відповідним чином для подачі на вхід мережі (нормування, стандартизація і т.д.);
- b. вибрати систему кодування вихідних значень.

3. Конструювання, навчання і оцінка якості мережі:

- a. обрати топологію мережі: кількість шарів, число нейронів в шарах і т.д.;
- b. вибрати функцію активації нейронів;
- c. обрати алгоритм навчання мережі;
- d. оцінити якість роботи мережі на основі підтверджуючої вибірки (частина навчальної вибірки) або іншим критерієм, оптимізувати архітектуру (зменшення ваг, проріджування простору ознак);
- e. обрати варіант мережі, який забезпечує найкращу здатність до узагальнення і оцінити якість роботи по тестовій вибірці.

4. Використання і діагностика:

- a. з'ясувати ступінь впливу різних чинників на прийняте рішення (евристичний підхід);
- b. переконатися, що мережа дає необхідну точність класифікації (кількість неправильно розпізнаних зразків невелика);
- c. при необхідності повернутися на етап 2, змінивши спосіб представлення зразків або змінивши базу даних;
- d. практично використовувати мережу для вирішення завдання [17].

3.6 Концептуальне моделювання

Концептуальне, чи інфологічне проектування – це процес створення моделі використовуваної на підприємстві інформації, що не залежить від будь-яких фізичних аспектів її представлення [41]. Інфологічне проектування має своїм результатом одержання семантичних моделей, що відбивають інформаційний зміст конкретної предметної області, тобто концептуальної моделі даних для аналізованої частини підприємства [42].

Дана модель створюється на основі аналізу інформації, записаної в специфікації користувача, і сформованої в ході підготовки до проектування користувальницької моделі бази даних. Ця модель абсолютно не залежить від логічних і фізичних подробиць реалізації бази даних. На цьому етапі виконується сприйняття реальної дійсності, абстрагування, вивчення й опис предметної області. У результаті цього визначаються об'єкти, властивості і зв'язки об'єктів, які мають відбиватися при проектуванні бази даних [43].

Розрізняють два головних підходи до моделювання даних при концептуальному проектуванні:

- семантичні моделі;
- об'єктні моделі [44].

Семантичні моделі головну увагу приділяють структурі даних. Найбільш поширеною семантичною моделлю є модель "сутність – зв'язок" (Entity Relationship model, ER-модель). ER-модель складається із сутностей, зв'язків, атрибутів, доменів атрибутів, ключів. Моделювання даних відображає логічну структуру даних, так само, як блок-схеми алгоритмів відображають логічну структуру програми [44].

Об'єктні моделі головну увагу приділяють поведінці об'єктів даних і засобам маніпуляції даними. Головне поняття таких моделей – об'єкт, тобто сутність, яка має стан і поведінку. Стан об'єкта визначається сукупністю його атрибутів, а поведінка об'єкта визначається сукупністю операцій специфікованих для нього [44].

Зближення цих моделей реалізується в розширеному ER-моделюванні (Extended Entity Relationship model, EER-модель) [45].

Спочатку створюємо концептуальну модель об'єктів (табл. 3.1), потім додаємо атрибути (табл. 3.2 – 3.7). Далі створюємо схему концептуальної моделі бази даних (рис. 3.8).

3.6.1 Опис сутностей

Таблиця 3.1 – об'єктів бази даних

Назва об'єкту	Характеристика об'єкта
Користувач (AspNetUsers)	Дані про користувача
Індивідуальні життєві показники користувача (DatasetModels)	Дані про індивідуальні життєві показники користувача
Результат біопсії (PredictionBiopsyModels)	Результат оцінювання ризику виникнення захворюваності на цервікальний рак – біопсія
Навчальна вибірка (TrainDatasetModels)	Дані про індивідуальні життєві показники користувачів навчальної вибірки для нейронної мережі
Тестова вибірка (TestDatasetModels)	Дані про індивідуальні життєві показники користувачів тестової вибірки для нейронної мережі
Міграції (__EFMigrationsHistory)	Міграції бази даних

Таблиця 3.2 – Таблиця атрибутів сутності «AspNetUsers»

Назва атрибута	Характеристика атрибута
Id	Атрибут, який виступає первинним ключем відношення AspNetUsers і позначає його порядковий номер
UserName	Ім'я користувача
NormalizedUserName	Нормалізоване ім'я користувача

Продовження таблиці 3.2

Назва атрибута	Характеристика атрибута
Email	Електронна адреса
NormalizedEmail	Нормалізована електронна адреса
EmailConfirmed	Підтвердження електронної адреси
PasswordHash	Захешований пароль
SecurityStamp	Печатка безпеки
ConcurrencyStamp	Штамп, який використовується для ідентифікації поточних версій даних.
PhoneNumber	Номер мобільного телефону
PhoneNumberConfirmed	Підтвердження мобільного телефону
TwoFactorEnabled	Двофакторна перевірка автентичності за допомогою SMS
LockoutEnd	Параметр ввімкнення блокування користувача протягом певного періоду часу
LockoutEnabled	Властивість, яка вказує чи заблокований користувач
AccessFailedCount	Число спроб невдалого входу в систему

Таблиця 3.3 – Таблиця атрибутів сутності «DatasetModels»

Назва атрибута	Характеристика атрибута
IDDataset	Атрибут, який виступає первинним ключем відношення WP і позначає його порядковий номер. Зв'язує таблиці WP і Configurations
Age	Вік користувача
NumberOfSexualPartners	Кількість статевих партнерів
AgeOfFirstSexualIntercourse	Вік першого статевого контакту
NumberOfPregnancies	Кількість вагітностей
Smokes	Параметр, який визначає чи палить користувач
AgeSmokes	Параметр, який визначає скільки років палить користувач

Продовження таблиці 3.3

Назва атрибута	Характеристика атрибута
PacksYearSmokes	Параметр, який визначає кількість пачок випалених цигарок на рік
UsingHormonalContraceptives	Параметр, який визначає чи вживав користувач гормональні контрацептиви
AgeUsingHormonalContraceptives	Параметр, який визначає тривалість у роках вживання гормональних контрацептивів
IUD	Параметр, який визначає використання внутрішньоматкової спіралі
AgeIUD	Параметр, який визначає кількість років використання внутрішньоматкової спіралі
STDs	Параметр, який визначає чи хворів користувач захворюваннями, що передаються статевим шляхом
NumberOfSTDs	Параметр, який визначає кількість захворювань, якими хворів користувач, що передаються статевим шляхом
STDs_Condylomatosis	Параметр, який визначає чи хворів користувач на кондиломатоз
STDs_CervicalCondylomatosis	Параметр, який визначає чи хворів користувач на цервікальний кондиломатоз
STDs_VaginalCondylomatosis	Параметр, який визначає чи хворів користувач на вагінальний кондиломатоз
STDs_VulvoPerinealCondylomatosis	Параметр, який визначає чи хворів користувач на вульво-перинеальний кондиломатоз
STDs_Syphilis	Параметр, який визначає чи хворів користувач сифілісом
STDs_PelvicInflammatoryDisease	Параметр, який визначає чи хворів користувач на запальні захворювання тазових органів
STDs_GenitalHerpes	Параметр, який визначає чи хворів користувач на генітальний герпес
STDs_MolluscumContagiosum	Параметр, який визначає чи хворів користувач на контагіозний молюск

Продовження таблиці 3.3

Назва атрибута	Характеристика атрибута
STDs_AIDS	Параметр, який визначає чи хворів користувач СНІДом
STDs_HIV	Параметр, який визначає чи хворів користувач ВІЛом
STDs_HepatitisB	Параметр, який визначає чи хворів користувач на гепатит Б
STDs_HPВ	Параметр, який визначає чи хворів користувач вірусом папіломи людини
STDs_NumberOfDiagnosis	Параметр, який визначає загальну кількість діагнозів
STDs_TimeSinceFirstDiagnosis	Параметр, який визначає час з моменту першого діагнозу, у роках
STDs_TimeSinceLastDiagnosis	Параметр, який визначає час, який минув від останнього діагнозу, у роках
Dx_Cancer	Параметр, який визначає чи був вже раніше продіагностований цервікальний рак
Dx_CIN	Параметр, який визначає чи є у користувача дисплазія шийки матки
Dx_HPВ	Параметр, який визначає чи був продіагностований вірус папіломи людини
Dx	Параметр, який визначає чи був раніше продіагностований рак
HinselmannTest	Параметр, який визначає результати кольпоскопії Хінсельмана
SchillerTest	Параметр, який визначає результати тесту Шилера
Citology	Параметр, який визначає результати цитології
FKIDUsers	Атрибут, який виступає зовнішнім ключем відношення DatasetModels зв'язує таблиці DatasetModels і AspNetUsers

Таблиця 3.4 – Таблиця атрибутів сутності «PredictionBiopsyModels»

Назва атрибута	Характеристика атрибута
IDPredictionBiopsy	Атрибут, який виступає первинним ключем відношення PredictionBiopsyModels і позначає його порядковий номер
PredictionBiopsy	Результат оцінювання ризику захворюваності
DateTime	Дата та час, коли був виконаний запит
FKIDUsers	Атрибут, який виступає зовнішнім ключем відношення PredictionBiopsyModels зв'язує таблиці PredictionBiopsyModels і AspNetUsers

Таблиця 3.5 – Таблиця атрибутів сутності «__EFMigrationsHistory»

Назва атрибута	Характеристика атрибута
MigrationId	Атрибут, який виступає первинним ключем відношення __EFMigrationsHistory і позначає його порядковий номер
ProductVersion	Версія міграції

Таблиця 3.6 – Таблиця атрибутів сутності «TestDatasetModels»

Назва атрибута	Характеристика атрибута
TestDatasetModels	Атрибут, який виступає первинним ключем відношення TestDatasetModels і позначає його порядковий номер
Age	Вік користувача
NumberOfSexualPartners	Кількість статевих партнерів
AgeOfFirstSexualIntercourse	Вік першого статевого контакту
NumberOfPregnancies	Кількість вагітностей
AgeSmokes	Параметр, який визначає скільки років палить користувач
PacksYearSmokes	Параметр, який визначає кількість пачок випалених цигарок на рік
AgeUsingHormonalContraceptives	Параметр, який визначає тривалість у роках вживання гормональних контрацептивів

Продовження таблиці 3.6

Назва атрибута	Характеристика атрибута
AgeIUD	Параметр, який визначає кількість років використання внутрішньоматкової спіралі
NumberOfSTDs	Параметр, який визначає кількість захворювань, якими хворів користувач, що передаються статевим шляхом
STDs_Condylomatosis	Параметр, який визначає чи хворів користувач на кондиломатоз
STDs_CervicalCondylomatosis	Параметр, який визначає чи хворів користувач на цервікальний кондиломатоз
STDs_VaginalCondylomatosis	Параметр, який визначає чи хворів користувач на вагінальний кондиломатоз
STDs_VulvoPerinealCondylomatosis	Параметр, який визначає чи хворів користувач на вульво-перинеальний кондиломатоз
STDs_Syphilis	Параметр, який визначає чи хворів користувач сифілісом
STDs_PelvicInflammatoryDisease	Параметр, який визначає чи хворів користувач на запальні захворювання тазових органів
STDs_GenitalHerpes	Параметр, який визначає чи хворів користувач на генітальний герпес
STDs_MolluscumContagiosum	Параметр, який визначає чи хворів користувач на контагіозний молюск
STDs_AIDS	Параметр, який визначає чи хворів користувач СНІДом
STDs_HIV	Параметр, який визначає чи хворів користувач ВІЛом
STDs_HepatitisB	Параметр, який визначає чи хворів користувач на гепатит Б
STDs_HPВ	Параметр, який визначає чи хворів користувач вірусом папіломи людини
STDs_NumberOfDiagnosis	Параметр, який визначає загальну кількість діагнозів
STDs_TimeSinceFirstDiagnosis	Параметр, який визначає час з моменту першого діагнозу, у роках

Продовження таблиці 3.6

Назва атрибута	Характеристика атрибута
STDs_TimeSinceLastDiagnosis	Параметр, який визначає час, який минув від останнього діагнозу, у роках
Biopsy	Параметр, який визначає чи є у користувача рак шийки матки
Smokes_0	Параметр, який визначає що користувач не палить (його значення 1, значення Smokes_1 – 0)
Smokes_1	Параметр, який визначає що користувач палить (його значення 1, значення Smokes_0 – 0)
UsingHormonalContraceptives_0	Параметр, який визначає що користувач не вживав гормональні контрацептиви (його значення 1, значення UsingHormonalContraceptives_1 – 0)
UsingHormonalContraceptives_1	Параметр, який визначає що користувач вживав гормональні контрацептиви (його значення 1, значення UsingHormonalContraceptives_0 – 0)
IUD_0	Параметр, який визначає використання внутрішньоматкової спіралі (його значення 1, значення IUD_1 – 0)
IUD_1	Параметр, який визначає використання внутрішньоматкової спіралі (його значення 1, значення IUD_0 – 0)
STDs_0	Параметр, який визначає чи хворів користувач захворюваннями, що передаються статевим шляхом (його значення 1, значення STDs_1 – 0)
STDs_1	Параметр, який визначає чи хворів користувач захворюваннями, що передаються статевим шляхом (його значення 1, значення STDs_0 – 0)
Dx_Cancer_0	Параметр, який визначає чи був вже раніше продіагнований цервікальний рак (його значення 1, значення Dx_Cancer_1 – 0)

Продовження таблиці 3.6

Назва атрибута	Характеристика атрибута
Dx_Cancer_1	Параметр, який визначає чи був вже раніше продіагностований цервікальний рак (його значення 1, значення Dx_Cancer_0 – 0)
Dx_CIN_0	Параметр, який визначає чи є у користувача дисплазія шийки матки (його значення 1, значення Dx_CIN_1 – 0)
Dx_CIN_1	Параметр, який визначає чи є у користувача дисплазія шийки матки (його значення 1, значення Dx_CIN_0 – 0)
Dx_HPВ_0	Параметр, який визначає чи був продіагностований вірус папіломи людини (його значення 1, значення Dx_HPВ_1 – 0)
Dx_HPВ_1	Параметр, який визначає чи був продіагностований вірус папіломи людини (його значення 1, значення Dx_HPВ_0 – 0)
Dx_0	Параметр, який визначає чи був раніше продіагностований рак (його значення 1, значення Dx_1 – 0)
Dx_1	Параметр, який визначає чи був раніше продіагностований рак (його значення 1, значення Dx_0 – 0)
HinselmannTest_0	Параметр, який визначає результати кольпоскопії Хінсельмана (його значення 1, значення HinselmannTest_1 – 0)
HinselmannTest_1	Параметр, який визначає результати кольпоскопії Хінсельмана (його значення 1, значення HinselmannTest_0 – 0)
SchillerTest_0	Параметр, який визначає результати тесту Шилера (його значення 1, значення SchillerTest_1 – 0)
SchillerTest_1	Параметр, який визначає результати тесту Шилера (його значення 1, значення SchillerTest_0 – 0)
Citology_0	Параметр, який визначає результати цитології (його значення 1, значення Citology_1 – 0)

Продовження таблиці 3.6

Назва атрибута	Характеристика атрибута
Citology_1	Параметр, який визначає результати цитології (його значення 1, значення Citology_0 – 0)

Таблиця 3.7 – Таблиця атрибутів сутності «TrainDatasetModels»

Назва атрибута	Характеристика атрибута
IDTrainDataset	Атрибут, який виступає первинним ключем відношення TrainDatasetModels і позначає його порядковий номер
Age	Вік користувача
NumberOfSexualPartners	Кількість статевих партнерів
AgeOfFirstSexualIntercourse	Вік першого статевого контакту
NumberOfPregnancies	Кількість вагітностей
AgeSmokes	Параметр, який визначає скільки років палить користувач
PacksYearSmokes	Параметр, який визначає кількість пачок випалених цигарок на рік
AgeUsingHormonalContraceptives	Параметр, який визначає тривалість у роках вживання гормональних контрацептивів
AgeIUD	Параметр, який визначає кількість років використання внутрішньоматкової спіралі
NumberOfSTDs	Параметр, який визначає кількість захворювань, якими хворів користувач, що передаються статевим шляхом
STDs_Condylomatosis	Параметр, який визначає чи хворів користувач на конділоматоз
STDs_CervicalCondylomatosis	Параметр, який визначає чи хворів користувач на цервікальний конділоматоз
STDs_VaginalCondylomatosis	Параметр, який визначає чи хворів користувач на вагінальний конділоматоз
STDs_VulvoPerinealCondylomatosis	Параметр, який визначає чи хворів користувач на вульво-перинеальний конділоматоз

Продовження таблиці 3.7

Назва атрибута	Характеристика атрибута
STDs_Syphilis	Параметр, який визначає чи хворів користувач сифілісом
STDs_PelvicInflammatoryDisease	Параметр, який визначає чи хворів користувач на запальні захворювання тазових органів
STDs_GenitalHerpes	Параметр, який визначає чи хворів користувач на генітальний герпес
STDs_MolluscumContagiosum	Параметр, який визначає чи хворів користувач на контагіозний молюск
STDs_AIDS	Параметр, який визначає чи хворів користувач СНІДом
STDs_HepatitisB	Параметр, який визначає чи хворів користувач на гепатит Б
STDs_HPВ	Параметр, який визначає чи хворів користувач вірусом папіломи людини
STDs_NumberOfDiagnosis	Параметр, який визначає загальну кількість діагнозів
STDs_TimeSinceFirstDiagnosis	Параметр, який визначає час з моменту першого діагнозу, у роках
STDs_TimeSinceLastDiagnosis	Параметр, який визначає час, який минув від останнього діагнозу, у роках
Biopsy	Параметр, який визначає чи є у користувача рак шийки матки
Smokes_0	Параметр, який визначає що користувач не палить (його значення 1, значення Smokes_1 – 0)
Smokes_1	Параметр, який визначає що користувач палить (його значення 1, значення Smokes_0 – 0)
UsingHormonalContraceptives_0	Параметр, який визначає що користувач не вживав гормональні контрацептиви (його значення 1, значення UsingHormonalContraceptives_1 – 0)

Продовження таблиці 3.7

Назва атрибута	Характеристика атрибута
UsingHormonalContraceptives_1	Параметр, який визначає що користувач вживав гормональні контрацептиви (його значення 1, значення UsingHormonalContraceptives_0 – 0)
IUD_0	Параметр, який визначає використання внутрішньоматкової спіралі (його значення 1, значення IUD_1 – 0)
IUD_1	Параметр, який визначає використання внутрішньоматкової спіралі (його значення 1, значення IUD_0 – 0)
STDs_0	Параметр, який визначає чи хворів користувач захворюваннями, що передаються статевим шляхом (його значення 1, значення STDs_1 – 0)
STDs_1	Параметр, який визначає чи хворів користувач захворюваннями, що передаються статевим шляхом (його значення 1, значення STDs_0 – 0)
Dx_Cancer_0	Параметр, який визначає чи був вже раніше продіагнований цервікальний рак (його значення 1, значення Dx_Cancer_1 – 0)
Dx_Cancer_1	Параметр, який визначає чи був вже раніше продіагнований цервікальний рак (його значення 1, значення Dx_Cancer_0 – 0)
Dx_CIN_0	Параметр, який визначає чи є у користувача дисплазія шийки матки (його значення 1, значення Dx_CIN_1 – 0)
Dx_CIN_1	Параметр, який визначає чи є у користувача дисплазія шийки матки (його значення 1, значення Dx_CIN_0 – 0)
Dx_HPВ_0	Параметр, який визначає чи був продіагнований вірус папіломи людини (його значення 1, значення Dx_HPВ_1 – 0)
Dx_HPВ_1	Параметр, який визначає чи був продіагнований вірус папіломи людини (його значення 1, значення Dx_HPВ_0 – 0)

Продовження таблиці 3.7

Назва атрибута	Характеристика атрибута
Dx_0	Параметр, який визначає чи був раніше продіагнований рак (його значення 1, значення Dx_1 – 0)
Dx_1	Параметр, який визначає чи був раніше продіагнований рак (його значення 1, значення Dx_0 – 0)
HinselmannTest_0	Параметр, який визначає результати кольпоскопії Хінсельмана (його значення 1, значення HinselmannTest_1 – 0)
HinselmannTest_1	Параметр, який визначає результати кольпоскопії Хінсельмана (його значення 1, значення HinselmannTest_0 – 0)
SchillerTest_0	Параметр, який визначає результати тесту Шилера (його значення 1, значення SchillerTest_1 – 0)
SchillerTest_1	Параметр, який визначає результати тесту Шилера (його значення 1, значення SchillerTest_0 – 0)
Citology_0	Параметр, який визначає результати цитології (його значення 1, значення Citology_1 – 0)
Citology_1	Параметр, який визначає результати цитології (його значення 1, значення Citology_0 – 0)

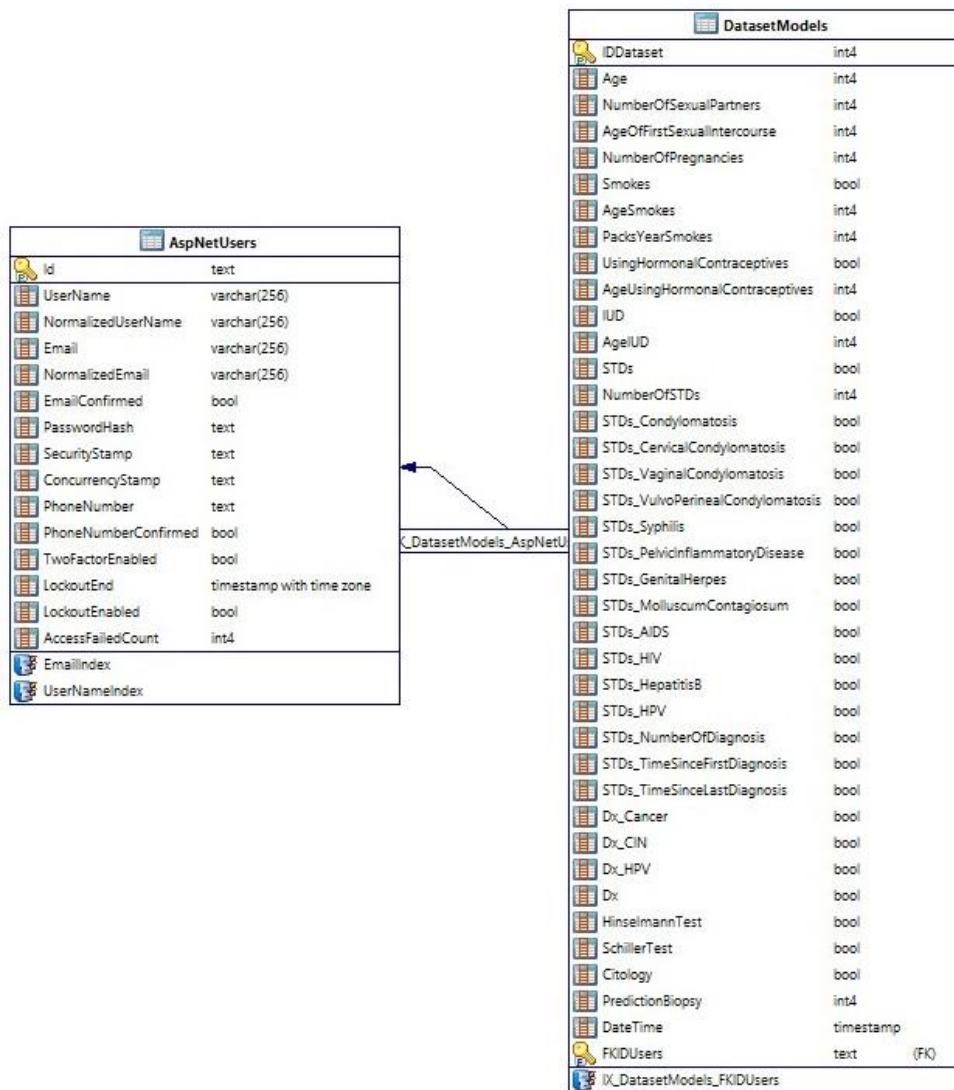


Рисунок 3.8 – Концептуальна модель бази даних

3.6.2 Зв'язки між об'єктами

Користувач (**AspNetUsers**) має набір даних, які є його індивідуальними життєвими показниками (**DatasetModels**).

Індивідуальні життєві показники (**DatasetModels**) належать одному користувачу (**AspNetUsers**).

3.7 Логічне проектування

Логічне проектування — це розроблення структур зберігання, методів доступу й логічної структури системи баз даних без прив'язки до конкретної системи управління базами даних (СУБД) [46].

Ціль етапу – створення логічної моделі даних на рівні моделі записів шляхом уточнення і перетворення концептуальної моделі з урахуванням особливостей обраної організації даних у цільовий СУБД (наприклад, реляційна чи мережева модель). Для реляційних моделей на цьому етапі фіксується набір і структура таблиць, що представляють сутності, визначаються ключі і проводиться нормалізація таблиць [47, 48].

3.7.1 Нормалізація таблиць

Проаналізувавши сутність «AspNetUsers» визначаємо, що вона не потребує нормалізації, так як перебуває в атомарному стані та всі неключові атрибути відношення залежать тільки від первинного ключа.

Виконаємо аналіз сутності «DatasetModels». Атрибути «PredictionBiopsy» та «DateTime» функціонально залежать від первинного ключа, ці дані є результуючими, тому доцільно виділити їх у окрему сутність «PredictionBiopsyModels».

Атрибути «TestDatasetModels», «TrainDatasetModels», «__EFMigrationsHistory» є додатковими і не включаються в основну схему базу даних додатку. Атрибути «TestDatasetModels», «TrainDatasetModels» призначені для збереження в них початкових навчальних та тренувальних даних для виконання навчання розроблюваної нейронної мережі. Атрибут «__EFMigrationsHistory» є окремим та використовується для збереження міграцій, які виконувалися у ході написання програмного коду. Міграції – це механізм ASP NET Core, призначений для

швидкого оновлення бази даних. Всі зміни у базі даних можна виконувати за допомогою програмного коду, а потім виконувати міграції, які автоматично оновлять структуру та властивості бази даних, відповідно до попередньо внесених змін.

На рисунку 3.9 зображена логічна модель бази даних.

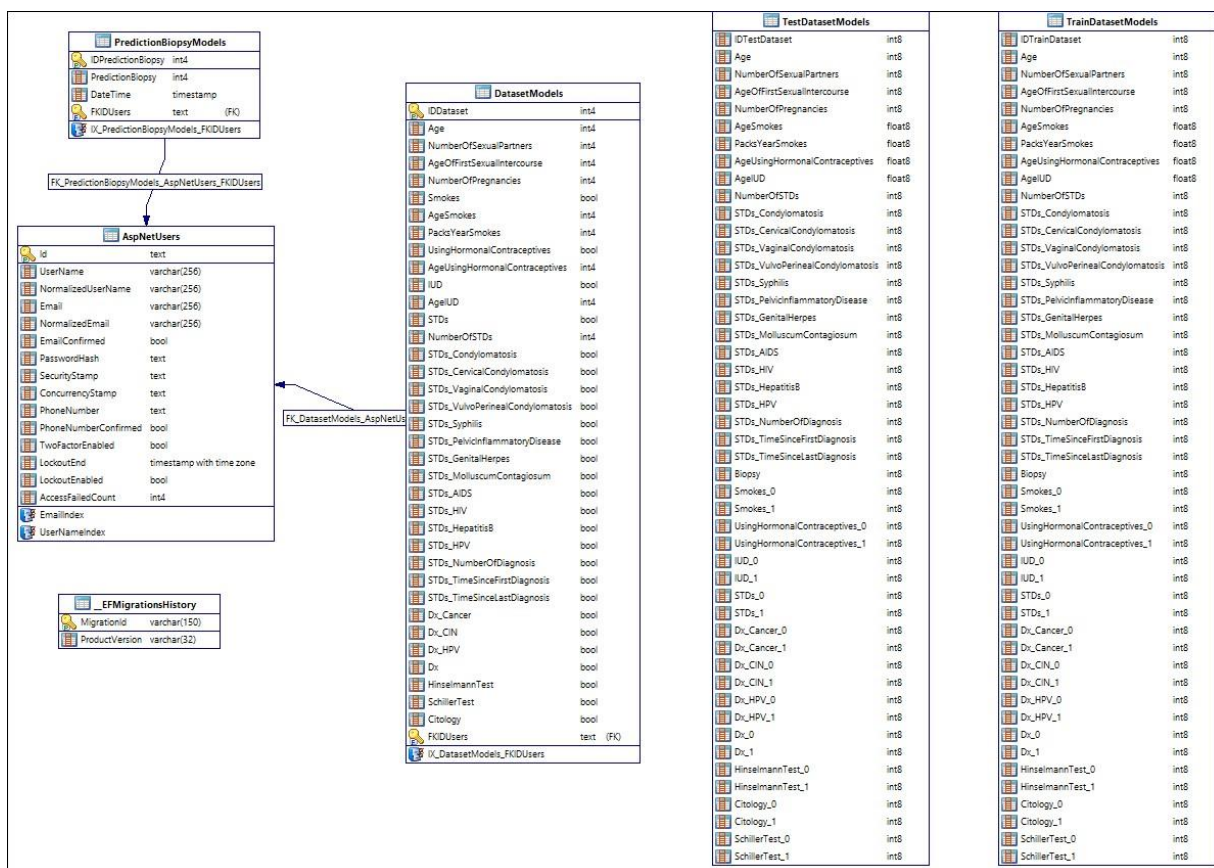


Рисунок 3.9 – Логічна модель бази даних

3.7.2 Функціональні залежності

Функціональна залежність — концепція, що лежить в основі багатьох питань, пов'язаних з реляційними базами даних, включаючи, зокрема, їхнє проектування. Математично являє собою бінарне відношення між множинами атрибутів даного відношення і є, по суті, зв'язком типу «один-до-багатьох».

Функціональні залежності відображають зв'язки між атрибутами, які властиві реальному об'єкту. Атрибут В функціонально залежить від А, якщо кожному значенню А відповідає в точності одне значення В.

Функціональні залежності розроблюваної бази даних:

- сутність «AspNetUsers»: «UserName», «NormalizedUserName», «Email», «NormalizedEmail», «EmailConfirmed», «PasswordHash», «SecurityStamp», «ConcurrencyStamp», «PhoneNumber», «PhoneNumberConfirmed», «TwoFactorEnabled», «LockoutEnd», «LockoutEnabled», «AccessFailedCount» → {«Id»};
- сутність «DatasetModels»: «Age», «NumberOfSexualPartners», «AgeOfFirstSexualIntercourse», «NumberOfPregnancies», «Smokes», «AgeSmokes», «PacksYearSmokes», «UsingHormonalContraceptives», «AgeUsingHormonalContraceptives», «IUD», «AgeIUD», «STDs», «NumberOfSTDs», «STDs_Condylomatosis», «STDs_CervicalCondylomatosis», «STDs_VaginalCondylomatosis», «STDs_VulvoPerinealCondylomatosis», «STDs_Syphilis», «STDs_PelvicInflammatoryDisease», «STDs_GenitalHerpes», «STDs_MolluscumContagiosum», «STDs_AIDS», «STDs_HIV», «STDs_HepatitisB», «STDs_HPВ», «STDs_NumberOfDiagnosis», «STDs_TimeSinceFirstDiagnosis», «STDs_TimeSinceLastDiagnosis», «Dx_Cancer», «Dx_CIN», «Dx_HPВ», «Dx», «HinselmannTest», «SchillerTest», «Citology» → {«IDDataset», «FKIDUsers»};
- сутність «PredictionBiopsyModels»: «PredictionBiopsy», «DateTime» → {«IDPredictionBiopsy», «FKIDUsers»};
- сутність «__EFMigrationsHistory»: «ProductVersion» → {«MigrationId»};
- сутність «TrainDatasetModels»: «Age», «NumberOfSexualPartners», «AgeOfFirstSexualIntercourse», «NumberOfPregnancies», «AgeSmokes», «PacksYearSmokes», «AgeUsingHormonalContraceptives», «AgeIUD», «NumberOfSTDs», «STDs_Condylomatosis»,

«STDs_CervicalCondylomatosis», «STDs_VaginalCondylomatosis»,
 «STDs_VulvoPerinealCondylomatosis», «STDs_Syphilis»,
 «STDs_PelvicInflammatoryDisease», «STDs_GenitalHerpes»,
 «STDs_MolluscumContagiosum», «STDs_AIDS», «STDs_HIV»,
 «STDs_HepatitisB», «STDs_HPВ», «STDs_NumberOfDiagnosis»,
 «STDs_TimeSinceFirstDiagnosis», «STDs_TimeSinceLastDiagnosis»,
 «Biopsy», «Smokes_0», «Smokes_1», «UsingHormonalContraceptives_0»,
 «UsingHormonalContraceptives_1», «IUD_0», «IUD_1», «STDs_0»,
 «STDs_1», «Dx_Cancer_0», «Dx_Cancer_1», «Dx_CIN_0», «Dx_CIN_1»,
 «Dx_HPВ_0», «Dx_HPВ_1», «Dx_0», «Dx_1», «HinselmannTest_0»,
 «HinselmannTest_1», «Citology_0», «Citology_1», «SchillerTest_0»,
 «SchillerTest_1» → {«IDTrainDataset»};

- **сyтнiсть «TestDatasetModels»:** «Age», «NumberOfSexualPartners»,
 «AgeOfFirstSexualIntercourse», «NumberOfPregnancies», «AgeSmokes»,
 «PacksYearSmokes», «AgeUsingHormonalContraceptives», «AgeIUD»,
 «NumberOfSTDs», «STDs_Condylomatosis»,
 «STDs_CervicalCondylomatosis», «STDs_VaginalCondylomatosis»,
 «STDs_VulvoPerinealCondylomatosis», «STDs_Syphilis»,
 «STDs_PelvicInflammatoryDisease», «STDs_GenitalHerpes»,
 «STDs_MolluscumContagiosum», «STDs_AIDS», «STDs_HIV»,
 «STDs_HepatitisB», «STDs_HPВ», «STDs_NumberOfDiagnosis»,
 «STDs_TimeSinceFirstDiagnosis», «STDs_TimeSinceLastDiagnosis»,
 «Biopsy», «Smokes_0», «Smokes_1», «UsingHormonalContraceptives_0»,
 «UsingHormonalContraceptives_1», «IUD_0», «IUD_1», «STDs_0»,
 «STDs_1», «Dx_Cancer_0», «Dx_Cancer_1», «Dx_CIN_0», «Dx_CIN_1»,
 «Dx_HPВ_0», «Dx_HPВ_1», «Dx_0», «Dx_1», «HinselmannTest_0»,
 «HinselmannTest_1», «Citology_0», «Citology_1», «SchillerTest_0»,
 «SchillerTest_1» → {«IDTestDataset»}.

3.7.3 Визначення первинних та зовнішніх ключів

Для сутності «AspNetUsers»: «Id» – первинний.

Для сутності «DatasetModels»: «IDDataset» – первинний, «FKIDUsers» – зовнішній;

Для сутності «PredictionBiopsyModels»: «IDPredictionBiopsy» – первинний, «FKIDUsers» – зовнішній;

Для сутності «__EFMigrationsHistory»: «MigrationId» – первинний.

Для сутності «TrainDatasetModels»: «IDTrainDataset» – первинний.

Для сутності «TestDatasetModels»: «IDTestDataset» – первинний.

3.8 Фізичне моделювання

Фізичне проектування — це проектування бази даних у конкретній СУБД.

На основі логічної моделі створюємо фізичну модель бази даних. Перший етап включає в себе створення базових таблиць, додавання колонок до базових таблиць, визначення типів даних для колонок відповідно дозволеним для даної СУБД типам даних, визначення первинних ключових полів таблиць, створення унікального індексу первинного ключа, визначення обмеження на дані за необхідності, створення єднальних таблиць.

3.8.1 Створення базових таблиць

На даному етапі проектування бази даних розроблюють код для створення таблиць бази даних з визначеними первинними ключами, обмеженнями та типами даних. Скрипт для створення таблиць у базі даних наведений у додатку Б.

3.8.2 Створення унікального індексу первинного ключа

Виконання створення унікального індексу первинного ключа для встановлення зав'язків між таблицями.

```
CREATE UNIQUE INDEX "AspNetUsersIndex" ON public."AspNetUsers" USING btree ("Id");

CREATE UNIQUE INDEX "IDDatasetIndex" ON public."DatasetModels" USING btree ("IDDataset");

CREATE UNIQUE INDEX "IDPredictionBiopsyIndex" ON public."PredictionBiopsyModels" USING btree ("IDPredictionBiopsy");

CREATE UNIQUE INDEX "MigrationIdIndex" ON public."__EFMigrationsHistory" USING btree ("MigrationId");

CREATE UNIQUE INDEX "IDTrainDatasetIndex" ON public."TrainDatasetModels" USING btree ("IDTrainDataset");

CREATE UNIQUE INDEX "IDTestDatasetIndex" ON public."TestDatasetModels" USING btree ("IDTestDataset");

CREATE INDEX "IX_DatasetModels_FKIDUsers" ON public."DatasetModels" USING btree ("FKIDUsers");

CREATE INDEX "IX_PredictionBiopsyModels_FKIDUsers" ON public."PredictionBiopsyModels" USING btree ("FKIDUsers");
```

3.8.3 Створення зовнішніх ключів

Написання коду для додавання зовнішніх ключів до таблиць бази даних та їх посилення на первинні ключі основних таблиць.

```
ALTER TABLE ONLY public."AspNetUsers"
  ADD CONSTRAINT "PK_AspNetUsers" PRIMARY KEY ("Id");

ALTER TABLE ONLY public."DatasetModels"
  ADD CONSTRAINT "PK_DatasetModels" PRIMARY KEY ("IDDataset");

ALTER TABLE ONLY public."PredictionBiopsyModels"
  ADD CONSTRAINT "PK_PredictionBiopsyModels" PRIMARY KEY ("IDPredictionBiopsy");

ALTER TABLE ONLY public."__EFMigrationsHistory"
  ADD CONSTRAINT "PK__EFMigrationsHistory" PRIMARY KEY ("MigrationId");

ALTER TABLE ONLY public."DatasetModels"
  ADD CONSTRAINT "FK_DatasetModels_AspNetUsers_FKIDUsers" FOREIGN KEY ("FKIDUsers")
REFERENCES public."AspNetUsers"("Id") ON DELETE RESTRICT;

ALTER TABLE ONLY public."PredictionBiopsyModels"
  ADD CONSTRAINT "FK_PredictionBiopsyModels_AspNetUsers_FKIDUsers" FOREIGN KEY ("FKIDUsers")
REFERENCES public."AspNetUsers"("Id") ON DELETE RESTRICT;
```

Результатом роботи на даному етапі роботи є моделювання роботи розроблюваної інформаційної технології. Було здійснено побудову діаграм у нотації IDEF, а саме діаграми у нотації IDEF з декомпозицією першого рівня, у нотації DFD та IDEF3. Було розроблено архітектуру інформаційної системи, побудовано діаграму варіантів використання та діаграму послідовності. Було розроблено алгоритм створення нейронної мережі. Також було описано розроблену базу даних та виконане концептуальне, логічне та фізичне моделювання.

4 РОЗРОБКА ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

4.1 Обробка даних

Для оцінки ризику виникнення захворювання цервікального раку в залежності від індивідуальних життєвих показників жінки була розроблена нейронна мережа методом багатосарового персептрону. Вхідні дані склали масив результатів аналізів 858 жінок. Даний масив вхідної інформації був отриманий з офіційного сайту «Machine Learning Repository» за посиланням <https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Factors%29#> (рис. 4.1).

UCI Machine Learning Repository

Machine Learning Repository
Center for Machine Learning and Intelligent Systems

About Citation Policy Donate a Data Set Contact

Search

Repository Web

View ALL Data Sets

Cervical cancer (Risk Factors) Data Set

Download [Data Folder](#), [Data Set Description](#)

Abstract: This dataset focuses on the prediction of indicators/diagnosis of cervical cancer. The features cover demographic information, habits, and historic medical records.

Data Set Characteristics:	Multivariate	Number of Instances:	858	Area:	Life
Attribute Characteristics:	Integer, Real	Number of Attributes:	36	Date Donated	2017-03-03
Associated Tasks:	Classification	Missing Values?	Yes	Number of Web Hits:	81808

Source:

Kelwin Fernandes (kafo_at_inesclec_dot_pt) - INESC TEC & FEUP, Porto, Portugal.
Jaime S. Cardoso - INESC TEC & FEUP, Porto, Portugal.
Jessica Fernandes - Universidad Central de Venezuela, Caracas, Venezuela.

Data Set Information:

The dataset was collected at 'Hospital Universitario de Caracas' in Caracas, Venezuela. The dataset comprises demographic information, habits, and historic medical records of 858 patients. Several patients decided not to answer some of the questions because of privacy concerns (missing values).

Attribute Information:

(int) Age
(int) Number of sexual partners
(int) First sexual intercourse (age)
(int) Num of pregnancies
Peak Smoker

Рисунок 4.1 – Сторінка сайту «Machine Learning Repository»

Умовно результат захворювання поділено на дві категорії – позитивний (1), якщо захворювання присутнє в організмі жінки та негативний (0), якщо захворювання відсутнє.

У якості вхідного шару нейронів використовувалися наступні параметри пацієнтів:

- вік жінки;
- кількість статевих партнерів;
- вік першого статевого контакту;
- кількість вагітностей;
- паління;
- кількість років паління;
- кількість пачок випалених цигарок на рік;
- вживання гормональних контрацептивів;
- тривалість у роках вживання гормональних контрацептивів;
- використання внутрішньоматкової спіралі;
- тривалість використання внутрішньоматкової спіралі;
- наявність захворювань, що передаються статевим шляхом;
- кількість захворювань, що передаються статевим шляхом;
- захворюваність на кондиломатоз;
- захворюваність на цервікальний кондиломатоз;
- захворюваність на вагінальний кондиломатоз;
- захворюваність на вульво-перинеальний кондиломатоз;
- захворюваність сифілісом;
- захворюваність на запальні захворювання тазових органів;
- захворюваність на генітальний герпес;
- захворюваність на контагіозний моллюск;
- захворюваність СНІДом;
- захворюваність ВІЛом;
- захворюваність гепатитом Б;

- захворюваність вірусом папіломи людини;
- загальна кількість діагнозів;
- час з моменту першого діагнозу, у роках;
- час, який минув від останнього діагнозу, у роках;
- ранішнє діагностування цервікального раку;
- захворюваність на дисплазію шийки матки;
- захворюваність на вірус папіломи людини;
- ранішнє діагностування іншого виду раку;
- результати кольпоскопії Хінсельмана;
- результати тесту Шилера;
- результати цитології.

Кожній ознаці було задане те чи інше цифрове дискретне значення в залежності від ступеня її прояви або наявності.

Вихідною ознакою є результат аналізу на біопсію, яка підтверджує чи спростовує наявність раку шийки матки в організмі жінки.

Розроблюваний програмний код базується на розробках Huang, Peng-Hsuan для використання бібліотеки Keras для побудови моделі неронної мережі [49].

4.1.1 Обробка вхідних даних

Обробка вхідних даних відбувалася за допомогою мови програмування Python. Вхідна вибірка у її початковому стані зберігалася у файлі з розширенням «csv» (рис. 4.2).

The image shows a screenshot of an Excel spreadsheet titled 'risk_factors_cervical_cancer.xlsx'. The spreadsheet contains a large table of data with columns labeled A through AC and rows numbered 1 through 40. The data includes various numerical values and some text, representing initial data for a study. The columns are labeled with variables such as Age, Number of First sexual partners, Number of Smokes, Smokes (packs per year), Hormonal Contraception, Hormonal IUD, IUD (years), STDs (num), STDs (cond), STDs (cerv), STDs (vagi), STDs (vulv), STDs (syph), STDs (pelv), STDs (geni), STDs (omni), STDs (AIDS), STDs (HIV), STDs (HepB), STDs (HPV), STDs (Num), STDs (Tim), and STDs (Cancer). The rows contain numerical data points for each of these variables across 40 different subjects or observations.

Рисунок 4.2 – Початкові дані

Для того, щоб зручно було працювати з мовою програмування Python було встановлено спеціальний пакет програм «Anaconda3». Це безкоштовне відкрите (open-source) програмне забезпечення, яке дозволяє працювати з такими мовами програмування як Python/R для реалізації машинного навчання на операційних системах Linux, Windows, and Mac OS X. Це стандарт для розробки, тестування та навчання, що дозволяє виконувати наступні операції:

- швидко завантажувати 1,500 + пакети Python/R пакетів даних;
- керувати бібліотеками, залежностями та середовищами з Conda;
- розробляти та тренувати машинне навчання та моделі глибокого навчання з scikit-learn, TensorFlow та Theano;
- аналізувати дані масштабовано та продуктивно за допомогою Dask, NumPy, pandas та Numba;
- візуалізувати результати за допомогою Matplotlib, Bokeh, Datashader і Holoviews [50].

На рисунку 4.3 зображено пакети програм, у середовищі яких можна працювати, інсталивавши «Anaconda Distribution».

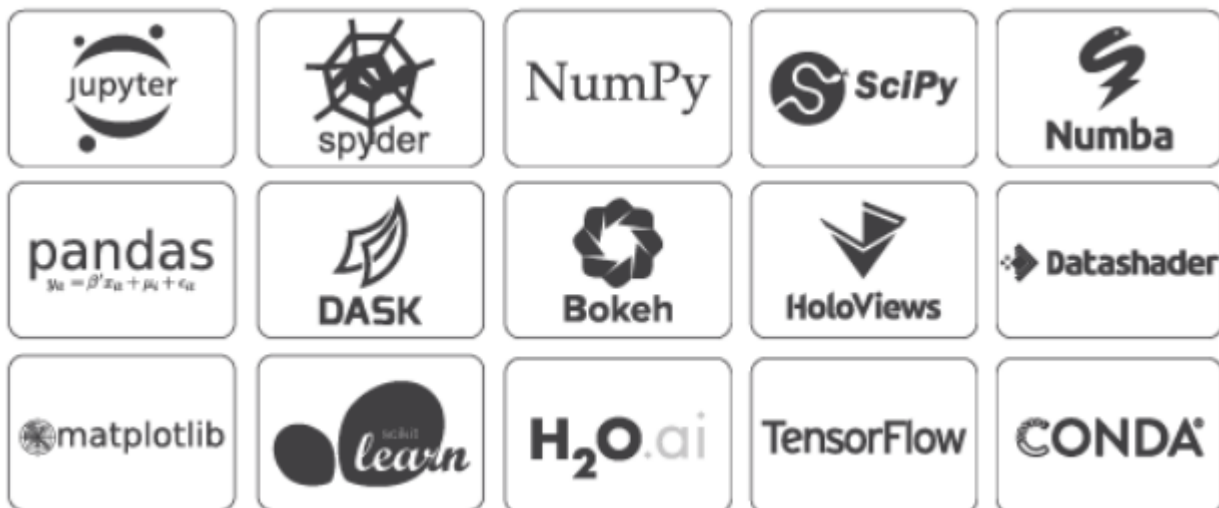


Рисунок 4.3 – Пакети програм, що включає в себе «Anaconda Distribution»

Для реалізації програмного коду Python робота буде виконуватися за допомогою програмного забезпечення «Jupyter Notebook». На рисунку 4.4 зображено початкову сторінку, яку завантажує даний програмний продукт.

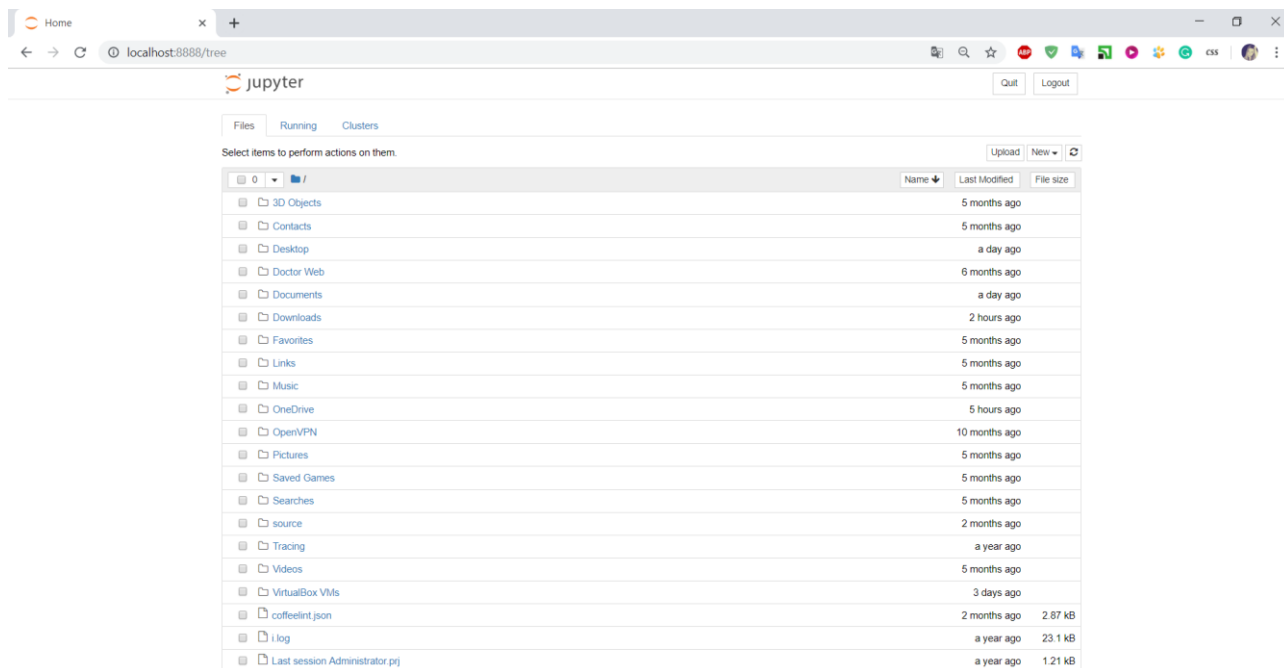


Рисунок 4.4 – Початкова сторінка завантаження «Jupyter Notebook»

Для того, щоб створити новий виконуючий файл необхідно натиснути на кнопку «New» та обрати з випадаючого списку «Python 3» (рис. 4.5).



Рисунок 4.5 – Створення нового файлу для Python скрипту

Виконаємо попередній огляд даних. Для цього потрібно завантажити масив вхідних даних до Python скрипту. Це можна виконати за допомогою команди (рис. 4.6):

```
initial_data = pd.read_csv("data/risk_factors_cervical_cancer.csv")
```

Для потатку потрібно імпортувати стандартні бібліотеки для роботи з даними:

```
import numpy as np
import pandas as pd
```

NumPy є основним пакетом для наукових обчислень з Python. Він містить:

- потужний об'єкт N-розмірного масиву;
- складні функції;
- інструменти для інтеграції C / C ++ і коду Fortran;
- лінійну алгебру, перетворення Фур'є і випадкові числа.

NumPy також може використовуватися як ефективний багатовимірний контейнер загальних даних. Можуть бути визначені довільні типи даних. Це дозволяє йому легко і швидко інтегруватися з широким спектром баз даних.

NumPy має ліцензією BSD (Berkeley Software Distribution – програмна ліцензія університету Берклі – ліцензійна угода, вперше застосована для поширення в UNIX-подібних операційних систем BSD), що дозволяє його повторне використання з невеликими обмеженнями [51].

Pandas – ліцензована бібліотека BSD з відкритим вихідним кодом, що надає високопродуктивні, прості у використанні структури даних та інструменти аналізу даних для мови програмування Python [52].

```
In [1]: import numpy as np
import pandas as pd
```

1. Огляд даних

```
In [2]: initial_data = pd.read_csv("data/risk_factors_cervical_cancer.csv")
initial_data.head()
```

Out[2]:

	Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes	Smokes (years)	Smokes (packs/year)	Hormonal Contraceptives	Hormonal Contraceptives (years)	IUD	...	STDs: Time since first diagnosis	STDs: Time since last diagnosis	Dx:Cancer	Dx:CIN
0	18	4.0	15.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	?	?	0	0
1	15	1.0	14.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	?	?	0	0
2	34	1.0	?	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	?	?	0	0
3	52	5.0	16.0	4.0	1.0	37.0	37.0	1.0	3.0	0.0	...	?	?	1	0
4	46	3.0	21.0	4.0	0.0	0.0	0.0	1.0	15.0	0.0	...	?	?	0	0

5 rows x 36 columns

Рисунок 4.6 – Завантаження вхідних даних до Python скрипту

Початкові вхідні дані необхідно обробити, так як вони містять порожні значення, «?» або неправильний тип даних. Виконаємо заміну «?» на числовий тип даних NaN за допомогою команди:

```
initial_data_with_NaN = initial_data.replace('?', np.nan)
```

Далі потрібно змінити тип ланих з «object» на числовий тип «int64» або «float64»:

```
df_data = df_data.apply(lambda col:pd.to_numeric(col, errors='coerce'))
```

Перевіримо тип даних вхідних ознак (рис. 4.7).

```

In [10]: # Результат перетворення даних
df_data.infer_objects().dtypes

Out[10]: Age                                int64
Number of sexual partners                  float64
First sexual intercourse                   float64
Num of pregnancies                         float64
Smokes                                    float64
Smokes (years)                           float64
Smokes (packs/year)                       float64
Hormonal Contraceptives                   float64
Hormonal Contraceptives (years)           float64
IUD                                        float64
IUD (years)                               float64
STDs                                       float64
STDs (number)                             float64
STDs:condylomatosis                       float64
STDs:cervical condylomatosis              float64
STDs:vaginal condylomatosis               float64
STDs:vulvo-perineal condylomatosis        float64
STDs:syphilis                             float64
STDs:pelvic inflammatory disease          float64
STDs:genital herpes                       float64
STDs:molluscum contagiosum                float64
STDs:AIDS                                 float64
STDs:HIV                                  float64
STDs:Hepatitis B                           float64
STDs:HPV                                  float64
STDs: Number of diagnosis                  int64
STDs: Time since first diagnosis            float64
STDs: Time since last diagnosis            float64
Dx:Cancer                                 int64
Dx:CIN                                    int64
Dx:HPV                                    int64
Dx                                         int64
Hinselmann                                int64
Schiller                                  int64
Citology                                  int64
Biopsy                                    int64
..

```

Рисунок 4.7 – Результат перетворення даних

Потрібно заповнити всі комірки значеннями. Тому для числових змінних NaN значення заповнюємо значеннями медіани попередніх стовпців, а для категоріальних змінних заповнюємо будь-яким значеннями (рандомно), витягнутими з того є стовпця або «0»:

```

df_data['Number of sexual partners'] = df_data['Number of sexual partners'].fillna(df_data['Number of sexual partners'].median())
df_data['First sexual intercourse'] = df_data['First sexual intercourse'].fillna(df_data['First sexual intercourse'].median())
df_data['Num of pregnancies'] = df_data['Num of pregnancies'].fillna(df_data['Num of pregnancies'].median())
df_data['Smokes'] = df_data['Smokes'].fillna(1)
df_data['Smokes (years)'] = df_data['Smokes (years)'].fillna(df_data['Smokes (years)'].median())
df_data['Smokes (packs/year)'] = df_data['Smokes (packs/year)'].fillna(df_data['Smokes (packs/year)'].median())
df_data['Hormonal Contraceptives'] = df_data['Hormonal Contraceptives'].fillna(1)
df_data['Hormonal Contraceptives (years)'] = df_data['Hormonal Contraceptives (years)'].fillna(df_data['Hormonal Contraceptives (years)'].median())
df_data['IUD'] = df_data['IUD'].fillna(0)
df_data['IUD (years)'] = df_data['IUD (years)'].fillna(0)
df_data['STDs'] = df_data['STDs'].fillna(1)
df_data['STDs (number)'] = df_data['STDs (number)'].fillna(df_data['STDs (number)'].median())
df_data['STDs:condylomatosis'] = df_data['STDs:condylomatosis'].fillna(df_data['STDs:condylomatosis'].median())
df_data['STDs:cervical condylomatosis'] = df_data['STDs:cervical condylomatosis'].fillna(df_data['STDs:cervical condylomatosis'].median())
df_data['STDs:vaginal condylomatosis'] = df_data['STDs:vaginal condylomatosis'].fillna(df_data['STDs:vaginal condylomatosis'].median())

```

```

df_data['STDs:vulvo-perineal condylomatosis'] = df_data['STDs:vulvo-perineal
condylomatosis'].fillna(df_data['STDs:vulvo-perineal condylomatosis'].median())
df_data['STDs:syphilis'] = df_data['STDs:syphilis'].fillna(df_data['STDs:syphilis'].median())
df_data['STDs:pelvic inflammatory disease'] = df_data['STDs:pelvic inflammatory
disease'].fillna(df_data['STDs:pelvic inflammatory disease'].median())
df_data['STDs:genital herpes'] = df_data['STDs:genital herpes'].fillna(df_data['STDs:genital herpes'].median())
df_data['STDs:molluscum contagiosum'] = df_data['STDs:molluscum contagiosum'].fillna(df_data['STDs:molluscum
contagiosum'].median())
df_data['STDs:AIDS'] = df_data['STDs:AIDS'].fillna(df_data['STDs:AIDS'].median())
df_data['STDs:HIV'] = df_data['STDs:HIV'].fillna(df_data['STDs:HIV'].median())
df_data['STDs:Hepatitis B'] = df_data['STDs:Hepatitis B'].fillna(df_data['STDs:Hepatitis B'].median())
df_data['STDs:HPV'] = df_data['STDs:HPV'].fillna(df_data['STDs:HPV'].median())
df_data['STDs: Time since first diagnosis'] = df_data['STDs: Time since first diagnosis'].fillna(df_data['STDs: Time
since first diagnosis'].median())
df_data['STDs: Time since last diagnosis'] = df_data['STDs: Time since last diagnosis'].fillna(df_data['STDs: Time
since last diagnosis'].median())

```

На рисунку 4.8 зображено результуючий заповнений масив вхідних даних.

In [12]: df_data

Out [12]:

	Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes	Smokes (years)	Smokes (packs/year)	Hormonal Contraceptives	Hormonal Contraceptives (years)	IUD	...	STDs: Time since first diagnosis	STDs: Time since last diagnosis	Dx: Cancer	Dx:
0	18	4.0	15.0	1.0	0.0	0.000000	0.00	0.0	0.00	0.0	...	4.0	3.0	0	
1	15	1.0	14.0	1.0	0.0	0.000000	0.00	0.0	0.00	0.0	...	4.0	3.0	0	
2	34	1.0	17.0	1.0	0.0	0.000000	0.00	0.0	0.00	0.0	...	4.0	3.0	0	
3	52	5.0	16.0	4.0	1.0	37.000000	37.00	1.0	3.00	0.0	...	4.0	3.0	1	
4	46	3.0	21.0	4.0	0.0	0.000000	0.00	1.0	15.00	0.0	...	4.0	3.0	0	
5	42	3.0	23.0	2.0	0.0	0.000000	0.00	0.0	0.00	0.0	...	4.0	3.0	0	
6	51	3.0	17.0	6.0	1.0	34.000000	3.40	0.0	0.00	1.0	...	4.0	3.0	0	
7	26	1.0	26.0	3.0	0.0	0.000000	0.00	1.0	2.00	1.0	...	4.0	3.0	0	
8	45	1.0	20.0	5.0	0.0	0.000000	0.00	0.0	0.00	0.0	...	4.0	3.0	1	
9	44	3.0	15.0	2.0	1.0	1.266973	2.80	0.0	0.00	0.0	...	4.0	3.0	0	
10	44	3.0	26.0	4.0	0.0	0.000000	0.00	1.0	2.00	0.0	...	4.0	3.0	0	
11	27	1.0	17.0	3.0	0.0	0.000000	0.00	1.0	8.00	0.0	...	4.0	3.0	0	
12	45	4.0	14.0	6.0	0.0	0.000000	0.00	1.0	10.00	1.0	...	4.0	3.0	0	
13	44	2.0	25.0	2.0	0.0	0.000000	0.00	1.0	5.00	0.0	...	4.0	3.0	0	
14	43	2.0	18.0	5.0	0.0	0.000000	0.00	0.0	0.00	1.0	...	4.0	3.0	0	
15	40	3.0	18.0	2.0	0.0	0.000000	0.00	1.0	15.00	0.0	...	4.0	3.0	0	
16	41	4.0	21.0	3.0	0.0	0.000000	0.00	1.0	0.25	0.0	...	4.0	3.0	0	
17	43	3.0	15.0	8.0	0.0	0.000000	0.00	1.0	3.00	0.0	...	4.0	3.0	0	
18	42	2.0	20.0	2.0	0.0	0.000000	0.00	1.0	7.00	1.0	...	4.0	3.0	0	
19	40	2.0	27.0	2.0	0.0	0.000000	0.00	0.0	0.00	1.0	...	4.0	3.0	0	
20	43	2.0	18.0	4.0	0.0	0.000000	0.00	1.0	15.00	0.0	...	4.0	3.0	0	
21	41	3.0	17.0	4.0	0.0	0.000000	0.00	1.0	10.00	0.0	...	21.0	21.0	0	

Рисунок 4.8 – Результуючий заповнений масив вхідних даних

Створимо новий dataframe зі значеннями тільки 0 та 1 для категоріальних змінних. У цьому випадку всі порожні категоріальні змінні заповнюються даними зі значеннями 0 та 1. Наприклад, якщо значення 'Smokes' = 0, то буде створено дві нові колонки зі значеннями 'Smokes_0' = 1 та 'Smokes_1' = 0, а якщо значення 'Smokes' = 1, то – 'Smokes_0' = 0 та 'Smokes_1' = 1:

```
df_data = pd.get_dummies(data=df_data, columns=['Smokes','Hormonal
Contraceptives','IUD','STDs','Dx:Cancer','Dx:CIN','Dx:HPV','Dx','Hinselmann','Citology','Schiller'])
```

Виконаємо перевірку на порожні значення за допомогою команди:

```
df_data.isnull().sum()
```

Результат обробки даних представлений на рисунках 4.9 – 4.10.

```
In [15]: # Перевіримо значення на NaN
df_data.isnull().sum()

Out[15]: Age 0
Number of sexual partners 0
First sexual intercourse 0
Num of pregnancies 0
Smokes (years) 0
Smokes (packs/year) 0
Hormonal Contraceptives (years) 0
IUD (years) 0
STDs (number) 0
STDs:condylomatosis 0
STDs:cervical condylomatosis 0
STDs:vaginal condylomatosis 0
STDs:vulvo-perineal condylomatosis 0
STDs:syphilis 0
STDs:pelvic inflammatory disease 0
STDs:genital herpes 0
STDs:molluscum contagiosum 0
STDs:AIDS 0
STDs:HIV 0
STDs:Hepatitis B 0
STDs:HPV 0
STDs: Number of diagnosis 0
STDs: Time since first diagnosis 0
STDs: Time since last diagnosis 0
Biopsy 0
Smokes_0.0 0
Smokes_1.0 0
Hormonal Contraceptives_0.0 0
Hormonal Contraceptives_1.0 0
IUD_0.0 0
IUD_1.0 0
STDs_0.0 0
STDs_1.0 0
Dx:Cancer_0 0
Dx:Cancer_1 0
Dx:CIN_0 0
```

Рисунок 4.9 – Перевірка масиву на порожні значення

```
In [16]: df_data

Out[16]:
```

	Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes (years)	Smokes (packs/year)	Hormonal Contraceptives (years)	IUD (years)	STDs (number)	STDs:condylomatosis	...	Dx:HPV_0	Dx:HPV_1
0	18	4.0	15.0	1.0	0.000000	0.00	0.00	0.00	0.0	0.0	...	1	0
1	15	1.0	14.0	1.0	0.000000	0.00	0.00	0.00	0.0	0.0	...	1	0
2	34	1.0	17.0	1.0	0.000000	0.00	0.00	0.00	0.0	0.0	...	1	0
3	52	5.0	16.0	4.0	37.000000	37.00	3.00	0.00	0.0	0.0	...	0	1
4	46	3.0	21.0	4.0	0.000000	0.00	15.00	0.00	0.0	0.0	...	1	0
5	42	3.0	23.0	2.0	0.000000	0.00	0.00	0.00	0.0	0.0	...	1	0
6	51	3.0	17.0	6.0	34.000000	3.40	0.00	7.00	0.0	0.0	...	1	0
7	26	1.0	26.0	3.0	0.000000	0.00	2.00	7.00	0.0	0.0	...	1	0
8	45	1.0	20.0	5.0	0.000000	0.00	0.00	0.00	0.0	0.0	...	0	1
9	44	3.0	15.0	2.0	1.266973	2.80	0.00	0.00	0.0	0.0	...	1	0

Рисунок 4.10 – Фінальний опрацьований масив вхідних дани

Отримані дані збережемо до окремого csv-файлу для подальшої роботи з ними за допомогою команди:

```
df_data.to_csv('Processed_data.csv', header = True, index = False)
```

4.1.2 Розробка навчальної та тестової вибірок

Створюємо новий Python скрипт для розробки навчальної та тестової вибірок. Імпортуємо у нього оброблені дані з файлу «Processed_data.csv». Всього масив даних складає 858 записів пацієнтів про індивідуальні показники організму. Для того, щоб навчання нейронної мережі було виконане коректно, включимо у навчальну вибірку більше половини оброблених записів, а саме 600. Таким чином, тренувальна вибірка буде складатися з 258 записів. Необхідно визначити функціональні та цільові мітки. Як було зазначено раніше, функціональні мітки будуть складатися параметри пацієнтів, цільовою міткою є біопсія.

Функціональні мітки тренувальної та тестової вибірок:

```
df_train_feature = df_train[['Age', 'Number of sexual partners', 'First sexual intercourse',
    'Num of pregnancies', 'Smokes (years)', 'Smokes (packs/year)',
    'Hormonal Contraceptives (years)', 'IUD (years)', 'STDs (number)',
    'STDs:condylomatosis', 'STDs:cervical condylomatosis',
    'STDs:vaginal condylomatosis', 'STDs:vulvo-perineal condylomatosis',
    'STDs:syphilis', 'STDs:pelvic inflammatory disease',
    'STDs:genital herpes', 'STDs:molluscum contagiosum', 'STDs:AIDS',
    'STDs:HIV', 'STDs:Hepatitis B', 'STDs:HPV', 'STDs: Number of diagnosis',
    'STDs: Time since first diagnosis', 'STDs: Time since last diagnosis',
    'Smokes_0.0', 'Smokes_1.0',
    'Hormonal Contraceptives_0.0', 'Hormonal Contraceptives_1.0', 'IUD_0.0',
    'IUD_1.0', 'STDs_0.0', 'STDs_1.0', 'Dx:Cancer_0', 'Dx:Cancer_1',
    'Dx:CIN_0', 'Dx:CIN_1', 'Dx:HPV_0', 'Dx:HPV_1', 'Dx_0', 'Dx_1',
    'Hinselmann_0', 'Hinselmann_1', 'Citology_0', 'Citology_1', 'Schiller_0', 'Schiller_1']]
```

```
df_test_feature = df_test[['Age', 'Number of sexual partners', 'First sexual intercourse',
    'Num of pregnancies', 'Smokes (years)', 'Smokes (packs/year)',
    'Hormonal Contraceptives (years)', 'IUD (years)', 'STDs (number)',
    'STDs:condylomatosis', 'STDs:cervical condylomatosis',
    'STDs:vaginal condylomatosis', 'STDs:vulvo-perineal condylomatosis',
    'STDs:syphilis', 'STDs:pelvic inflammatory disease',
    'STDs:genital herpes', 'STDs:molluscum contagiosum', 'STDs:AIDS',
    'STDs:HIV', 'STDs:Hepatitis B', 'STDs:HPV', 'STDs: Number of diagnosis',
    'STDs: Time since first diagnosis', 'STDs: Time since last diagnosis',
    'Smokes_0.0', 'Smokes_1.0',
    'Hormonal Contraceptives_0.0', 'Hormonal Contraceptives_1.0', 'IUD_0.0',
    'IUD_1.0', 'STDs_0.0', 'STDs_1.0', 'Dx:Cancer_0', 'Dx:Cancer_1',
    'Dx:CIN_0', 'Dx:CIN_1', 'Dx:HPV_0', 'Dx:HPV_1', 'Dx_0', 'Dx_1',
    'Hinselmann_0', 'Hinselmann_1', 'Citology_0', 'Citology_1', 'Schiller_0', 'Schiller_1']]
```

Цільові мітки тренувальної та тестової вибірок:

```
train_label = np.array(df_train['Biopsy'])
test_label = np.array(df_test['Biopsy'])
```

Виконаємо нормалізацію даних. Зробимо діапазон значень даних у межах від $\{0...1\}$:

```
minmax_scale = MinMaxScaler(feature_range=(0, 1))
train_feature = minmax_scale.fit_transform(df_train_feature)
test_feature = minmax_scale.fit_transform(df_test_feature)
```

Збережемо отримані дані до окремих файлів: для навчальної вибірки – «Train_data.csv», для тестової вибірки – «Test_data.csv».

4.1.3 Кореляція між параметрами вхідних даних

Для того, щоб визначити залежність індивідуальних показників пацієнта один від одного побудуємо графік їх кореляції. Для цього імпортуємо бібліотеки для зображення теплової карти:

```
import matplotlib.pyplot as plt
import seaborn as sns
```

Теплова карта – це графічне представлення даних, де окремі значення, що містяться в матриці, представлені у вигляді кольорів [53].

Виконаємо наступні команди для побудови контуру вхідних даних (рис. 4.11):

```
corrmat = data.corr()
f, ax = plt.subplots(figsize=(15, 15))
sns.heatmap(corrmat, vmax=1, square=True, cmap='gist_rainbow')
```

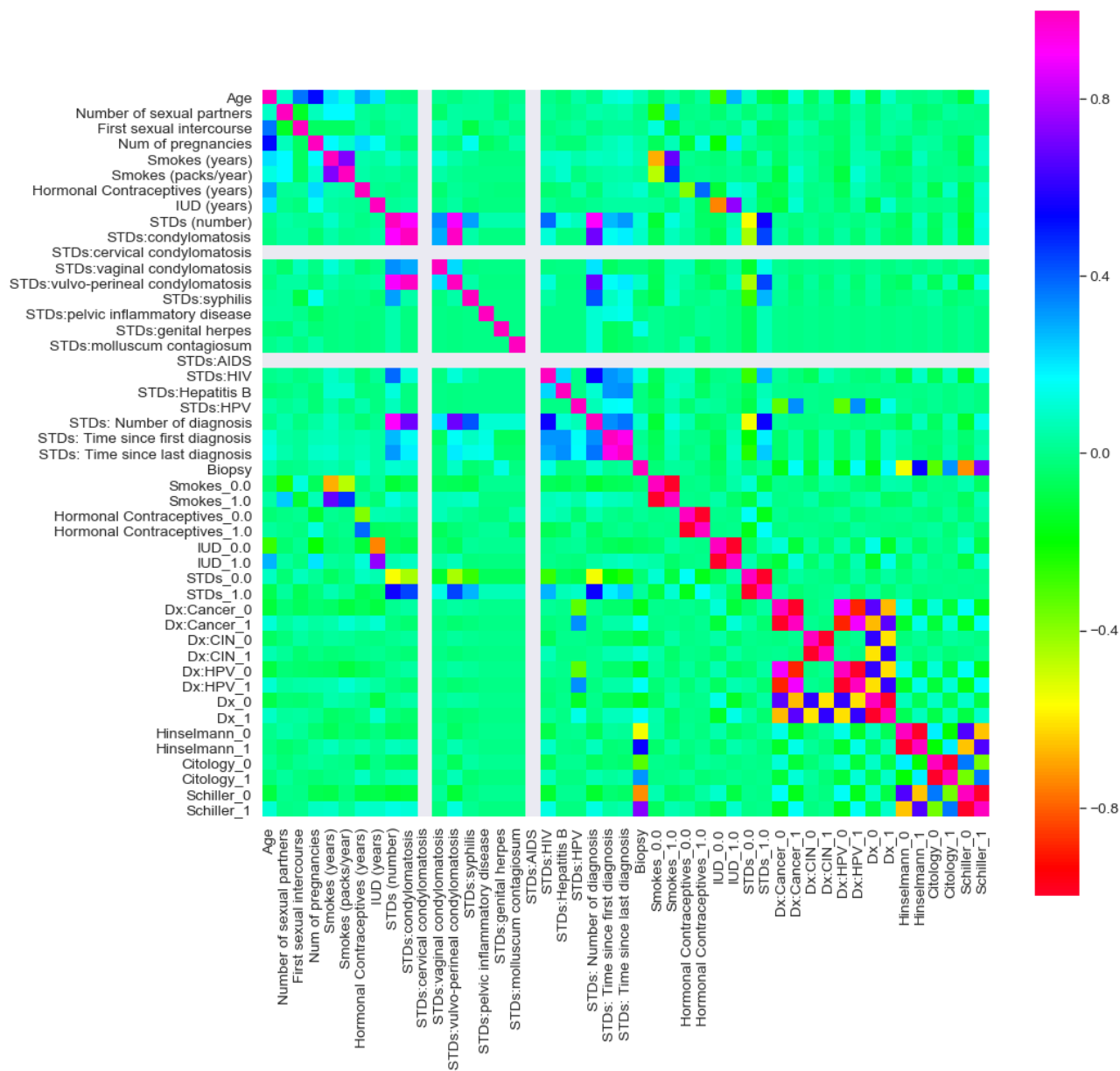


Рисунок 4.11 – Вхідний контур даних

Для того, щоб простежити кореляцію елементів – ознак пацієнтів в залежності від кластеризації їх розміщення, виконаємо наступне (рис. 4.12):

```

k = 46 # кількість змінних на тепловій мапі
cols = corrmatrix.nlargest(k, 'Biopsy')['Biopsy'].index
cm = np.corrcoef(data[cols].values.T)

plt.figure(figsize=(15,15))

sns.set(font_scale=1.25)
hm = sns.heatmap(cm, cbar=True, annot=True, square=True, cmap='jet', fmt='.2f', annot_kws={'size': 10},
yticklabels = cols.values, xticklabels = cols.values)
plt.show()

```

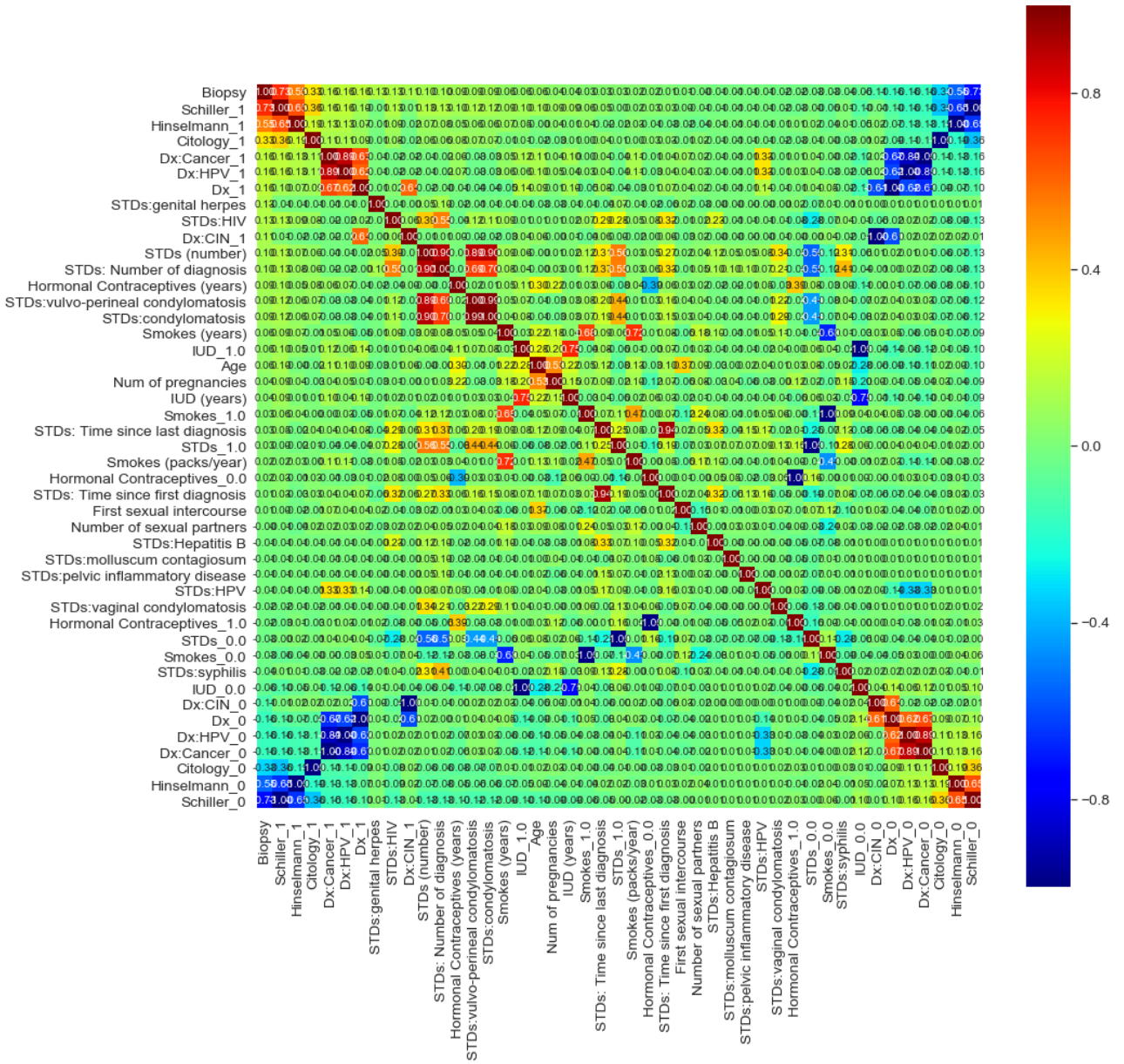


Рисунок 4.12 – Теплова карта кореляції показників пацієнтів

З графіку можемо зробити такий висновок, що коефіцієнти, які мають найбільшу позитивну кореляцію з біопсією це "Schiller_1" (позитивний результат тесту Шиллера), "Hinselmann_1" (позитивний результат тесту Хінсельмана) та "Citology_1" (позитивний результат аналізу на цитологію). Більшість показників не корелюються один від одного, що дозволяє зробити більш точні та достовірні результати.

4.2 Створення бази даних

Створення бази даних було реалізоване за допомогою об'єктно-реляційна система керування базами даних PostgreSQL. Фундаментальна характеристика об'єктно-реляційної бази даних – це підтримка об'єктів і їх поведінки, включаючи типи даних, функції, операції, домени та індекси. Це робить PostgreSQL досить гнучким і надійним. Також за допомогою нього можна створювати, зберігати та вилучати складні структури даних [54].

Для зручної роботи з базою даних PostgreSQL було встановлено десктопний додаток «PostgreSQL 11», який включає в себе open-source інструмент для керування базою даних PostgreSQL «pgAdmin» [55]. Після завантаження інструменту додаток завантажує початкову сторінку, зображену на рисунку 4.13.

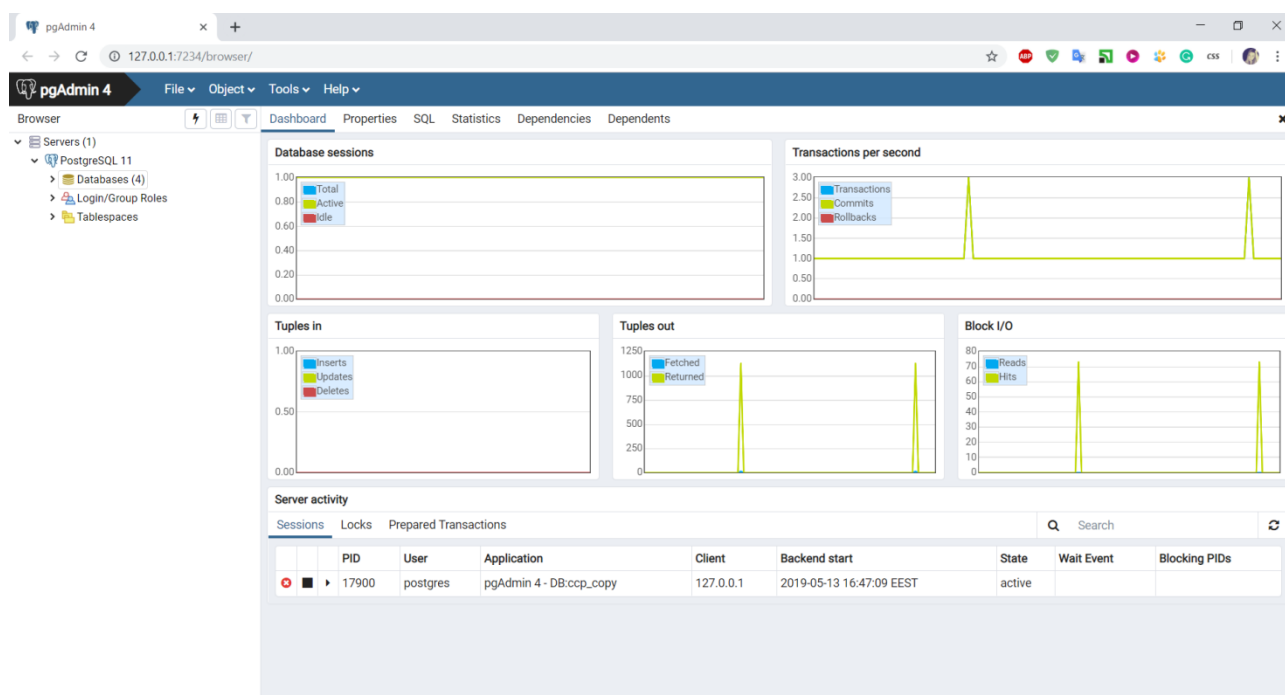


Рисунок 4.13 – Початкова сторінка завантаження додатку «PostgreSQL 11»

Для того, щоб створити нову базу даних потрібно перемістити курсор на пункт «Databases», натиснути правою кнопкою миші, обрати пункт «Create» – «Database...» (рис. 4.14).

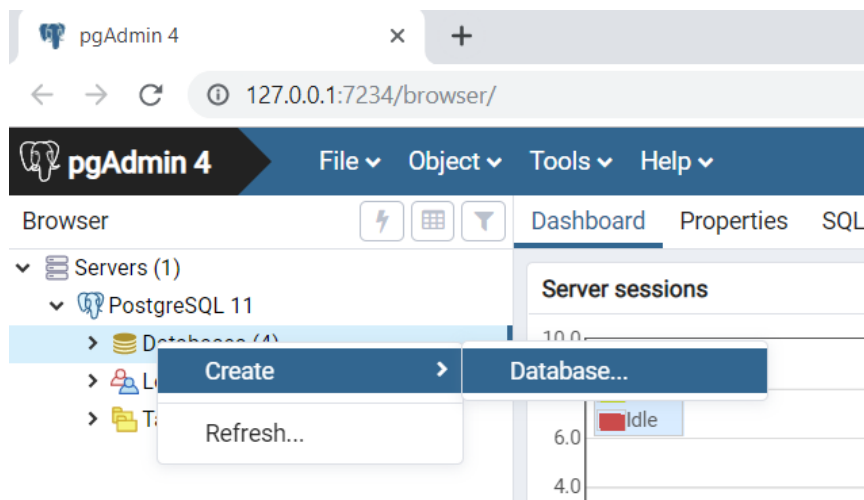


Рисунок 4.14 – Створення бази даних

У відкритому діалоговому вікні «Create – Database» вводимо назву бази даних «ccp» та обираємо роль («Owner») – «postgres» та натискаємо кнопку «Save» (рис. 4.15).

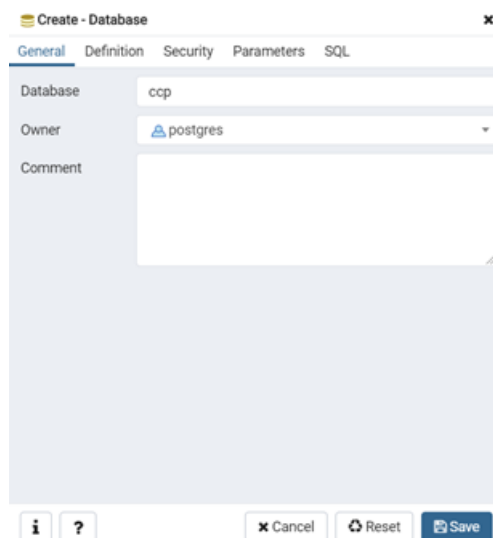


Рисунок 4.15 – Введення назви бази даних

У подальшому робота з базою даних буде проводитися за допомогою програмного коду, використовуючи середовище ASP NET Core.

4.3 Імпорт даних до бази даних

Оброблені дані були збережені в окремі csv-файли для навчальної та тестової вибірок. Для початку було зчитано з файлів «Train_data.csv» та «Test_data.csv» готові масиви даних до Python скрипту:

```
train_data = pd.read_csv("Train_data.csv")
test_data = pd.read_csv("Test_data.csv")
```

Імпортуємо бібліотеку «psycopg2» – найбільш популярний адаптер бази даних PostgreSQL для мови програмування Python:

```
import psycopg2
from sqlalchemy import create_engine
```

Було виконане з'єднання з базою даних «ссп». Далі виконано імпорт оброблених даних до таблиць «TrainDatasetModels» та «TestDatasetModels» бази даних PostgreSQL «ссп»:

```
engine = create_engine('postgresql://postgres:root@localhost:5432/ссп')
train_data.to_sql('TrainDatasetModels', con=engine, if_exists = 'replace', index=False)
test_data.to_sql('TestDatasetModels', con=engine, if_exists = 'replace', index=False)
```

На рисунках 4.16 – 4.17 представлено заповнені таблиці «TrainDatasetModels» та «TestDatasetModels» бази даних PostgreSQL обробленими даними для тренування та тестування нейронної мережі.

The screenshot shows the pgAdmin 4 interface with the 'TrainDatasetModels' table selected. The table data is as follows:

IDTrainDataset	Age	NumberOfSexualPartners	AgeOfFirstSexualIntercourse	NumberOfPregnancies	AgeSmokes	PacksYearSmokes	AgeUsingHormonalContraceptives
1	15	1	14	1	0	0	
2	34	1	17	1	0	0	
3	52	5	16	4	37	37	
4	46	3	21	4	0	0	
5	42	3	23	2	0	0	
6	51	3	17	6	34	3.4	
7	26	1	26	3	0	0	
8	45	1	20	5	0	0	
9	44	3	15	2	1.266972909	2.8	
10	44	3	26	4	0	0	
11	27	1	17	3	0	0	
12	45	4	14	6	0	0	
13	44	2	25	2	0	0	
14	43	2	18	5	0	0	
15	40	3	18	2	0	0	
16	41	4	21	3	0	0	
17	43	3	15	8	0	0	

Рисунок 4.16 – Таблица «TrainDatasetModels» бази даних PostgreSQL

The screenshot shows the pgAdmin 4 interface with the 'TestDatasetModels' table selected. The table data is as follows:

IDTestDataset	Age	NumberOfSexualPartners	AgeOfFirstSexualIntercourse	NumberOfPregnancies	AgeSmokes	PacksYearSmokes	AgeUsingHormonalContraceptives
1	28	3	18	2	0	0	
2	34	3	16	2	0	0	
3	26	4	16	3	0	0	
4	23	3	17	2	0	0	
5	20	2	18	1	0	0	
6	29	5	17	5	0	0	
7	30	2	29	1	0	0	
8	19	2	16	1	0	0	
9	26	2	16	3	11	2.2	
10	21	2	18	3	0	0	
11	23	2	19	1	0	0	
12	21	2	15	1	1.266972909	0.05	
13	23	2	15	2	0	0	
14	24	1	23	1	0	0	
15	35	6	12	6	0	0	
16	25	2	16	2	8	0.16	
17	25	2	16	2	8	0.16	

Рисунок 4.17 – Таблица «TestDatasetModels» бази даних PostgreSQL

4.4 Розробка нейронної мережі

Було створено окремий Python скрипт для навчання та тестування нейронної мережі. Після підключення до бази даних PostgreSQL «ccp» (host=localhost; port=5432; database=ccp; user ID=postgres; password=root) та імпортування даних у змінні з таблиць «>» та «<»), шляхом виконання SQL запиту, було виконано побудову моделі нейронної мережі:

```
engine = create_engine('postgresql://postgres:root@localhost:5432/ccp')
train_data = pd.read_sql_query('SELECT "Age", "NumberOfSexualPartners", "AgeOfFirstSexualIntercourse",
"NumberOfPregnancies", \
    "AgeSmokes", "PacksYearSmokes", "AgeUsingHormonalContraceptives", "AgeIUD", \
    "NumberOfSTDs", "STDs_Condylomatosis", "STDs_CervicalCondylomatosis",
"STDs_VaginalCondylomatosis", \
    "STDs_VulvoPerinealCondylomatosis", "STDs_Syphilis", "STDs_PelvicInflammatoryDisease", \
    "STDs_GenitalHerpes", "STDs_MolluscumContagiosum", "STDs_AIDS", "STDs_HIV", "STDs_HepatitisB",
\
    "STDs_HPV", "STDs_NumberOfDiagnosis", "STDs_TimeSinceFirstDiagnosis",
"STDs_TimeSinceLastDiagnosis", \
    "Smokes_0", "Smokes_1", "UsingHormonalContraceptives_0", "UsingHormonalContraceptives_1",
"IUD_0", \
    "IUD_1", "STDs_0", "STDs_1", "Dx_Cancer_0", "Dx_Cancer_1", "Dx_CIN_0", "Dx_CIN_1",
"Dx_HPV_0", "Dx_HPV_1", \
    "Dx_0", "Dx_1", "HinselmannTest_0", "HinselmannTest_1", "Citology_0", "Citology_1", "SchillerTest_0", \
    "SchillerTest_1" FROM "TrainDatasetModels", con=engine)
train_label = pd.read_sql_query('SELECT "Biopsy" FROM "TrainDatasetModels", con=engine)
test_data = pd.read_sql_query('SELECT "Age", "NumberOfSexualPartners", "AgeOfFirstSexualIntercourse",
"NumberOfPregnancies", \
    "AgeSmokes", "PacksYearSmokes", "AgeUsingHormonalContraceptives", "AgeIUD", \
    "NumberOfSTDs", "STDs_Condylomatosis", "STDs_CervicalCondylomatosis",
"STDs_VaginalCondylomatosis", \
    "STDs_VulvoPerinealCondylomatosis", "STDs_Syphilis", "STDs_PelvicInflammatoryDisease", \
    "STDs_GenitalHerpes", "STDs_MolluscumContagiosum", "STDs_AIDS", "STDs_HIV", "STDs_HepatitisB",
\
    "STDs_HPV", "STDs_NumberOfDiagnosis", "STDs_TimeSinceFirstDiagnosis",
"STDs_TimeSinceLastDiagnosis", \
    "Smokes_0", "Smokes_1", "UsingHormonalContraceptives_0", "UsingHormonalContraceptives_1",
"IUD_0", \
    "IUD_1", "STDs_0", "STDs_1", "Dx_Cancer_0", "Dx_Cancer_1", "Dx_CIN_0", "Dx_CIN_1",
"Dx_HPV_0", "Dx_HPV_1", \
    "Dx_0", "Dx_1", "HinselmannTest_0", "HinselmannTest_1", "Citology_0", "Citology_1", "SchillerTest_0", \
    "SchillerTest_1" FROM "TestDatasetModels", con=engine)
test_label = pd.read_sql_query('SELECT "Biopsy" FROM "TestDatasetModels", con=engine)
```

Основним завданням було вибір оптимальної архітектури нейромережі, яка забезпечувала максимально точні результати. Вхідний шар має 46 входів – оброблені та провалідовані введені користувачем індивідуальні показники

організму. Вихідний шар має 1 вихід, так як результатом оцінювання виникнення ризику цервікального раку є значення позитивний {1} чи негативний {0}. У ході виконання експерименту для визначення кількості прихованих шарів було проведена досить велика кількість дослідів з різними варіаціями побудови моделі нейронної мережі. Для оцінювання якості її навчання було розглянуто декілька параметрів:

- accuracy – точність навчання моделі;
- sensitivity – є показником частки фактичних позитивних випадків, які були прогнозовані як позитивні (або справжні позитивні) [56];
- specificity – визначається як частка фактичних негативів, які отримали прогноз як негативний (або істинний негатив) [56];
- false_positive_rate – коефіцієнт помилкових позитивних значень;
- false_negative_rate – коефіцієнт помилкових негативних значень.

Вище значення sensitivity означатиме більш високе значення істинно позитивних і нижчої величини помилково негативних значень. Більш низьке значення sensitivity означає меншу величину істинно позитивних і більшу величину помилково негативних значень. Для сфери охорони здоров'я та фінансів бажані моделі з високим значенням sensitivity [56].

Більш високе значення specificity означає більшу величину істинно негативних значень і нижчу величину коефіцієнта помилков позитивних значень. Більш низьке значення specificity означатиме меншу величину істинно негативних і більш високе значення помилково позитивних значень [56].

Функції активації. Різні функції активації фактично мають різні властивості. Функцію активації між двома шарами нейронної мережі було обрано «relu», так як вона не насичується, а також швидше обчислюється. Використовуючи цю функцію, похідна для всіх позитивних входів має значення одиниці (1), тому градієнт для тих нейронів, не буде повністю змінений блоком активації і не сповільнить спуск градієнта.

Для вихідного шару мережі блок активації також залежить від завдання. Так як потрібним результатом є класифікація, щоб вихід мав одне значення {0} або {1}, тому доречним є використання функції «sigmoid» для його наближення, якщо класів було б більше одного, то функцією активації на вихідному шарі була б «softmax» [57].

Функція вартості (функція втрат). Метою навчання нейронної мережі є мінімізація функції вартості [58]. Вага і функція активації кожного нейрона призводять до основної функції, яка є нейронною мережею. Це функція вартості, яка повинна бути скоригована (за крок навчання) для отримання кращих результатів [59]. У якості функції втрат для розроблюваної нейромережі було обрано «binary crossentropy». Це активація сигмоїди та втрата крос-ентропії. На відміну від функцією активації «softmax», вона є незалежною для кожної векторної складової (класу), тобто втрата, що обчислюється для кожного компонента векторного виходу, не залежить від інших значень компонентів. Тому «softmax» використовується для багатозначної класифікації, якщо уявлення про елемент, що належить до певного класу, не має впливати на рішення іншого класу. «Sigmoid» називається втратою бінарної крос-ентропії, оскільки вона визначає задачу бінарної класифікації [60].

У таблиці 4.1 зображено експерименти, які проводились для навчання нейронної мережі з метою отримання найвищої точності її навчання. Завдяки ним було обрано нейромережу з п'ятьма шарами.

Таблиця 4.1 – Визначення архітектури нейронної мережі

№	Загальна кількість шарів	Кількість прихованих шарів	Кількість вузлів у прихованих шарах	Кількість епох	Масив вибірок, які будуть поширюватися по мережі	Точність навчання
1	3	1	100	20	200	86,92%
2	3	1	100	20	300	83,14%
3	3	1	100	50	200	87,73%
4	3	1	100	50	300	84,51%

Продовження таблиці 4.1

№	Загальна кількість шарів	Кількість прихованих шарів	Кількість вузлів у прихованих шарах	Кількість епох	Масив вибірок, які будуть поширюватися по мережі	Точність навчання
5	3	1	200	20	200	87,88%
6	3	1	200	20	300	88,37%
7	3	1	200	50	200	90,75%
8	3	1	200	50	300	90,60%
9	3	1	500	20	200	88,97%
10	3	1	500	20	300	87,17%
11	3	1	500	50	200	89,99%
12	3	1	500	50	300	89,66%
13	3	1	500	80	200	89,79%
14	3	1	500	80	300	89,46%
15	5	3	100 100 100	20	200	92,25%
16	5	3	100 100 100	20	300	91,02%
17	5	3	100 100 100	50	200	94,55%
18	5	3	100 100 100	50	300	93,92%
19	5	3	100 100 100	80	200	92,44%
20	5	3	100 100 100	80	300	92,10%

Продовження таблиці 4.1

№	Загальна кількість шарів	Кількість прихованих шарів	Кількість вузлів у прихованих шарах	Кількість епох	Масив вибірок, які будуть поширюватися по мережі	Точність навчання
21	5	3	200 200 200	20	200	90,00%
22	5	3	200 200 200	20	300	89,61%
23	5	3	200 200 200	50	200	90,23%
24	5	3	200 200 200	50	300	90,17%
25	5	3	200 200 200	80	200	89,35%
26	5	3	200 200 200	80	300	89,22%
27	5	3	500 500 500	20	200	88,54%
28	5	3	500 500 500	20	300	88,43%
29	5	3	500 500 500	50	200	89,66%

Продовження таблиці 4.1

№	Загальна кількість шарів	Кількість прихованих шарів	Кількість вузлів у прихованих шарах	Кількість епох	Масив вибірок, які будуть поширюватися по мережі	Точність навчання
30	5	3	500 500 500	50	300	89,33%
31	5	3	500 500 500	80	200	87,58%
32	5	3	500 500 500	80	300	87,29%

Отже, можемо зробити висновок, що найбільш точніший результат навчання зі значенням 94,55% має нейронна мережа з наступною архітектурою:

- загальна кількість шарів – 5;
- кількість прихованих шарів – 3;
- кількість вузлів у вхідному шарі – 46;
- кількість вузлів у прихованих шарах – 100;
- кількість вузлів у вихідному шарі – 1;
- кількість епох – 50;
- розмір масиву вибірок, які будуть поширюватися по мережі – 200.

Схематично архітектуру нейронної мережі зображено на рисунку 4.18.

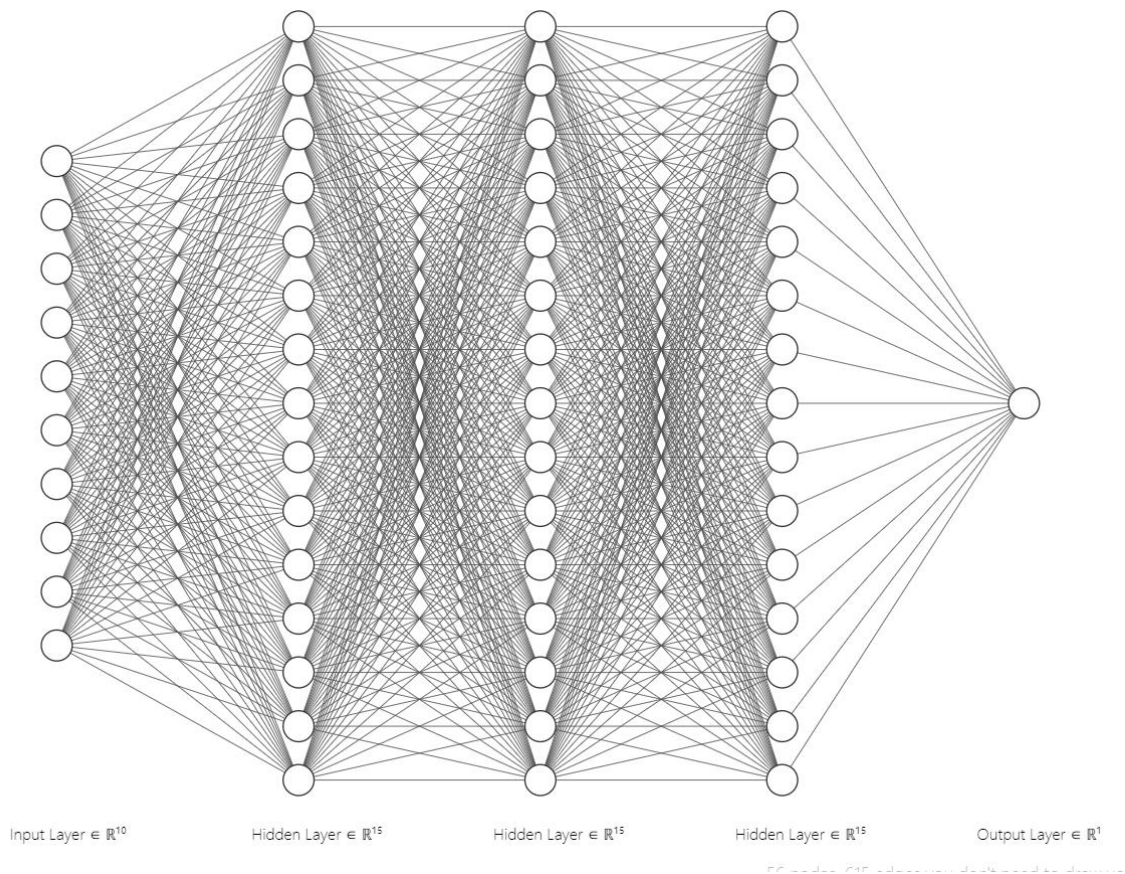


Рисунок 4.18 – Схема нейронної мережі

Розглянемо Python скрипт, який використовувався для розробки нейромережі. Після встановлення з'єднання з базою даних PostgreSQL, було виконано запити для отримання оброблених даних (вхідних параметрів та цільової змінної) з таблиць «TrainDatasetModels» та «TestDatasetModels». Далі виконувалася побудова, навчання та тестування моделі:

```
# виконуємо побудову моделі

# будуємо лінійний стек шарів моделі нейронної мережі
model = Sequential()

# створюємо вхідний шар
model.add(Dense(units=100,
                input_dim=n_cols,
                init='uniform',
                activation='relu'))

# створюємо перший прихований шар
model.add(Dense(units=100,
                init='uniform',
                activation='relu'))
```

```

# створюємо другий прихований шар
model.add(Dense(units=100,
                init='uniform',
                activation='relu'))

# створюємо третій прихований шар
model.add(Dense(units=100,
                init='uniform',
                activation='relu'))

# створюємо вихідний шар
model.add(Dense(units=1,
                init='uniform',
                activation='sigmoid'))

# відображення структури та параметрів нейронної мережі
print(model.summary())

# Визначення як виконувалися розрахунки
model.compile(loss='binary_crossentropy',
              optimizer='adam', metrics=['accuracy'])

# Тренування моделі
# Verbose=2, відображення втрат і точності змін своєчасно
train_history = model.fit(x=train_data, y=train_label,
                        validation_split=0.2, epochs=50,
                        batch_size=200, verbose=2)

# відображення втрат і точності після кожної епохи
show_train_history(train_history, 'acc', 'val_acc')
show_train_history(train_history, 'loss', 'val_loss')

```

На рисунку 4.19 зображено графік значень точності навчання неромережі від кількості пройдених епох.

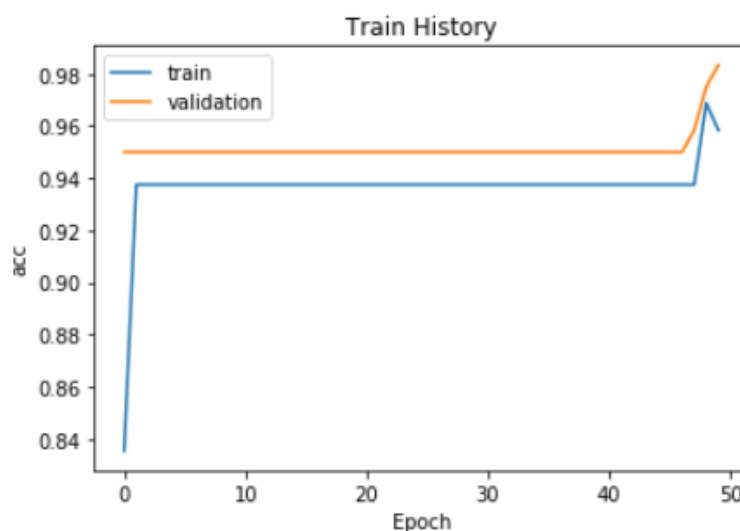


Рисунок 4.19 – Графік залежності точності навчання від кількості епох

На рисунку 4.20 зображено графік залежності мінімізації функції вартості «binary_crossentropy» від кількості пройдених епох.

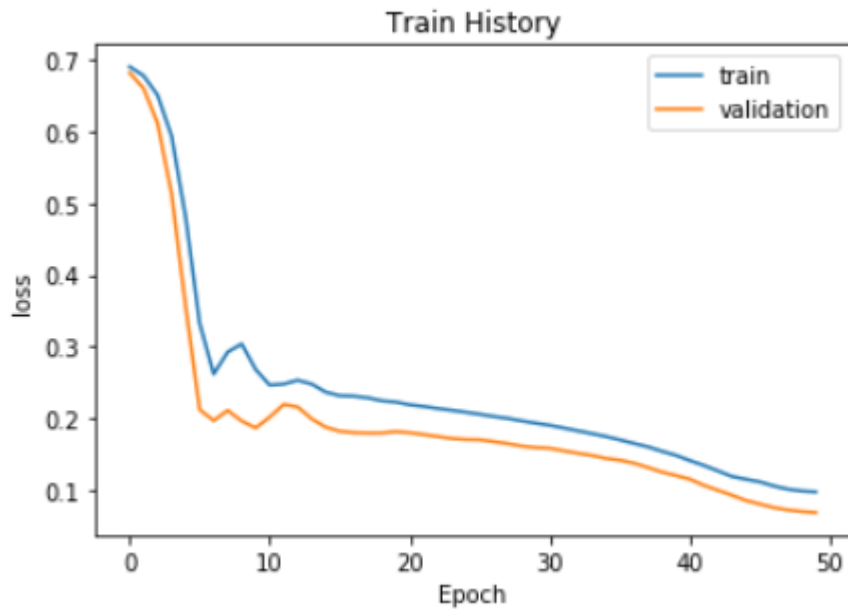


Рисунок 4.20 – Графік залежності мінімізації функції вартості від кількості епох

Далі було розраховано точність моделі прогнозування. За допомогою команди «evaluate» можна отримати значення точності навчання нейронної мережі (рис. 4.21):

```
scores = model.evaluate(test_data, test_label)
```

```
In [33]: # визначаємо точність навчання моделі
scores = model.evaluate(test_data, test_label)
print('\n')
print('accuracy=', scores[1])

257/257 [=====] - 0s 43us/step

accuracy= 0.9455252918287937
```

Рисунок 4.21 – Точність навчання моделі прогнозування

Потім було виконане порівняння значень результатів біопсії тестової вибірки та прогнозованих даних (рис. 4.22).

```
In [21]: # перевіряємо, які значення не співпадають
df_ans[ df_ans['Biopsy'] != df_ans['Prediction'] ]
```

```
Out[21]:
```

	Biopsy	Prediction
43	0	1
45	0	1
52	1	0
54	1	0
69	1	0
70	1	0
72	1	0
85	1	0
86	1	0
87	1	0
92	1	0
138	1	0
149	0	1
152	1	0
153	1	0
185	1	0
213	1	0

```
In [22]: df_ans['Prediction'].value_counts()
```

```
Out[22]: 0    249
         1     8
         Name: Prediction, dtype: int64
```

```
In [23]: df_ans['Biopsy'].value_counts()
```

```
Out[23]: 0    238
         1    19
         Name: Biopsy, dtype: int64
```

Рисунок 4.22 – Порівняння значень результатів тестової вибірки та прогнозованих даних

Було складено порівняльну матрицю результатів, отриманих у ході навчання прогнозованої моделі.

```
cols = ['Biopsy_1','Biopsy_0'] # реальні значення
rows = ['Prediction_1','Prediction_0'] # прогнозовані дані

B1P1 = len(df_ans[(df_ans['Prediction'] == df_ans['Biopsy']) & (df_ans['Biopsy'] == 1)])
B1P0 = len(df_ans[(df_ans['Prediction'] != df_ans['Biopsy']) & (df_ans['Biopsy'] == 1)])
B0P1 = len(df_ans[(df_ans['Prediction'] != df_ans['Biopsy']) & (df_ans['Biopsy'] == 0)])
B0P0 = len(df_ans[(df_ans['Prediction'] == df_ans['Biopsy']) & (df_ans['Biopsy'] == 0)])

conf = np.array([[B1P1,B0P1],[B1P0,B0P0]])
df_cm = pd.DataFrame(conf, columns = [i for i in cols], index = [i for i in rows])
```

На рисунку 4.23 представлено порівняльну матрицю результатів прогнозування та результатів тестової вибірки.

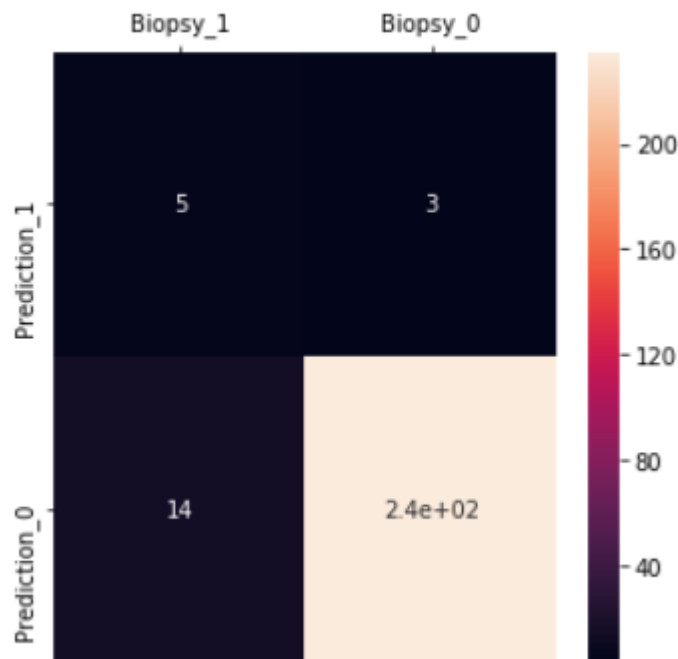


Рисунок 4.23 – Порівняльна матриця

Нижче представлено частину програмного коду з розрахунками параметрів sensitivity, specificity, false_positive_rate та false_negative_rate.

```
def model_efficacy(conf):
    total_num = np.sum(conf)
    sen = float(conf[0][0]/(conf[0][0]+conf[1][0]))
    spe = float(conf[1][1]/(conf[1][0]+conf[1][1]))
    false_positive_rate = conf[0][1]/(conf[0][1]+conf[1][1])
    false_negative_rate = conf[1][0]/(conf[0][0]+conf[1][0])

    #G = реальні значення (gold standard); P = прогнозовані значення (prediction)

    print('total_num: ',total_num)
    print('G1P1: ',conf[0][0])
    print('GOP1: ',conf[0][1])
    print('G1P0: ',conf[1][0])
    print('GOP0: ',conf[1][1])
    print('*****')
    print('sensitivity: ',sen)
    print('specificity: ',spe)
    print('false_positive_rate: ',false_positive_rate)
    print('false_negative_rate: ',false_negative_rate)

    return total_num, sen, spe, false_positive_rate, false_negative_rate

model_efficacy(conf)
```

Результати розрахунків представлені на рисунку 4.24.

```

total_num: 257
G1P1: 9
G0P1: 4
G1P0: 10
G0P0: 234
*****
sensitivity: 0.47368421052631576
specificity: 0.9590163934426229
false_positive_rate: 0.01680672268907563
false_negative_rate: 0.5263157894736842

Out[26]: (257,
          0.47368421052631576,
          0.9590163934426229,
          0.01680672268907563,
          0.5263157894736842)

```

Рисунок 4.24 – Результати розрахунків додаткових параметрів

Отже, значення показника частки фактичних позитивних випадків, які були прогнозовані як позитивні (або справжні позитивні) – sensitivity становить 47,36 %, частка фактичних негативів, які отримали прогноз як негативний (або істинний негатив) – specificity становить 95,90 %, значення коефіцієнту помилкових позитивних значень становить 1,6 %, а значення коефіцієнту помилкових негативних значень – 52,63 %.

Для того, щоб при запиті користувач анейромережа кожного разу не перенавчалася, то було виконано імпорт архітектури побудованої моделі (до json файлу) та її вагових коефіцієнтів (до файлу з розширенням «model»), які у подальшому будуть тільки завантажуватися до Python скрипту з прогнозуванням результатів оцінки ризику виявлення захворюваності раку шийки матки:

```

model_json = model.to_json()
with open("Model.json", "w") as json_file:
    json_file.write(model_json)
model.save_weights("Model_weights.model")

```


4.5 Створення web-додатку

Web-додаток було створено за допомогою програми Microsoft Visual Studio. Спочатку потрібно додати платформу розробки ASP.NET до Microsoft Visual Studio. Завантажуємо Microsoft Visual Studio Installer та обираємо «ASP.NET and web development» (рис. 4.25).

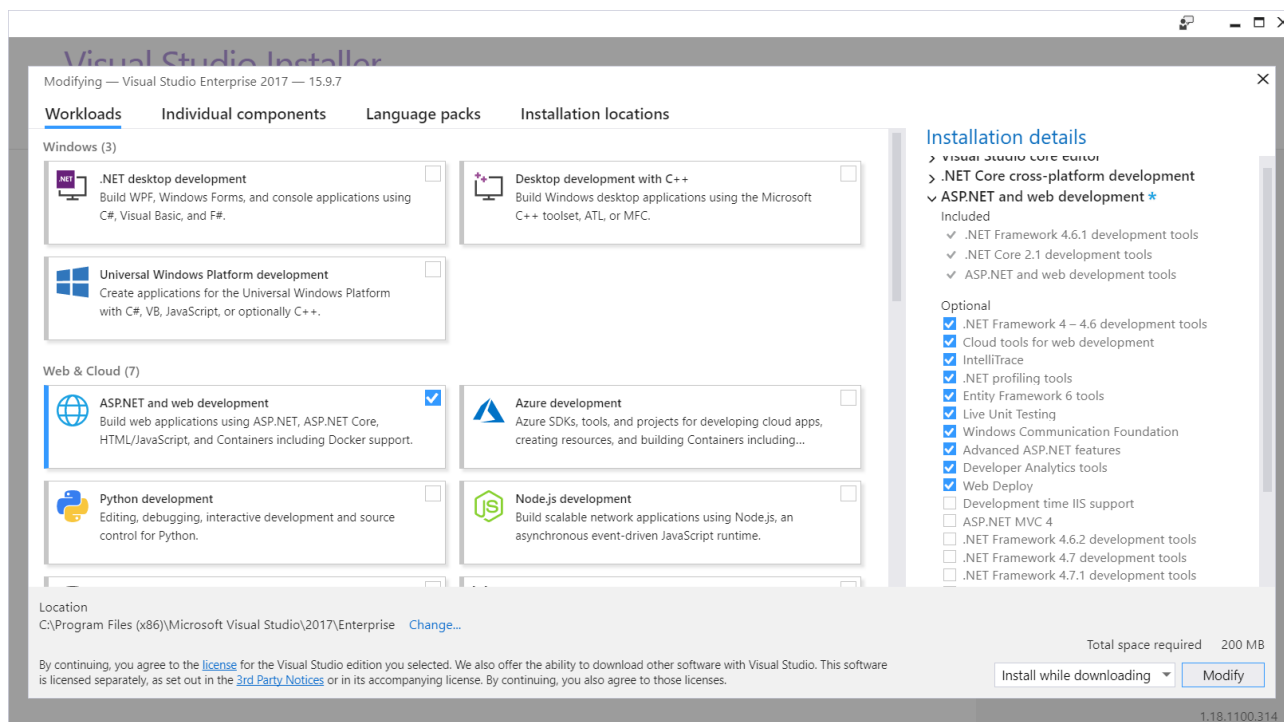


Рисунок 4.25 – Завантаження модулю «ASP.NET and web development»

Потім відкриваємо інструмент розробки MS Visual Studio та створюємо новий проект. Натискаємо на кнопку «File» – «New» – «Project...». У вікні «New Project» обираємо «.NET Core», тип додатку «ASP.NET Core Web Application», змінюємо назву додатку та локалізацію і натискаємо кнопку «ОК» (рис. 4.26).

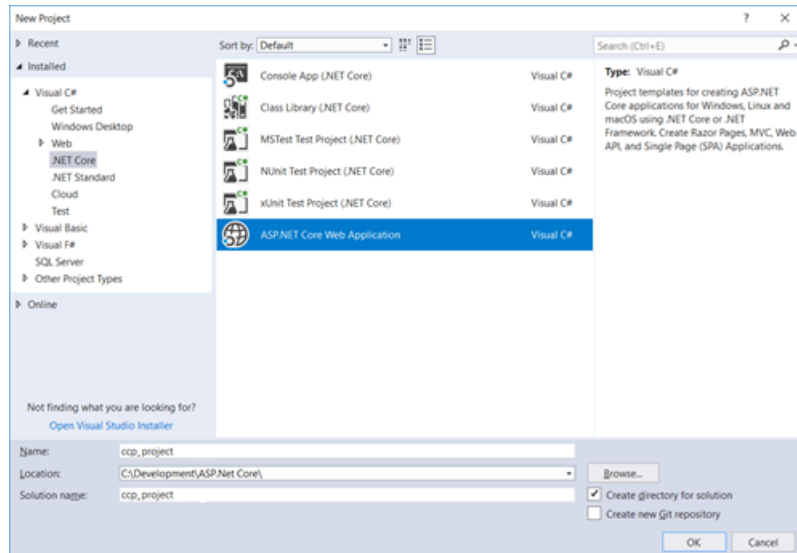


Рисунок 4.26 – Створення нового web-додатку

Далі обираємо тип додатку «Web Application (Model-View-Controller)» (рис. 4.27) та натискаємо кнопку «Change Authentication».

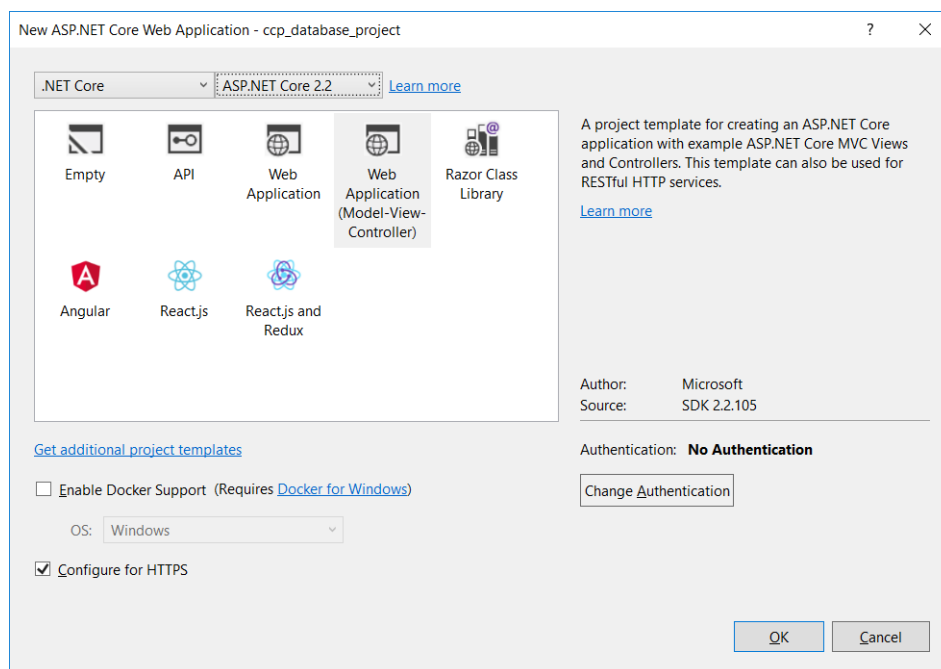


Рисунок 4.27 – Вибір типу додатку «Web Application (Model-View-Controller)»

У вікні «Change Authentication» обираємо пункт «Individual User Accounts» (рис. 4.28) та натискаємо кнопку «ОК».

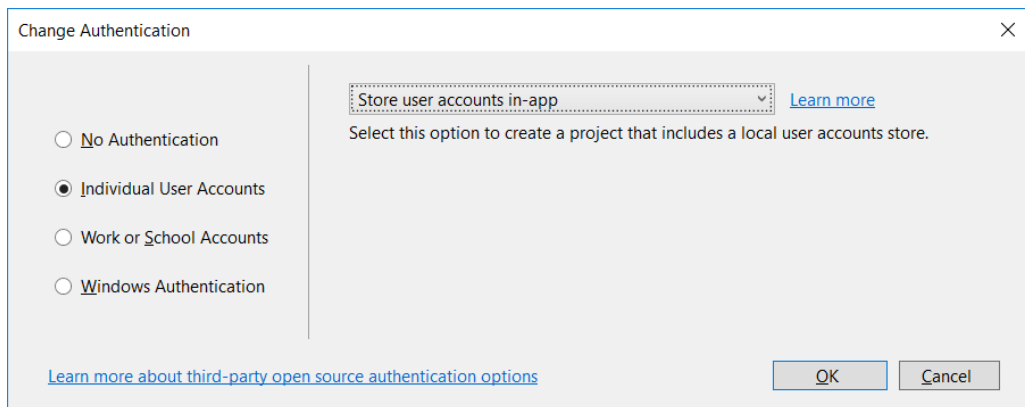


Рисунок 4.28 – Вибір типу авторизації для web-додатку

Таким чином було створено початкову версію проекту, яка зображена на рисунку 4.29.

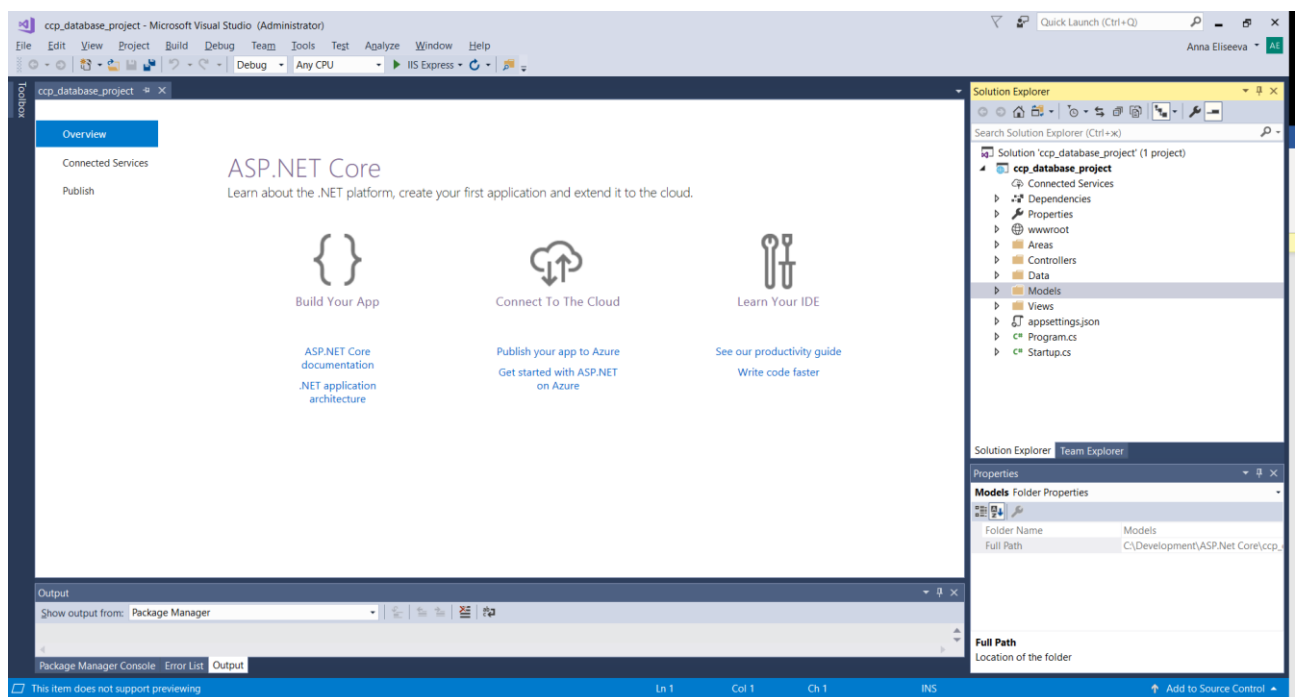


Рисунок 4.29 – Початкова версія web-додатку

4.5.1 Створення моделей (Models)

Моделі, які представлені у додатку виступають сутностями бази даних (рис. 4.30):

- DatasetModel – дані, які вводить користувач на сторінці створення запиту для оцінювання ризику виникнення захворюваності на цервікальний рак;
- PredictionBiopsyModel – результати прогнозування появи захворюваності;
- TrainDatasetModel – масив даних, призначений для навчання нейронної мережі;
- TestDatasetModel – масив даних, призначений для тренування нейронної мережі.

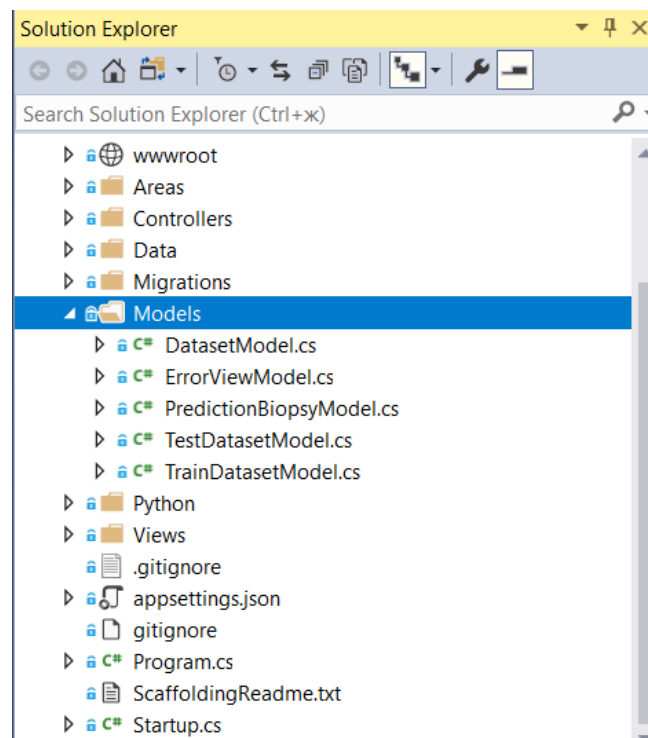


Рисунок 4.30 – Моделі web-додатку

Модель «DatasetModel» включає в себе 37 параметрів – індивідуальних показників організму користувача, які він додає при створенні запиту та ID для первинного ключа та зовнішнього ключа для користувача.

Моделі «TrainDatasetModel» та «TestDatasetModel» включають 48 параметрів – вже оброблені дані, які корегувалися за допомогою Python-скриптів та ID для первинного ключа.

Модель «PredictionBiopsyModel» включає 4 параметри: результати запиту, час виконання запиту, ID для первинного ключа та ID зовнішнього ключа для користувача.

Код програми для даних моделей зберігається у додатку В.

4.5.2 Створення представлення (Views)

Представлення, які були розроблені для web-додатку (рис. 4.31):

- Home;
- Shared;
- DatasetModels;
- PredictionBiopsyModels.

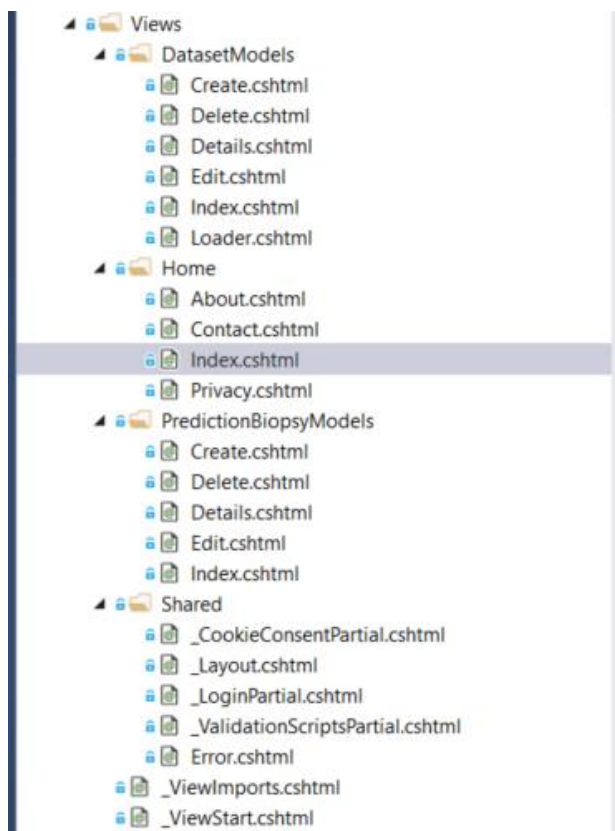


Рисунок 4.31 – Представлення web-додатку

У таблицях 4.2 – 4.5 описано основні представлення та завантажуючі сторінки їх функціоналу.

Таблиця 4.2 – Представлення «Home»

Сторінка	Опис
Index	Головна сторінка web-додатку
Contact	Сторінка для контакту з розробника
About	Сторінка з інформацією про розробника
Privacy	Сторінка з описом прав

Таблиця 4.3 – Представлення «Shared»

Сторінка	Опис
_Layout	Сторінка з основною розміткою для дочірніх сторінок web-додатку

Таблиця 4.4 – Представлення «DatasetModels»

Сторінка	Опис
Create	Сторінка для введення користувачем індивідуальних даних
Details	Сторінка для перегляду введених даних для запиту
Index	Сторінка для перегляду введених даних у вигляді таблиці
Loader	Сторінка для завантаження «progress bar» для запиту

Таблиця 4.5 – Представлення «PredictionBiopsyModels»

Сторінка	Опис
Delete	Видалення створеного запиту
Details	Перехід на сторінку Details представлення «DatasetModels» для перегляду введених даних для запиту

Продовження таблиці 4.5

Сторінка	Опис
Index	Сторінка для перегляду результатів у вигляді таблиці

4.5.3 Створення контролерів (Controllers)

Для функціонування web-додатку та керування моделями та представленнями було додано наступні контролери (рис. 4.32):

- HomeController;
- DatasetModelsController;
- PredictionBiopsyModelsController.

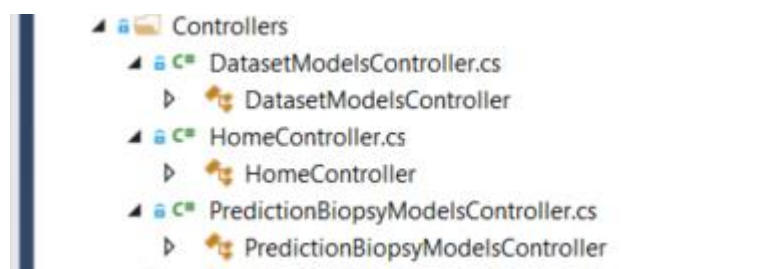


Рисунок 4.32 – Представлення web-додатку

У таблицях 4.6 – 4.8 описані основні методи контролерів, які вони виконують.

Таблиця 4.6 – Контролер «HomeController»

Метод	Опис
Index()	Метод для завантаження головної сторінки представлення «Home»
About()	Метод для завантаження сторінки «About» представлення «Home»
Contact()	Метод для завантаження сторінки «Contact» представлення «Home»

Продовження таблиці 4.6

Метод	Опис
Privacy()	Метод для завантаження сторінки «Privacy» представлення «Home»

Таблиця 4.7 – Контролер «DatasetModelsController»

Метод	Опис
Loader()	Метод для завантаження сторінки «Loader» представлення «DatasetModels»
Index()	Метод для завантаження сторінки «Index» представлення «DatasetModels»
Details()	Метод для завантаження сторінки «Details» представлення «DatasetModels»
Create()	Метод для завантаження сторінки «Create» представлення «DatasetModels»
DatasetModelExists()	Метод для перевірки існування запитів
ExecutePythonScript()	Метод для виконання Python скрипту з нейромережею

Таблиця 4.8 – Контролер «PredictionBiopsyModelsController»

Метод	Опис
Delete()	Метод для видалення результатів запиту
Index()	Метод для завантаження сторінки «Index» представлення «PredictionBiopsyModels»
Details()	Метод для завантаження сторінки «Details» представлення «PredictionBiopsyModels»
PredictionBiopsyModelExists()	Метод для перевірки існування результатів запитів

4.5.4 Підключення до бази даних

Підключення до бази даних PostgreSQL відбувається у файлі «appsettings.json». Код для створення з'єднання с базою даних «срр» представлений нижче:


```
"DefaultConnection": "Host=localhost;Port=5432;Database=ccp;User ID=postgres;Password=root"
```

Крім цього у файлі «Startup.cs» за допомогою методу «ConfigureServices()»

було підключено сервіси, що забезпечують з'єднання з базою даних PostgreSQL:

```
services.AddEntityFrameworkNpgsql().AddDbContext<ApplicationDbContext>(options =>
    options.UseNpgsql(Configuration.GetConnectionString("DefaultConnection"))
);

services.AddDefaultIdentity<IdentityUser>()
    .AddEntityFrameworkStores<ApplicationDbContext>();
```

4.5.5 Підключення зовнішніх модулів для авторизації користувача

Для авторизації користувача у web-додатку було налаштовано додаткові сервіси для більш комфортного та швидкого виконання даного процесу. Найбільш популярними додатками та обліковими записами, які використовують користувачі Інтернету є:

- Facebook;
- Google account;
- Microsoft account.

Саме дані сервіси були імплементовані у web-додаток. Для цього у файлі «Startup.cs» за допомогою методу «ConfigureServices()» було підключено вище вказані сервіси:

```
services.AddAuthentication()
    .AddFacebook(facebookOptions =>
    {
        facebookOptions.AppId = Configuration["Authentication:Facebook:AppId"];
        facebookOptions.AppSecret = Configuration["Authentication:Facebook:AppSecret"];
    })
    .AddGoogle(googleOptions =>
    {
        googleOptions.ClientId = Configuration["Authentication:Google:ClientId"];
        googleOptions.ClientSecret = Configuration["Authentication:Google:ClientSecret"];
    })
    .AddMicrosoftAccount(microsoftOptions =>
    {
        microsoftOptions.ClientId = Configuration["Authentication:Microsoft:ApplicationId"];
        microsoftOptions.ClientSecret = Configuration["Authentication:Microsoft:Password"];
    });
```

4.5.6 Валідація введених даних

Валідація даних забезпечується за допомогою реалізації моделей та додаткових JavaScript скриптів.

Обмеження діапазону допустимих введених даних користувачем реалізується у моделях за допомогою додаткових полів для їх атрибутів.

```
[Range(12, 100)]
public int Age { get; set; }
```

Дана команда забезпечує для поля віку введення даних з мінімальним значенням 12 років та максимальним віком – 100 років.

Так як деякі введені дані є залежними один від одного (наприклад, деякі поля для введення стають доступними при заповненні іншого параметру або вибору checkbox), то для перевірки правильного взаємозв'язку між ними було впроваджено додаткові JavaScript скрипти, які працюють у реальному часі:

- для перевірки вибору checkbox (true або false):

```
function IfCheckboxIsTrue1() {
var checkbox = document.getElementById("trigger1")
var hidden3 = document.getElementById("hidden_field3")
var hidden4 = document.getElementById("hidden_field4")

if (checkbox.checked == true) {
    hidden3.disabled = false;
    hidden4.disabled = false;
} else {
    hidden3.disabled = true;
    hidden4.disabled = true;
}
};
```

- перевірки кількості введених сексуальних партнерів:

```
function Bar() {
var sexpart = document.getElementById("sexpart").value;
var hidden1 = document.getElementById("hidden_field1")
var hidden2 = document.getElementById("hidden_field2")

if (sexpart > 0) {
    hidden1.disabled = false;
    hidden2.disabled = false;
} else {
    hidden1.disabled = true;
    hidden2.disabled = true;
}
}
```

4.5.7 Підключення Python-скрипту з нейронною мережею

Виконання обробки запиту користувача з прогнозуванням оцінки ризику появи захворюваності цервікального раку відбувається за допомогою підключеного Python скрипту. Він був доданий в окрему папку «Python» проекту, в якій розміщуються файли з основним скриптом, json-файлом архітектури та h5-файлом з налаштованими вагами попередньо навченої нейромережі (рис. 4.33).

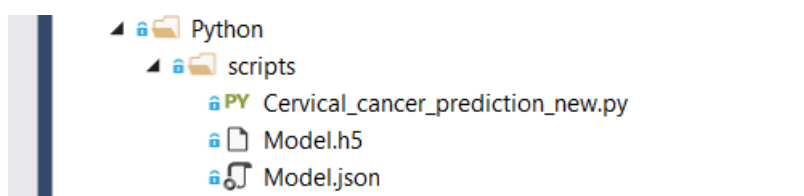


Рисунок 4.33 – Python-скрипт з навченою нейронною мережею

Розглянемо основні етапи роботи даного скрипту. Спочатку імпортуються всі необхідні бібліотеки для роботи команд. Потім встановлюється з'єднання з базою даних PostgreSQL «ссп», після чого виконується SQL запит для отримання значень введених параметрів користувача. Отримані дані проходять первинну обробку: перетворюються на числовий тип даних, налаштовуються значення для категоріальних змінних аналогічно до етапу обробки даних зазначених попередньо. Далі відбувається завантаження із зовнішнього json-файлу моделі нейронної мережі та з h5-файлу – налаштовані ваги:

```
json_file = open('Model.json', 'r')
loaded_model_json = json_file.read()
json_file.close()
loaded_model = model_from_json(loaded_model_json)
# load weights into new model
loaded_model.load_weights("Model.h5")
```

Потім відбувається компіляція нейромережі та виконання прогнозування:

```
loaded_model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
biopsy = loaded_model.predict_classes(data)
```

Отримані результати зберігаються в окремій змінній та за допомогою SQL запиту вивантажуються до таблиці «PredictionBiopsyModels» розробленої бази даних:

```
prediction.to_sql('PredictionBiopsyModels', con=engine, if_exists = 'append', index=False)
```

4.5.8 Міграції

Міграції – спеціальний функціонал середовища розробки ASP NET Core, за допомогою якого можна швидко працювати з базою даних, її таблицями та іншими операціями. Завдяки ним працювати з базами даних більш швидко та швидше, так як ASP NET Core самостійно формує необхідні файли з вже написаним програмним кодом для створення таблиць у тазах даних. Після створення моделей у проекті було виконано декілька міграцій, які представлені на рисунку 4.34.

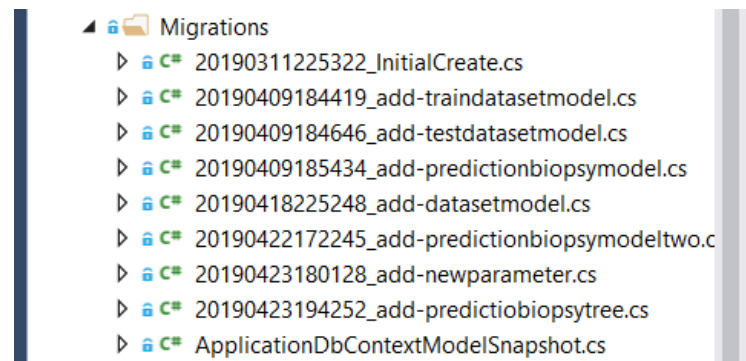


Рисунок 4.34 – Міграції проекту

Для того, щоб виконати міграцію у проєкті MS Visual Studio потрібно перейти на вкладку «Tool» – «NuGet Package Manager» – «Package Manager Console». Для того, щоб дозволити проєкту працювати з міграціями було виконано команду «enable-migrations». Для того, щоб додати першу міграцію у консолі було виконано наступну команду «Add-Migration InitialCreate». Далі було додано міграцію зі створенням таблиці для збереження масиву даних навчальної вибірки. У провіднику рішень було знайдено теку «Data», відкрито файл «ApplicationDbContext.cs» та у

клас «ApplicationDbContext : IdentityDbContext» додано команду зі створення тіблиці «TrainDatasetModels»:

```
public DbSet<TrainDatasetModel> TrainDatasetModels { get; set; }
```

Далі у консолі було виконано наступні дві команди:

```
Add-Migration add-traindatasetmodel
Update-Database
```

Перша з них створила міграцію «add-traindatasetmodel», у результаті чого в проєкті з'явився файл «20190409184419_add-traindatasetmodel.cs» (рис. 4.35)

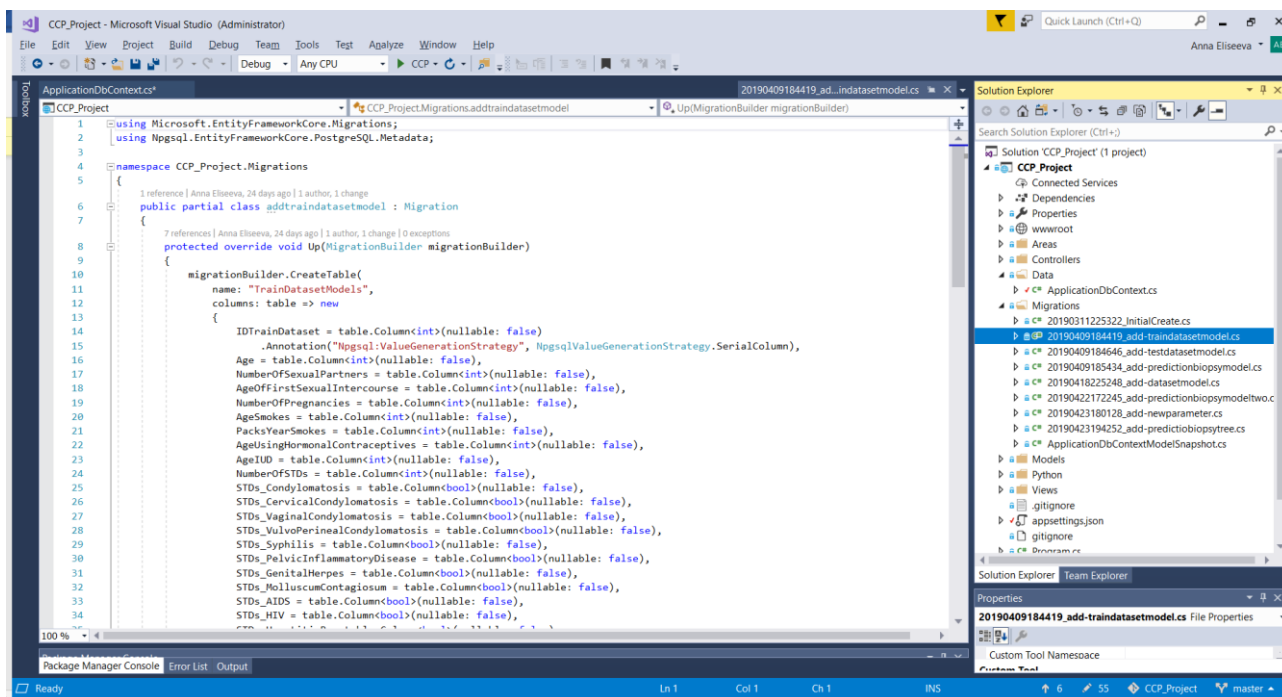


Рисунок 4.35 – Міграція «add-traindatasetmodel»

Друга команда виконує оновлення бази даних на сервері бази даних, у результаті чого створюється таблиця «TrainDatasetModels» (рис. 4.36).

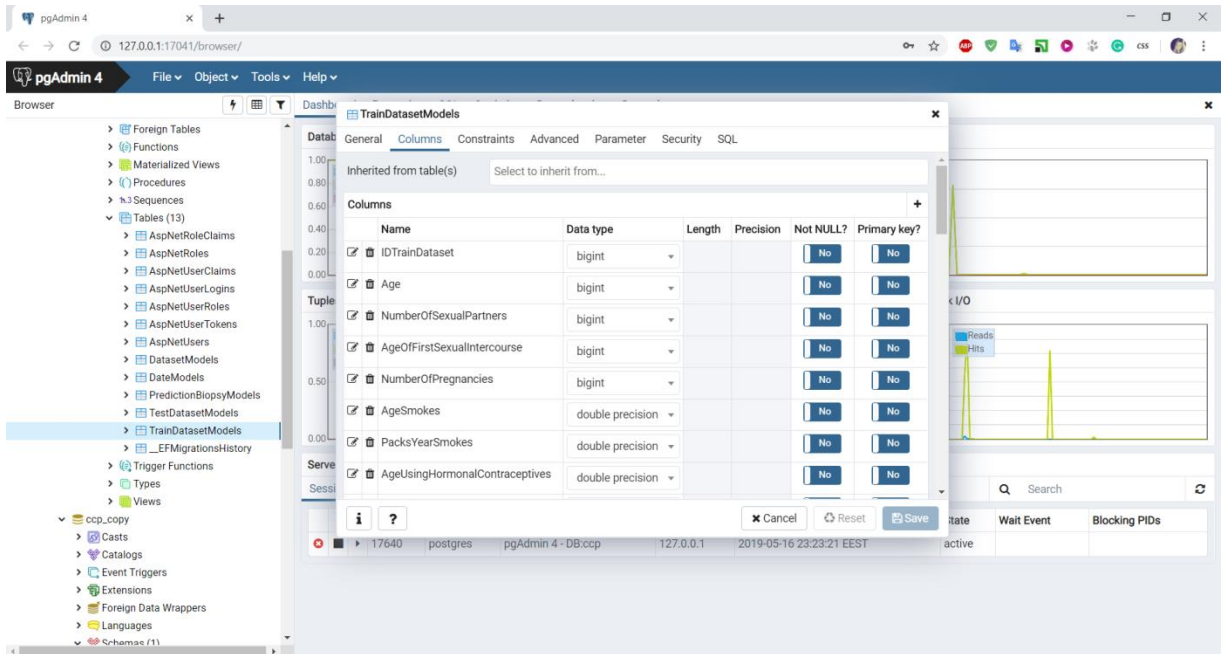


Рисунок 4.36 – Таблиця «TrainDatasetModels» бази даних «срр»

4.5.9 Модулі проекту

Для того, щоб проект працював було завантажено спеціальні функціональні модулі, які забезпечують виконання команд для ASP NET Core та підключення до серверу баз даних PostgreSQL (рис. 4.37).

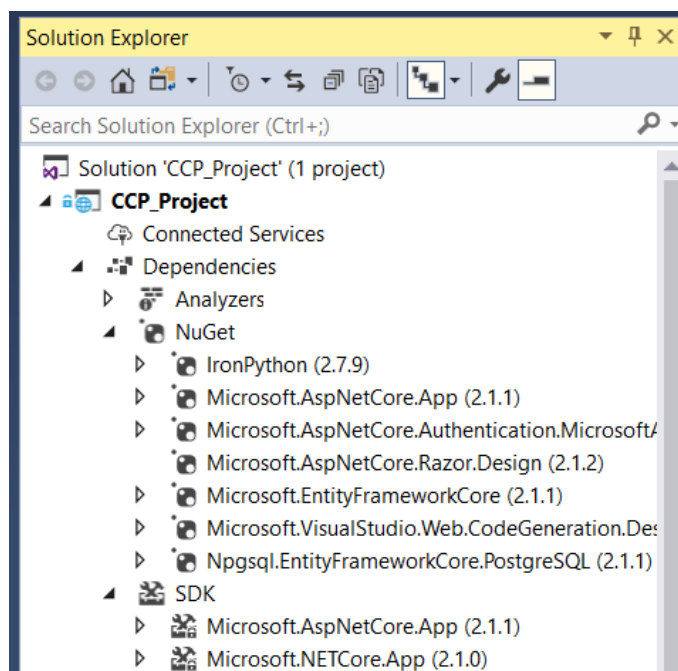


Рисунок 4.37 – Підключені модулі проекту

4.6 Реалізація роботи web-додатку

Після запуску проекту завантажується головна сторінка web-додатку, на якій збережена основна інформація про цервікальний рак, а саме (рис. 4.38):

- що таке рак шийки матки;
- які симптоми виникають в організмі при його появі;
- яким чином виконується діагностика;
- якими способами цього можна вилікувати.

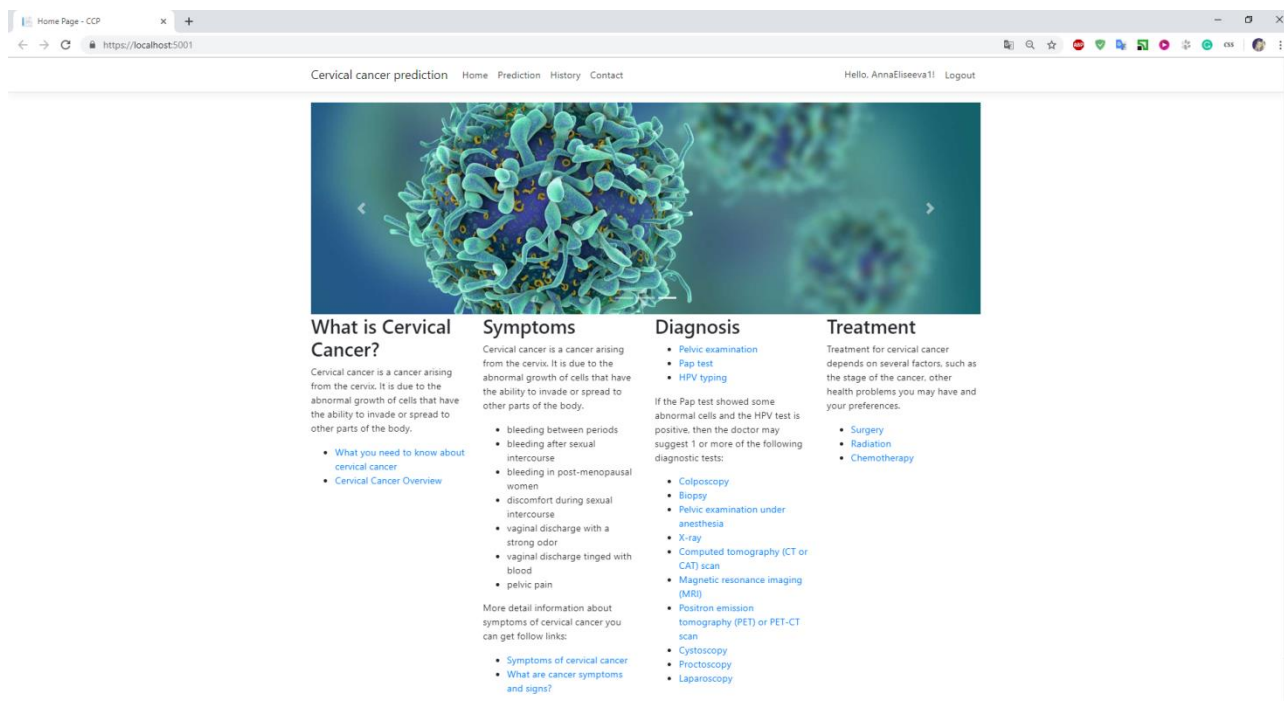


Рисунок 4.38 – Головна сторінка web-додатку

Було розроблено чотири вкладки:

- «Home»;
- «Prediction»;
- «History»;
- «Contact».

На вкладці «Prediction» користувач може ввести свої індивідуальні показники та створити запит для розрахунку оцінки ризику захворюваності цервікального раку. Створимо запит та перевіримо роботу програми. Введемо дані та перевіримо їх валідацію. На рисунку 4.39 зображено введення некоректного віку користувача.

The screenshot shows a web browser window with the URL <https://localhost:5001/DataSetModels/Create>. The page title is "Cervical cancer prediction" and the subtitle is "Individual characteristics". The form contains several sections with input fields:

- Age:** A text input field containing the value "11". Below it, a red error message reads: "The field Age must be between 12 and 100."
- Existing of sexual intercourse:** Includes "Number of sexual partners" (input: 0), "Age of first sexual intercourse" (input: 0), and "Number of pregnancies" (input: 0).
- Smoking:** Includes "Do you smoke?" (checkbox), "How many years do you smoke?" (input: 0), and "How many packs of cigarettes do you smoke per year?" (input: 0).
- Using hormonal contraceptives:** Includes "Do you use the hormonal contraceptives?" (checkbox) and "How many years do you use the hormonal contraceptives?" (input: 0).
- Intrauterine device:** Includes "Do you use the intrauterine device?" (checkbox) and "How many years do you use the intrauterine device?" (input: 0).
- Sexually transmitted diseases:** Includes "Do you have sexually transmitted diseases?" (checkbox), "Do you have cervical condylomatosis?" (checkbox), "How many sexually transmitted disease diagnoses have been determined?" (input: 0), "Do you have condylomatosis?" (checkbox), and "Do you have vulvo perineal condylomatosis?" (checkbox).

Рисунок 4.39 – Некоректний вік користувача

На рисунках 4.40 – 4.41 зображено введені користувачем дані та створення запиту.

The screenshot shows the same web browser window as Figure 4.39, but with valid data entered into the form:

- Age:** Input field contains "35".
- Existing of sexual intercourse:** "Number of sexual partners" is "5", "Age of first sexual intercourse" is "19", and "Number of pregnancies" is "2".
- Smoking:** "Do you smoke?" is checked. "How many years do you smoke?" is "2", and "How many packs of cigarettes do you smoke per year?" is "103".
- Using hormonal contraceptives:** "Do you use the hormonal contraceptives?" is checked. "How many years do you use the hormonal contraceptives?" is "4".
- Intrauterine device:** "Do you use the intrauterine device?" is checked. "How many years do you use the intrauterine device?" is empty.
- Sexually transmitted diseases:** "Do you have sexually transmitted diseases?" is checked. "Do you have cervical condylomatosis?" is checked. "How many sexually transmitted disease diagnoses have been determined?" is empty. "Do you have condylomatosis?" is checked. "Do you have vulvo perineal condylomatosis?" is checked. "Do you have syphilis?" is checked. "Do you have vaginal condylomatosis?" is checked. "Do you have genital herpes?" is checked.

Рисунок 4.40 – Введення даних користувачем

The screenshot shows a web browser window with the URL `https://localhost:5001/DatasetModel/Create`. The page contains a form for creating a dataset. The form is divided into several sections:

- Intrauterine device:** A question 'Do you use the intrauterine device?' with a radio button. A text input field contains the number '4'.
- Sexually transmitted diseases:** Multiple questions with radio buttons, including 'Do you have sexually transmitted diseases?', 'Do you have cervical condylomatosis?', 'Do you have syphilis?', 'Do you have molluscum contagiosum?', 'Do you have hepatitis B?', 'How much time has passed since the first diagnosis?', 'How many sexually transmitted disease diagnoses have been determined?', 'Do you have vaginal condylomatosis?', 'Do you have pelvic inflammatory disease?', 'Do you have AIDS (acquired immune deficiency syndrome)?', 'Do you have HPV (human papillomavirus)?', 'How much time has passed since the last diagnosis?', 'Do you have condylomatosis?', 'Do you have vulvo perineal condylomatosis?', 'Do you have genital herpes?', 'Do you have HIV (human immunodeficiency viruses)?', and 'Enter number of sexually transmitted diagnosis'.
- Diagnosis:** Questions with radio buttons: 'Have you ever been diagnosed with cancer?', 'Have you ever been diagnosed with HPV (human papillomavirus)?', and 'Have you ever been diagnosed with CIN (cervical intraepithelial neoplasia)?'.
- Clinical tests:** Questions with radio buttons: 'Do you have a positive Hinselmann test result?', 'Do you have a positive Schiller test result?', and 'Do you have a positive cytology result?'.

A 'Create' button is located at the bottom right of the form. A 'Back to List' link is at the bottom left. The footer contains the text '© 2019 - Created by Anna Eliseeva - Privacy'.

Рисунок 4.41 – Введення даних користувачем

Після створення запиту виконується його обробка, яка супроводжується завантажуючою стрічкою, для того, щоб користувач розумів, на якому етапі відбувається прогнозування (рис. 4.42).

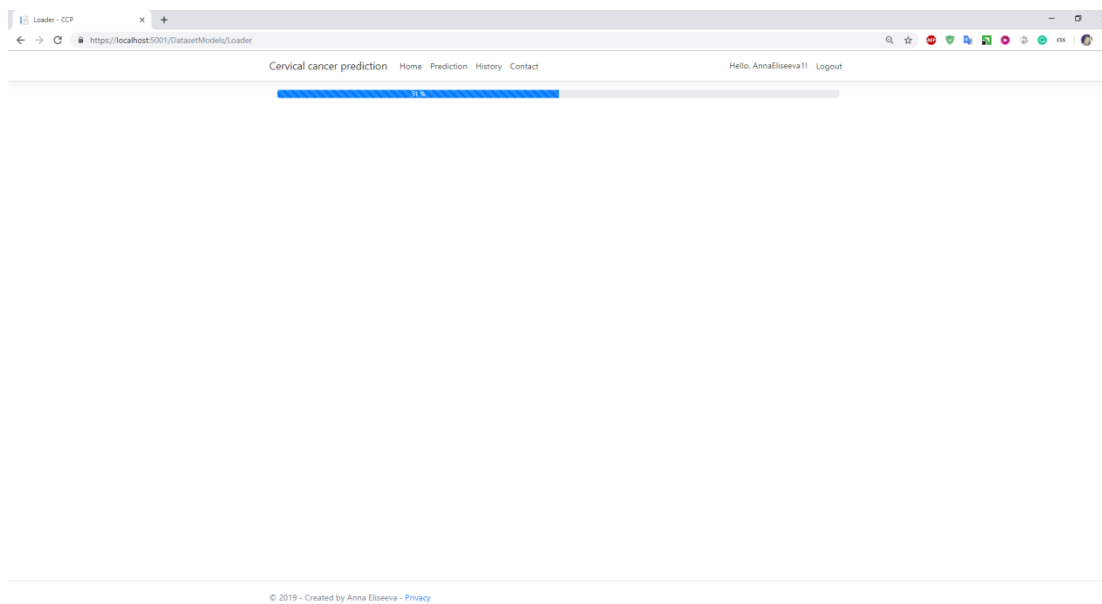


Рисунок 4.42 – Обробка запиту

На рисунку 4.43 зображено результат прогнозування (вкладка «History»), який є негативним, а також дату та час, коли запит було виконано. Користувач має змогу

видалити результати цього запиту (кнопка «Delete») та детально переглянути, які дані ним були введені при виконанні прогнозування (кнопка «Details»).

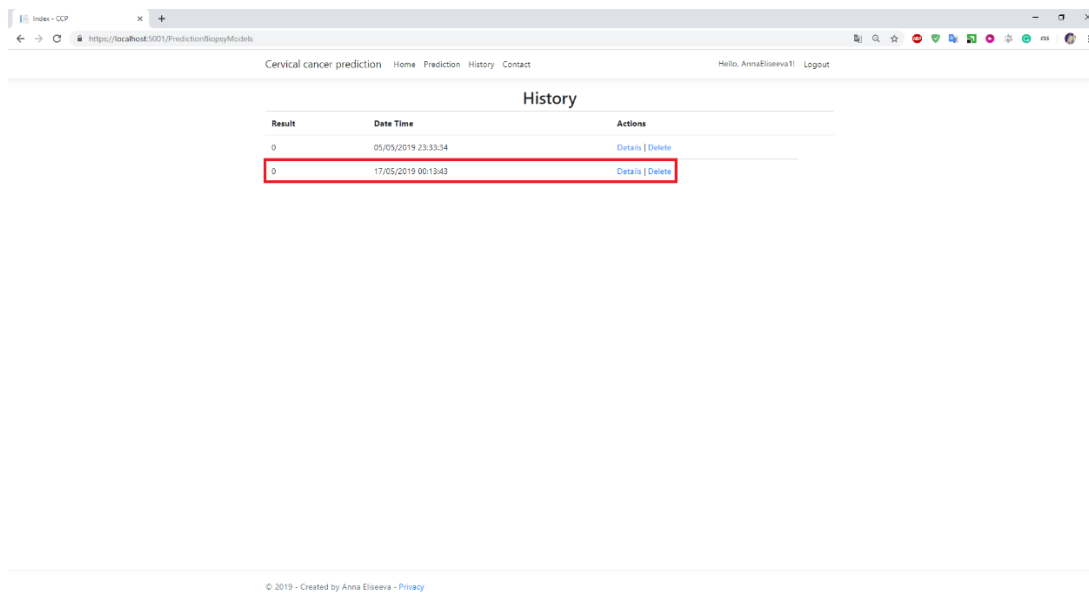


Рисунок 4.43 – Результат запиту

Переглянемо, які показники були введені користувачем за допомогою кнопки «Details» (рис. 4.44).

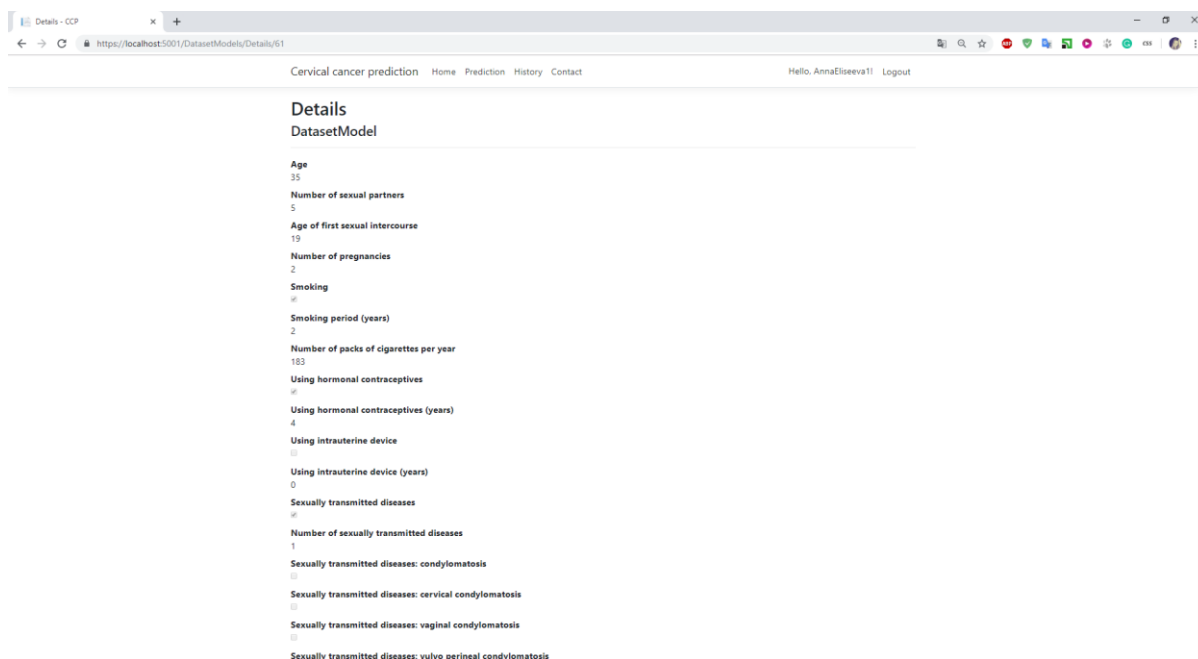


Рисунок 4.44 – Перегляд введених користувачем даних у запиті

Видалимо один із оброблених запитів. Завантажується сторінка з підтвердженням видалення користувацьких даних (рис. 4.45)

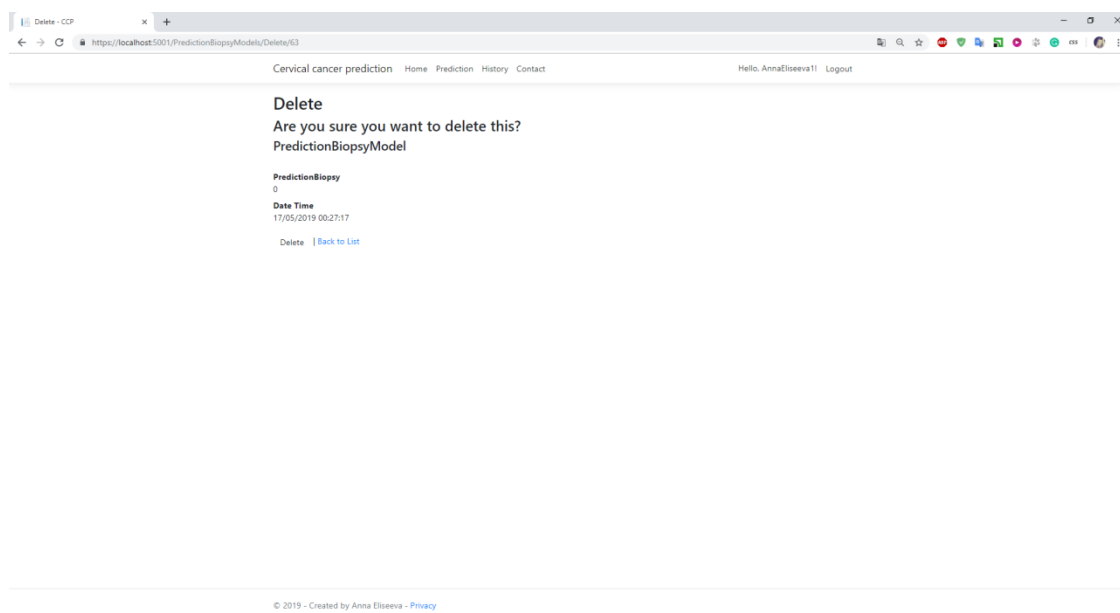


Рисунок 4.45 – Видалення даних

На рисунку 4.46 представлено вкладку «Contact», на якій відображається інформація про розробника web-додатку.

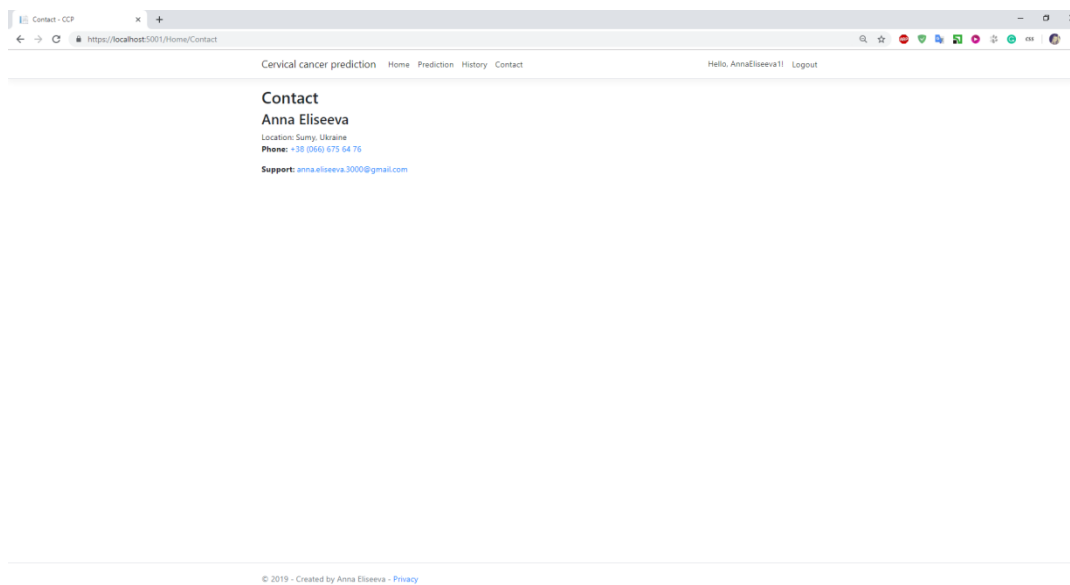


Рисунок 4.46 – Інформація про розробника

Крім цього була реалізована аторизація користувача та реєстрація. Якщо на головній сторінці додатку натиснути на кнопку «Login», то завантажується сторінка для авторизації користувача (рис. 4.47). Якщо користувач раніше був зареєстрований у системі, то його дані підтягуються атоматично або користувач може здійснити авторизацію за допомогою власних акаунтів Facebook, Google або Microsoft при їх наявності (рис. 4.48 – 4.49).

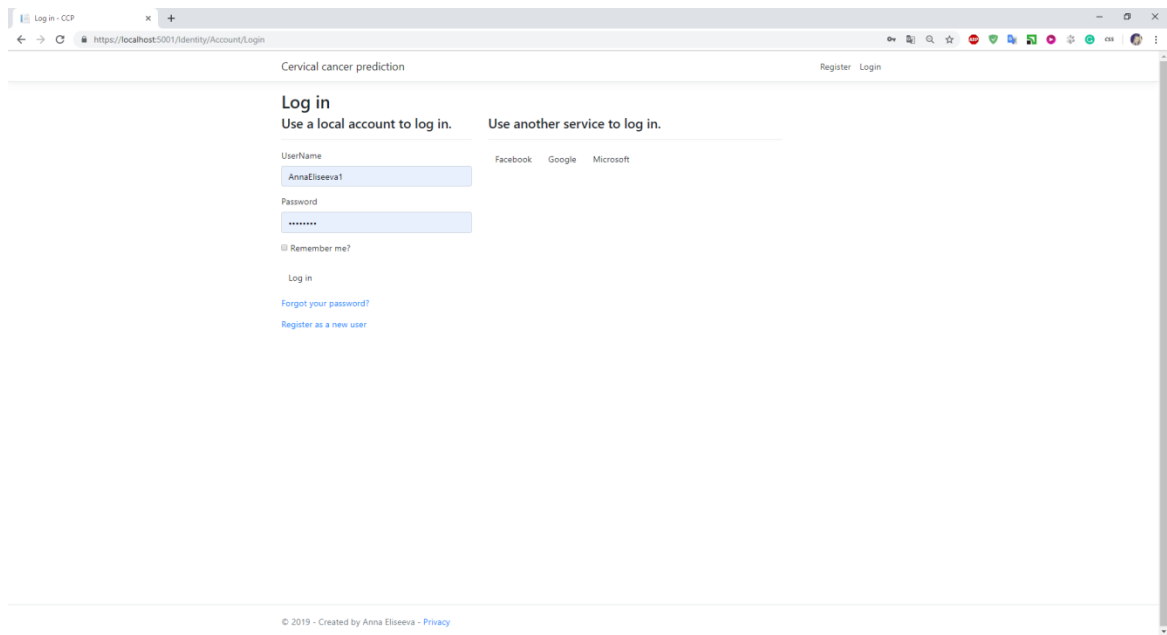


Рисунок 4.47 – Авторизація користувача

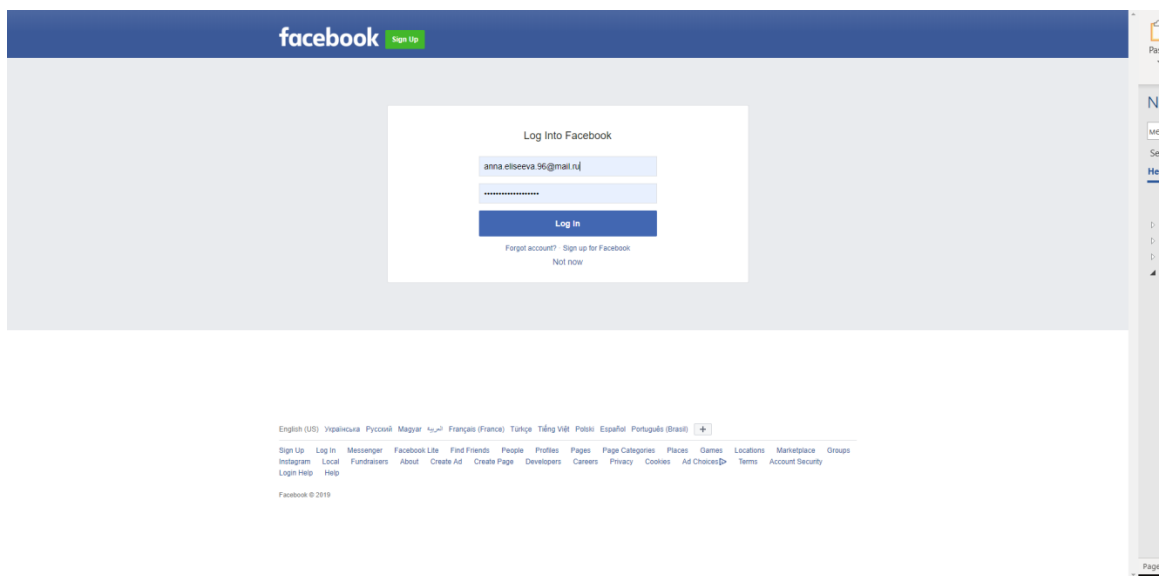


Рисунок 4.48 – Авторизація користувача за допомогою Facebook

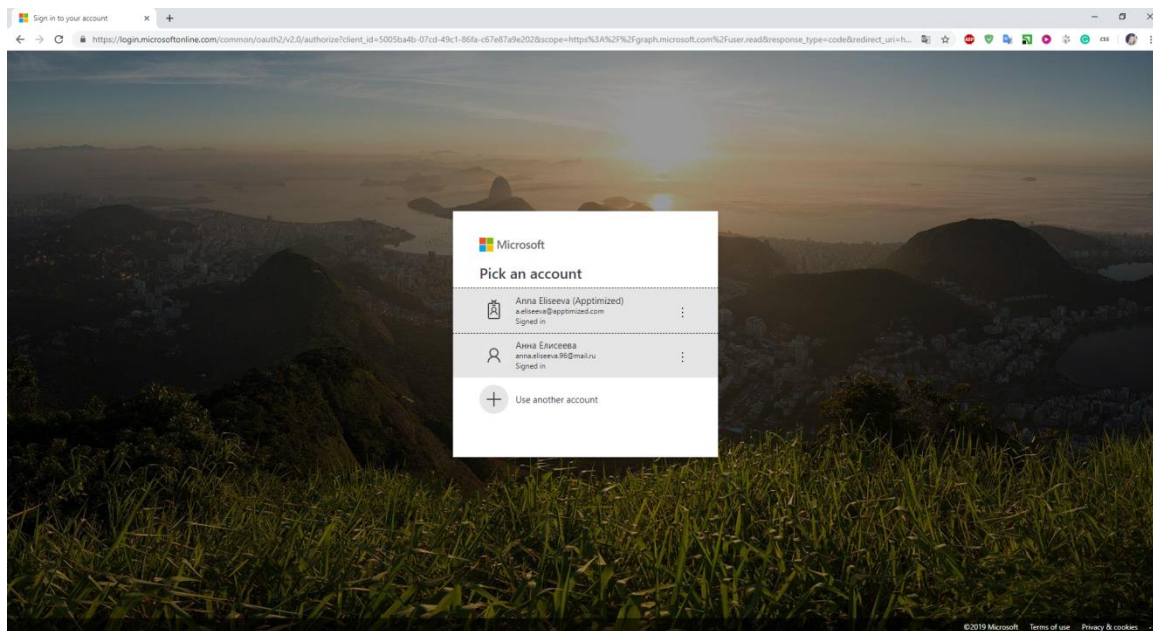


Рисунок 4.49 – Авторизація користувача за допомогою Microsoft

У результаті виконання даного етапу була описана розробка інформаційної технології. Було зазначено як оброблювався масив початкових даних, який використовувався для навчання нейронної мережі, було описано як створювалася база даних у PostgreSQL, як здійснювався імпорт даних до бази даних. Також було зазначено як виконувалася розробка нейромережі, створення web-додатку за допомогою програмного коду, реалізованого за допомогою середовища розробки ASP NET Core та показано реалізацію роботи готового програмного продукту.

ВИСНОВОК

У сучасній медицині широке застосування мають інформаційні технології при діагностиці і лікуванні різних захворювань. Основною задачею створення таких систем є поліпшення якісних показників діагностики та лікування. Тому роботи, спрямовані на пошук нових рішень при створенні таких систем, є актуальними. Незважаючи на всі плюси нейронних мереж, зустрічається досить велика кількість проблем при їх впровадженні в медицину. Це пояснюється тим, що нейронна мережа не дає ніяких документальних обґрунтувань того, чому вона вважає саме цей діагноз правильним.

Нейронні мережі представляють собою нелінійні системи, що дозволяють краще класифікувати дані, ніж звичайні лінійні методи. У системі для медичної діагностики вони дають можливість значно підвищити специфічність методу, не знижуючи його чутливості.

У зв'язку зі складністю діагностування онкологічних захворювань, а саме раку шийки матки, а також визначення ефективного методу лікування в залежності від стадії його розвитку з можливістю прогнозування повторної появи метастаз була визначена необхідність створення відповідної інформаційно-діагностичної технології.

Задачі, які було вирішено у ході виконання наукової роботи:

- виконано огляд сучасних методів автоматизованого аналізу і обробки статистичних даних;
- було обрано метод розпізнавання ракових пухлин, заснованого на нейронних мережах;
- було перевірено кореляцію різних факторів на результати оцінки ризику виникнення ракового захворювання;
- було побудовано нейронну мережу з відповідними параметрами для визначення результату про наявність чи відсутність раку шийки матки;

- виконано оцінку якості навчання нейронної мережі;
- розроблено інформаційну технологію з інтуїтивним інтерфейсом для зручного та швидкого доступу для виконання оціночного аналізу.

Для вирішення задачі аналізу ризиків та прогнозування розвитку хвороби використовують гнучкі нейронні мережі. На першому етапі за переліком ознак виконується задача класифікації діагностування раку шийки матки, як класифікація за допомогою штучних нейронних мереж типу багатошарового персептрону. Усі отримані дані зберігаються у реляційній базі даних – об’єктно-реляційній системі управління базами даних PostgreSQL. Початкова обробка даних, навчання, тренування та тестування розроблюваної нейронної мережі виконується за допомогою високорівневої мови програмування Python з використанням вбудованих бібліотек для машинного навчання. Для зручного представлення необхідної інформації було створено веб-додаток, який на основі заданих основних 36 параметрів, щодо поточного стану жінки, за допомогою навченої нейронної мережі обробляє отримані дані та прогнозує можливу оцінку розвитку цервіклярного раку.

СПИСОК ЛІТЕРАТУРИ

- 1 1 Dimitrios S. Neural Networks in Medicine [Електронний ресурс] / Siganos Dimitrios – Режим доступу: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol2/ds12/article2.html.
- 2 Ежов А. Нейронные сети в медицине / А. Ежов, В. Чечеткин. // Открытые системы архитекторам информационных систем. – 1997.
- 3 Прорыв в онкологии: нейронную сеть научили распознавать раковые клетки [Електронний ресурс] – Режим доступу: <https://rossaprimavera.ru/news/6d1ebc7d>.
- 4 Малышевская Е. Н. Системні дослідження та інформаційні технології / Е. Н. Малышевская // Анализ использования нейросетей для диагностики рака шейки матки по мультиспектральному изображению / Е. Н. Малышевская., 2010.
- 5 Смекалов В. Модель нейронной сети для прогнозирования результатов лечения предраковых заболеваний шейки матки / В. Смекалов, А. Сепп // Проблемы и перспективы современной науки / В. Смекалов, А. Сепп., 2008.
- 6 Classification of Cervical Cancer using Artificial Neural Networks / D.Anousouya, S. Ravi, J. Vaishnavi, S. Punitha // Twelfth International Multi-Conference on Information Processing-2016 / D.Anousouya, S. Ravi, J. Vaishnavi, S. Punitha., 2016. – С. 465–472.
- 7 Guiding the management of cervical cancer with convolutional neural networks [Електронний ресурс]. – 2019. – Режим доступу: <http://cs231n.stanford.edu/reports/2017/pdfs/920.pdf>.
- 8 Improvement of Features Extraction Process and Classification of Cervical Cancer for the NeuralPap System / S.Sulaimana, N. Mat-Isab, N. Othmanc, F. Ahmada // 19th International Conference on Knowledge Based and Intelligent Information and Engineering / S.Sulaimana, N. Mat-Isab, N. Othmanc, F. Ahmada., 2015.
- 9 Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems / H. A.Almubarak, J. R. Stanleya, R. Longb, S. Antanib., 2017.
- 10 Malli P. Machine learning Technique for detection of Cervical Cancer using k-NN and Artificial Neural Network / P. Malli, S. Nandyal // International Journal of

Emerging Trends & Technology in Computer Science (IJETTCS) / P. Malli, S. Nandyal., 2017. – С. 145–149.

11 Constructing of the risk classification model of cervical cancer by artificial neural network / X.Qiu, N. Tao, Y. Tan, X. Wu // Expert Systems with Applications / X.Qiu, N. Tao, Y. Tan, X. Wu., 2007. – С. 1094–1099.

12 Matsuo K. Survival outcome prediction in cervical cancer: Cox models versus deep-learning model / K. Matsuo, S. Purushotham, P. Jiang // American Journal of Obstetrics and Gynecology / K. Matsuo, S. Purushotham, P. Jiang., 2018.

13 Факторы риска рака шейки матки [Электронный ресурс] – Режим доступа: <https://kasvaja.net/рак-шейки-матки/фактори-риска-рака-шейки-матки/>.

14 Семь факторов риска рака шейки матки, о которых надо знать [Электронный ресурс]. – 2018. – Режим доступа: <https://krasotka.postimees.ee/4423785/sem-faktorov-riska-raka-sheyki-matki-о-kotoryh-nado-znat>.

15 Смерних О. Факторы риска рака шейки матки. Псевдоэрозия. [Электронный ресурс] / О. Смерних – Режим доступа: <http://www.genskie-nogki.ru/health/bolezni-ginekologii/factory-riska-raka-shjejki-matki-psjevdoerozija.php>.

16 Факторы риска развития рака шейки матки [Электронный ресурс] – Режим доступа: <http://www.rusmedserv.com/cervicalcancer/risk-factors/>.

17 Starikov A. Применение нейронных сетей для задач классификации [Электронный ресурс] / Starikov – Режим доступа: <https://basegroup.ru/community/articles/classification>.

18 A Beginner's Guide to Multilayer Perceptrons (MLP) [Электронный ресурс] – Режим доступа: <https://skymind.ai/wiki/multilayer-perceptron>.

19 Нейросетевое моделирование: многослойный персептрон [Электронный ресурс] – Режим доступа: <http://www.ievbras.ru/ecostat/Kiril/Library/Book1/Content394/Content394.htm>.

20 Jason B. Binary Classification Tutorial with the Keras Deep Learning Library [Электронный ресурс] / Jason. – 2016. – Режим доступа: <https://machinelearningmastery.com/binary-classification-tutorial-with-the-keras-deep-learning-library/>.

21 Jason B. How to Choose Loss Functions When Training Deep Learning Neural Networks [Электронный ресурс] / Brownlee Jason. – 2019. – Режим доступа: <https://machinelearningmastery.com/how-to-choose-loss-functions-when-training-deep-learning-neural-networks/>.

22 Usage of loss functions [Электронный ресурс] – Режим доступа: <https://keras.io/losses/>.

23 Binary crossentropy [Электронный ресурс] – Режим доступа: <https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/loss-functions/binary-crossentropy>.

24 Rick A. Введение в ASP.NET Core [Электронный ресурс] / Rick. – 2019. – Режим доступа: <https://docs.microsoft.com/ru-ru/aspnet/core/?view=aspnetcore-2.2>.

25 Rick A. Выбор между ASP.NET 4.x и ASP.NET Core [Электронный ресурс] / Rick. – 2018. – Режим доступа: <https://docs.microsoft.com/ru-ru/aspnet/core/fundamentals/choose-aspnet-framework?view=aspnetcore-2.2#framework-selection>.

26 What is PostgreSQL? [Электронный ресурс] – Режим доступа: <https://www.postgresql.org/about/>.

27 Светлова И. Что такое PostgreSQL? Плюсы и минусы бесплатной базы данных [Электронный ресурс] / Ирина Светлова. – 2018. – Режим доступа: <https://oracle-patches.com/common/3214-что-такое-postgresql>.

28 "What is Python? Executive Summary" Python Software Foundation [Электронный ресурс] – Режим доступа: <https://www.python.org/doc/essays/blurbs/>.

29 Антонов А. Язык программирования Python и пакеты для машинного обучения и Data Mining [Электронный ресурс] / А. Антонов. – 2014. – Режим доступа: <http://robotosha.ru/algorithm/python-packages-for-machine-learning-and-data-mining.html>.

30 Git для начинающих. Часть 1. Что такое системы контроля версий? [Электронный ресурс]. – 2017. – Режим доступа: <https://devpractice.ru/git-for-beginners-part-1-what-is-vcs/>.

31 Введение - Основы Git [Электронный ресурс] – Режим доступа: <https://git-scm.com/book/ru/v1/Введение-Основы-Git>.

32 Знакомство с нотацией IDEF0 и пример использования [Електронний ресурс] – Режим доступу: <https://habr.com/ru/company/trinion/blog/322832/>.

33 Моделі потоків даних (DFD-моделі): призначення, місце застосування в системному аналізі, правила побудови, приклади [Електронний ресурс]. – 2015. – Режим доступу: <https://studopedia.org/8-191272.html>.

34 Методологія dfd [Електронний ресурс]. – 2016. – Режим доступу: <https://studfiles.net/preview/5609405/page:4/>.

35 Information Integration for Concurrent Engineering (IICE) IDEF3 Process Description Capture Method Report [Електронний ресурс]. – 2019. – Режим доступу: https://www.researchgate.net/profile/Christopher_Menzel2/publication/215439476_Information_Integration_for_Concurrent_Engineering_IICE_IDEF3_Process_Description_Capture_Method_Report/links/00b7d534717522ba41000000/Information-Integration-for-Concurrent-En.

36 Архитектура системы [Електронний ресурс] – Режим доступу: https://ru.wikipedia.org/wiki/Архитектура_системы, вільний.

37 Діаграма прецедентів [Електронний ресурс] – Режим доступу: https://uk.wikipedia.org/wiki/Діаграма_прецедентів.

38 Єфремов М. Ф. ПРОЕКТУВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ З ВИКОРИСТАННЯМ UML [Електронний ресурс] / М. Ф. Єфремов, Ю. М. Єфремов, В. М. Єфремов // Житомирський державний технологічний університет – Режим доступу до ресурсу: <https://conf.ztu.edu.ua/wp-content/uploads/2016/06/12.pdf>.

39 Діаграма послідовності [Електронний ресурс] // Национальный технический университет Украины «Киевский политехнический институт». – 2015. – Режим доступу до ресурсу: <https://studfiles.net/preview/3751299/page:10/>.

40 Об'єктно-орієнтоване моделювання програмних систем. Навчальний посібник – Львів: Видавничий центр ЛНУ імені Івана Франка, 2007. – 108 с.

41 Етап концептуального проектування. Основні поняття концептуального проектування. Концептуальне проектування. Об'єкти і їх властивості. Взаємовідношення об'єктів. [Електронний ресурс] // Каменец-Подольський національний університет ім. І. Огиенко. – 2016. – Режим доступу до ресурсу: <https://studfiles.net/preview/5454386/page:5/>.

42 Розробка бази даних - База знань Allbest [Електронний ресурс] – Режим доступу:

https://knowledge.allbest.ru/programming/2c0a65635b3bd79b4d43b89421216d27_0.html.

43 Моделі баз даних [Електронний ресурс] – Режим доступу до ресурсу: www.100balov.com/data/ukr/IInfo_dlya_stydena__21/1_4393.doc.

44 Етапи проектування [Електронний ресурс] – Режим доступу до ресурсу: http://wiki.fizmat.tnpu.edu.ua/index.php/Етапи_проектування.

45 Enhanced entity–relationship model [Електронний ресурс] – Режим доступу: <http://www.cs.toronto.edu/~jm/2507S/Notes04/EER.pdf>.

46 Проектування бази даних [Електронний ресурс] – Режим доступу до ресурсу:

https://elearning.sumdu.edu.ua/free_content/lectured:89b3d175c06a6b137e410cb14821d0e94549ad5a/latest/44233/index.html.

47 Етапи проектування та моделювання БД [Електронний ресурс] – Режим доступу до ресурсу: <http://ua.textreferat.com/referat-23369-3.html>.

48 Толстохатько В. А. Бази даних: проектування та використання для обліку нерухомого майна. Навч. посібник / В. А. Толстохатько, О. С. Поморцева, І. М. Патракеєв. – Харків: Х. : ХНУМГ, 2014. – 174 с.

49 Peng-Hsuan H. For beginners: using Keras to build models [Електронний ресурс] / Peng-Hsuan. – 2017. – Режим доступу: <https://www.kaggle.com/randyrose2017/for-beginners-using-keras-to-build-models/comments>.

50 Anaconda Distribution [Електронний ресурс] – Режим доступу: <https://www.anaconda.com/distributio>

51 NumPy [Електронний ресурс]. – 2019. – Режим доступу: <https://www.numpy.org/>.

52 Python Data Analysis Library [Електронний ресурс]. – 2019. – Режим доступу: <https://pandas.pydata.org/>.

53 HEATMAP [Електронний ресурс]. – 2019. – Режим доступу: <https://python-graph-gallery.com/heatmap/>.

- 54 Smith L. Чем PostgreSQL лучше других SQL баз данных с открытым исходным кодом. Часть 1 [Электронный ресурс] / Lisa Smith. – 2016. – Режим доступа: <https://habr.com/ru/post/282764/>.
- 55 PgAdmin 4 Documentation [Электронный ресурс]. – 2019. – Режим доступа: <https://www.pgadmin.org/docs/pgadmin4/4.x/>.
- 56 Kumar A. ML Metrics: Sensitivity vs. Specificity [Электронный ресурс] / Ajitesh Kumar. – 2018. – Режим доступа до ресурсу: <https://dzone.com/articles/ml-metrics-sensitivity-vs-specificity-difference>.
- 57 Выбор из другой функции стоимости и функции активации нейронной сети [Электронный ресурс]. – 2015. – Режим доступа до ресурсу: <http://qaru.site/questions/264833/choosing-from-different-cost-function-and-activation-function-of-a-neural-network>.
- 58 Поллак Г. А. Применение генетических алгоритмов для обучения нейронных сетей / Г. А. Поллак, 2011. – С. 174–178.
- 59 Что такое функция стоимости в нейронной сети? [Электронный ресурс]. – 2015. – Режим доступа до ресурсу: <http://qaru.site/questions/4994816/what-is-cost-function-in-neural-network>.
- 60 Gómez R. Understanding Categorical Cross-Entropy Loss, Binary Cross-Entropy Loss, Softmax Loss, Logistic Loss, Focal Loss and all those confusing names [Электронный ресурс] / Raúl Gómez. – 2018. – Режим доступа до ресурсу: https://gombu.github.io/2018/05/23/cross_entropy_loss/.
- 61 Батенко Л. П. Управління проектами: Навч. посібник. / Л. П. Батенко, О. А. Загородніх, В. Ліщинська. – Київ: КНЕУ, 2003. – 231 с.
- 62 Грей К. Управление проектами: Практическое руководство/Пер. с англ. / К. Грей, Э. Ларсон., 2003. – 528 с.
- 63 Локк Д. Основы Управления Проектами / Пер. с англ. / Д. Локк., 2004. – 253 с.

ДОДАТОК А

4.7 Планування робіт

4.7.1 Ідентифікація мети ІТ-проекту методом SMART

Більш детально описати мету проекту можна за допомогою SMART –методу, що впливає з розшифровки термінів, які формують його назву: конкретна – Specific, вимірювана – Measurable, досяжна – Achievable, реалістична – Relevant, обмежена у часі – Time-framed (табл. А.1).

Матеріальний об’єкт, який буде виступати в якості продукту проекту – веб-додаток з інтуїтивно зрозумілим інтерфейсом для зручного маніпулювання власними даними для користувача.

Проблема, яка буде вирішена за рахунок застосування продукту проекту – користувач на основі введених достовірних даних про стан його організму зможе отримати результати прогнозування імовірності захворіти на цервікальний рак.

Таблиця А.1 – Деталізація мети методом SMART

Термін	Пояснення	Значення
Specific (конкретна)	Вимагає щоб сформульована мета давала чітке якісне уявлення про специфічні унікальні та інноваційні властивості майбутнього продукту проекту порівняно з іншими його альтернативами.	Створити інформаційну технологію, яка буде представляти веб-додаток для прогнозування виникнення онкологічного захворювання РШМ в умовах невизначеностей з урахуванням індивідуальних показників жінки.

Продовження таблиці А.1

Термін	Пояснення	Значення
Measurable (вимірювана)	Передбачає показників вартості які вимірюються. При відсутності фізичних способів та інструментарію виміру використовуються експерти – як інструмент для виміру.	Результатом роботи проекту є оцінка точності прогнозованих даних.
Achievable (досяжна)	Встановлює, що мета повинна впливати з реальних проблем, місії, стратегічних планів, планів розвитку, а також узгоджуватись з інтересами зацікавлених сторін проекту.	У сучасному світі онкологічні захворювання є досить поширеною проблемою, тому мати змогу заздалегідь оцінити ризики появи такого захворювання є актуальним питанням для кожного.
Relevant (реалістична)	Показує, що мета є такою, яку можливо досягти з урахуванням реально доступних ресурсних можливостей та обмежень (людських, фінансових тощо).	У наявності є доступні ресурсні можливості, які були визначені на основі проведеного аналізу вже наявних досліджень експертів у сфері медицини.
Time-framed (обмежена у часі)	Зумовлює необхідність «прив'язки» мети до певних обґрунтованих термінів її досягнення (або початку та тривалості дій по її досягненню).	Мета має часове обмеження. Робота повинна бути виконана за встановлений термін згідно з календарним планом.

4.7.2 Планування змісту структури робіт ІТ-проекту

Планування змісту структури робіт ІТ-проекту (WBS). Виробнича структура (WBS - work breakdown structure) – це ієрархічна структура, побудована з метою логічного розподілу усіх робіт з виконання проекту і подана у графічному вигляді. Це сукупність декількох рівнів, кожний з яких формується в результаті розподілу роботи попереднього рівня на її складові. На верхньому першому рівні WBS фіксується продукт проекту. Наступний II рівень відповідає діям або основним

заходам для досягнення продукту проекту. Потім триває розбивка цих дій доти, поки не відбувається виконання дій елементарних робіт.

Структура декомпозиції робіт (WBS) у проектному менеджменті та системотехніці є орієнтованою на доконане виконання проекту декомпозицією проекту на менші частки. Структура декомпозиції робіт є ключовою часткою робіт по проекту, яка організовує командну роботу по проекту у керовані частини [61].

На рисунку Б.1 представлено WBS структуру розроблюваного проекту, яка містить перелік операцій необхідний для реалізації поставленої мети.

Планування структури організації для впровадження готового проекту (OBS). Наступним кроком розробки структури проекту є визначення організаційної структури (OBS) проекту.

Внутрішня організаційна структура проекту (Organization Breakdown Structure - OBS) — це система зв'язків між окремими виконавцями і групами, які працюють над проектом як окремі організаційні одиниці всередині проектної команди. При цьому зовнішні стосунки виконавців і груп з «материнськими» підрозділами або компаніями не беруться до уваги. На верхньому рівні OBS розташована команда проекту. На наступному рівні фіксуються виконавці: організації, відділи тощо.

На рисунку Б.2 представлено OBS структуру даного проекту, яка включає перелік учасників та виконавців проекту відповідно до їх обов'язків.

4.7.3 Побудова календарного графіку виконання ІТ - проекту

Для того щоб мати реальне уявлення про тривалість виконання робіт з урахуванням обмеженості у використанні ресурсів, на підставі часткової мережевої моделі, а також, проекту в цілому з урахуванням вихідних та святкових днів, будують календарний графік робіт. Він є реальним розподілом робіт з пакету за календарними датами, тобто своєрідним розкладом виконання робіт.

Діаграма Ганта – горизонтальна лінійна діаграма, на якій задачі проекту представляються протяжними в часі відрізками, що характеризуються датами початку та закінчення, затримками і, можливо, іншими тимчасовими параметрами [62].

Кожен відрізок відповідає окремому завданню або підзадачі. Завдання і підзадачі, складові плану, розміщуються по вертикалі. Початок, кінець і довжина відрізка на шкалі часу відповідають початку, кінцю і тривалості завдання. На деяких діаграмах Ганта також показується залежність між завданнями.

У представленій діаграмі Ганта усі підзадачі виконуються послідовно у зв'язку з обмеженими трудовими ресурсами. Виконавці витрачали у середньому 1 год/день на виконання робіт.

На рисунках Б.3 - Б.4 представлено діаграму Ганта розроблюваного проекту.

Формування бюджету проекту. Бюджет проекту – план, який включає кількісні показники та відображає витрати, що необхідні для досягнення поставленої мети. У бюджеті зазначаються оцінені результати календарного плану та стратегія реалізації проекту.

Кошторис продукту проекту – це загальні майбутні витрати, які необхідні безпосередньо для створення продукту проекту. Тобто це витрати на фінансування всіх робіт, передбачених WBS - структурою проекту.

Бюджет продукту проекту – це кошторис продукту проекту, розподілений в часі на основі календарного плану реалізації робіт або за окремими WBS елементами.

План фінансування – це кошторис продукту проекту в розрізі основних джерел фінансування робіт з проекту [63].

Трудові ресурси:

- Єлісеєва А.Р. – 50 грн/год;
- Шендрик В.В. – 100 грн/год.

У таблиці А.2 зображений розрахований бюджет проекту.

Таблиця А.2 – Бюджет проекту

Назва роботи	Продовжність (днів)	Трудові затрати (год)	Витрати (грн)
Інформаційна технологія оцінювання ризику захворюваності на цервікальний рак за допомогою нейронних мереж	148	176	10 350,00₴
Планування робіт	30	47	3 250,00₴
Постановка завдання проекту	9	16	1 200,00₴
Аналіз предметної області	9	16	1 200,00₴
Аналіз проведених досліджень	4	8	600,00₴
Аналіз наявних аналогів програмних продуктів	4	8	600,00₴
Постановка задачі та вибір методів розробки	6	10	700,00₴
Опис мети та задач дослідження	2	4	300,00₴
Аналіз методів дослідження	2	4	300,00₴
Вибір засобів реалізації	2	2	100,00₴
Планування проекту	14	21	1 350,00₴
Ідентифікація мети ІТ-проекту методом SMART	5	6	450,00₴
Планування змісту структури робіт ІТ-проекту	5	6	450,00₴
Побудова календарного графіку виконання ІТ - проекту	5	3	150,00₴
Формування бюджету проекту	5	3	150,00₴
Планування ризиків ІТ-проекту	5	3	150,00₴
Проектування проекту	41	39	1 950,00₴
Структурно-функціональне моделювання	12	12	600,00₴

Продовження таблиці А.2

Назва роботи	Продовжність (днів)	Трудові затрати (год)	Витрати (грн)
Побудова діаграм в нотації сімейства IDEF	3	3	150,00₴
Побудова архітектури інформаційної технології	3	3	150,00₴
Побудова UML-діаграм	3	3	150,00₴
Розробка моделей бази даних	3	3	150,00₴
Проектування інтерфейсу	27	27	1 350,00₴
Розробка макету	12	12	600,00₴
Створення прототипу	15	15	750,00₴
Розробка проекту	62	73	4 200,00₴
Розробка веб-додатку	35	35	1 750,00₴
Створення дизайну	15	15	750,00₴
Програмування основного функціоналу	20	20	1 000,00₴
Створення бази даних	7	7	350,00₴
Створення міграцій	3	3	150,00₴
Заповнення бази даних вхідними даними	2	2	100,00₴
Заповнення бази даних вихідними даними	2	2	100,00₴
Розробка нейронної мережі	20	31	2 100,00₴
Формування вихідної вибірки даних	3	4	250,00₴
Попередня обробка і нормалізація вихідної вибірки	3	3	150,00₴
Поділ вихідної вибірки на навчальну і тестову складові	2	2	100,00₴
Визначення структури нейронної мережі	3	6	450,00₴

Продовження таблиці А.2

Назва роботи	Продовжність (днів)	Трудові затрати (год)	Витрати (грн)
Налаштування параметрів нейронної мережі та алгоритму її навчання	4	8	600,00₴
Навчання нейронної мережі	3	6	450,00₴
Тестування нейронної мережі	1	1	50,00₴
Оцінка точності навчання нейронної мережі	1	1	50,00₴
Тестування проекту	6	6	300,00₴
Модульне тестування	2	2	100,00₴
Інтеграційне тестування	2	2	100,00₴
Системне тестування	2	2	100,00₴
Здача проекту	9	11	650,00₴
Оформлення документаций проекту	6	6	300,00₴
Створення інструкції користувача	3	3	150,00₴
Створення інструкції адміністратора	3	3	150,00₴
Здача проекту	3	5	350,00₴
Представлення презентації проекту	1	1	50,00₴
Дослідна експлуатація	2	4	300,00₴

Розмір бюджету проекту: 10 350,00 грн.

На рисунках Б.5 - Б.6 зображений бюджет проекту в залежності від трудових ресурсів.

4.7.4 Планування ризиків проекту

Планування та реалізація проектів відбувається в умовах невизначеності, що породжується зміною внутрішнього та зовнішнього середовищ. Під невизначеністю

розуміють відсутність повної та достовірної інформації про умови реалізації проекту.

Невизначеність, що пов'язана з можливістю виникнення в ході реалізації проекту несприятливих умов, ситуацій та наслідків називається ризиком.

У ринковій економіці ризик є невід'ємним атрибутом господарювання. Невизначеність приводить до того, що уникнути ризику неможливо. Але це не значить, що слід шукати такі рішення, в яких завчасно відомий результат, вони, як правило, неефективні. Необхідно навчитися передбачати ризик, оцінювати його розміри, планувати заходи щодо його запобігання. Ризик є складною економічно-управлінською категорією, при визначенні якої має місце ряд протиріч.

Управління ризиком – це процес реагування на події та зміни ризиків в процесі виконання проекту.

При цьому важливим є проведення моніторингу ризиків. Моніторинг ризиків включає контроль ризиків протягом всього життєвого циклу проекту. Якісний моніторинг ризиків забезпечує управління інформацією, яка допомагає приймати ефективні рішення до настання ризикових подій.

На першому етапі ефективним інструментом, який допоможе визначити всі можливі ризики, може стати карта процесів. У ній визначаються завдання проекту, цілі та процеси (заходи), які необхідно здійснити для їх досягнення. Побудова такої карти дозволяє охопити процеси в їх взаємозв'язку, а потім визначити найбільш істотні ризики. Далі за допомогою якісного аналізу об'єктивних і суб'єктивних факторів, що впливають на збільшення ступеня ризику, визначають процеси, які найбільшою мірою схильні до ризиків і перешкоджають досягненню цілей.

На рисунку Б.7 зображені ризики проекту.

На другому етапі проводиться кількісна оцінка виявлених ризиків, яка може бути виражена відносним або абсолютним рівнем витрат і вимірюється ймовірністю виникнення ризику та ступенем його впливу при виникненні (рис. А.1).

Процесс оценки идентифицированных рисков в ходе реализации проекта **2**



Рисунок А.1 – Методи оцінки ризиків

Для визначення ступеня впливу розраховують величину втрат у відсотках від планової величини прибутку (рис. А.2).

3 Классификация рисков по величине потерь

Виды рисков	Величина потерь	
	I_q (баллы)	I_q (в % от плановой прибыли по объекту)
Минимальные	1	$0 < I_q \in 10$
Низкие	2	$10 < I_q \in 40$
Средние	3	$40 < I_q \in 60$
Высокие	4	$60 < I_q \in 90$
Максимальные	5	$90 < I_q \in 100$

Рисунок А.2 – Класифікація ризиків за величиною втрат

Потім, перемножуючи значення ймовірності виникнення та ступеня впливу, отримуємо індекс ризику – показник величини ймовірних втрат в балах, який визначається за допомогою матриці «Ймовірність – Втрати» і дає можливість оцінити ступінь впливу й рівні ризику (табл. А.3).

Таблиця А.3 – Оцінка ризиків проекту

№		Ризик	Імовірність виникнення	Втрати	Ступінь впливу		Рівень ризику	
1	1.1 Процеси проекту	1.1.1	Некоректно складене ТЗ	3	4	12	Істотний	Недопустимий
		1.1.2	Зрив плану робіт проекту	1	5	5	Незначний	Виправданий
	1.2 Людські ресурси	1.2.1	Непорозуміння між замовником та розробниками	1	1	1	Ігноруючий	Прийнятний
		1.2.2	Захворювання розробників	2	5	10	Помірний	Виправданий
		1.2.3	Сімейні обставини розробників	1	4	4	Ігноруючий	Прийнятний
		1.2.4	Низька мотивація розробників	3	4	12	Істотний	Недопустимий
	1.3 Фінансування	1.3.1	Обмеженість бюджетних коштів	4	5	20	Критичний	Недопустимий
		1.3.2	Зниження заробітної плати ІТ-працівникам	2	4	8	Незначний	Виправданий
2.1	Невиконання учасниками проекту своїх зобов'язань	1	5	5	Незначний	Виправданий		

Продовження таблиці А.3

№		Ризик	Імовірність виникнення	Втрати	Ступінь впливу		Рівень ризику	
2	Управління проектом	2.2	Неефективне планування кошторисної вартості проекту	2	3	6	Незначний	Виправданий
		2.3	Затримка введення проекту в експлуатацію	1	4	4	Ігноруючий	Прийнятний
		2.4	Неоптимальне планування управління проектом	3	5	15	Істотний	Недопустимий
		2.5	Низький рівень кваліфікації виконавців проекту	2	5	10	Помірний	Виправданий
3	Технічні	3.1	Проблеми з технічним обладнанням	3	5	15	Істотний	Недопустимий
		3.2	Несумісність СУБД розробника та замовника	1	5	5	Незначний	Виправданий
		4.1	Підвищення цін на апаратне та програмне забезпечення	4	1	4	Ігноруючий	Прийнятний

Продовження таблиці А.3

№		Ризик	Імовірність виникнення	Втрати	Ступінь впливу		Рівень ризику
4 Зовнішні	4.2	Ризик негативних соціальних наслідків	2	1	2	Ігноруючий	Прийнятний
	4.3	Зміна вимог замовника у процесі реалізації проекту	5	5	25	Критичний	Недопустимий
	4.4	Природні катастрофи	1	5	5	Незначний	Виправданий

Запропонуємо варіанти запобігання (що потрібно робити, щоб запобігти виникненню) та мінімізації витрат (що потрібно робити, у разі настання ризику). Виберемо ті ризики, які одночасно є істотними та недопустимими, або критичними та недопустимими (табл. А.4).

На рисунку А.3 представлені проаналізовані ризику розроблюваного проекту.

Таблиця А.4 – Варіанти запобігання та мінімізація наслідків ризиків

Ризик	Варіант запобігання	Мінімізація наслідків
Некоректно складене ТЗ	Чітке визначення вимог до програмного продукту та апаратного забезпечення із замовником	Виправлення неправильно сформульованих вимог

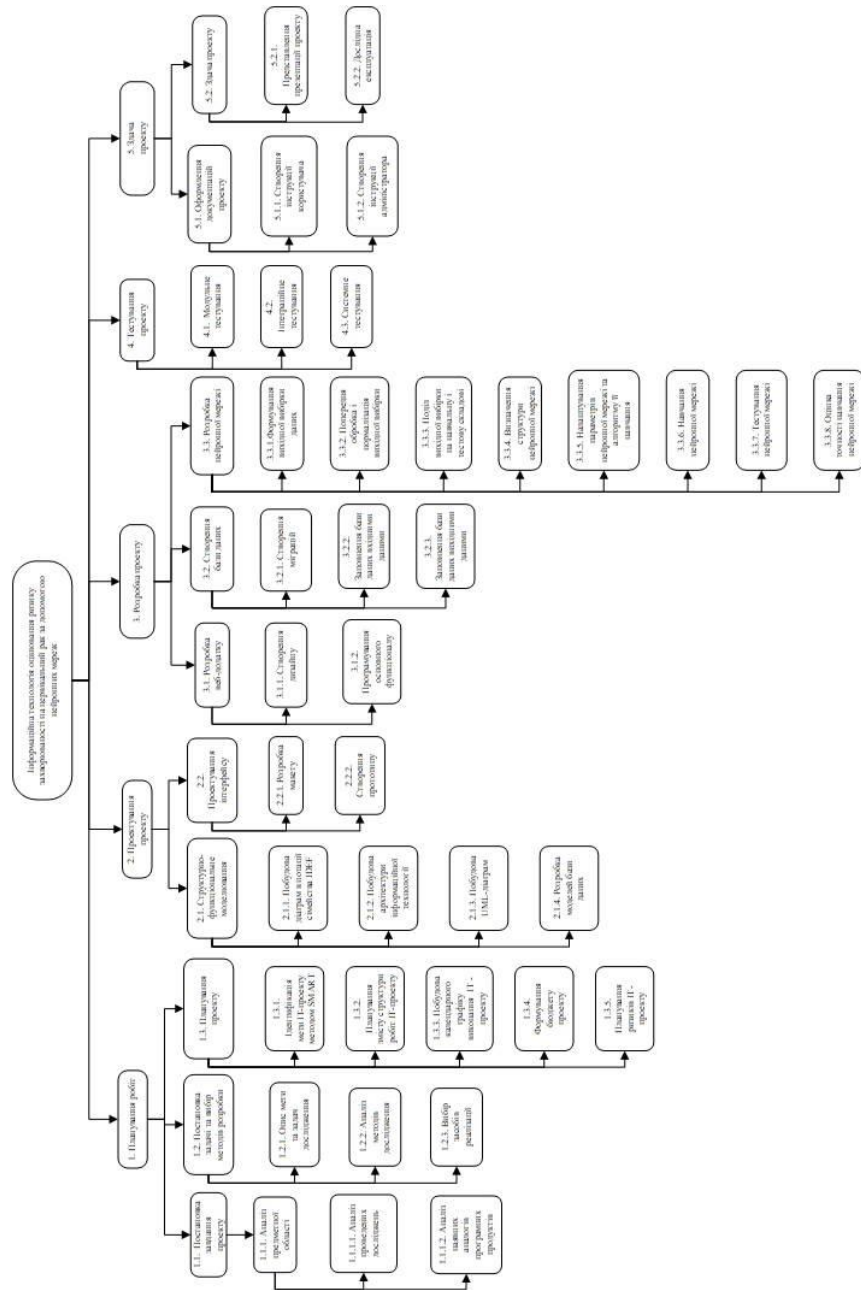
Продовження таблиці А.4

Ризик	Варіант запобігання	Мінімізація наслідків
Низька мотивація розробників	Заохочуванням до ефективної роботи розробників шляхом надання премій та відповідної заробітної плати	Надання премій розробникам у ході виконання проекту, забезпечення комфортних умов праці
Обмеженість бюджетних коштів	Ефективно планувати бюджет проекту	Збільшення надаваних коштів для найбільш важливих для проекту робіт, визначених замовником
Неоптимальне планування управління проектом	Складання оптимального плану управління проектом в залежності від виконуваних робіт	Максимально оптимальна корекція залишкового часу виконуваного проекту
Проблеми з технічним обладнанням	Придбання якісного апаратного забезпечення та вчасне проведення його діагностики	Швидке виправлення технічних проблем професіоналами
Зміна вимог замовника у процесі реалізації проекту	Максимальне передбачення усіх можливих вимог та виключних ситуацій відповідно до побажань замовника, керуючись життєвим досвідом	Реалізація нових умов з мінімальним впливом на розроблений проект та пояснення замовнику, що за затримку реалізації проекту відповідає він

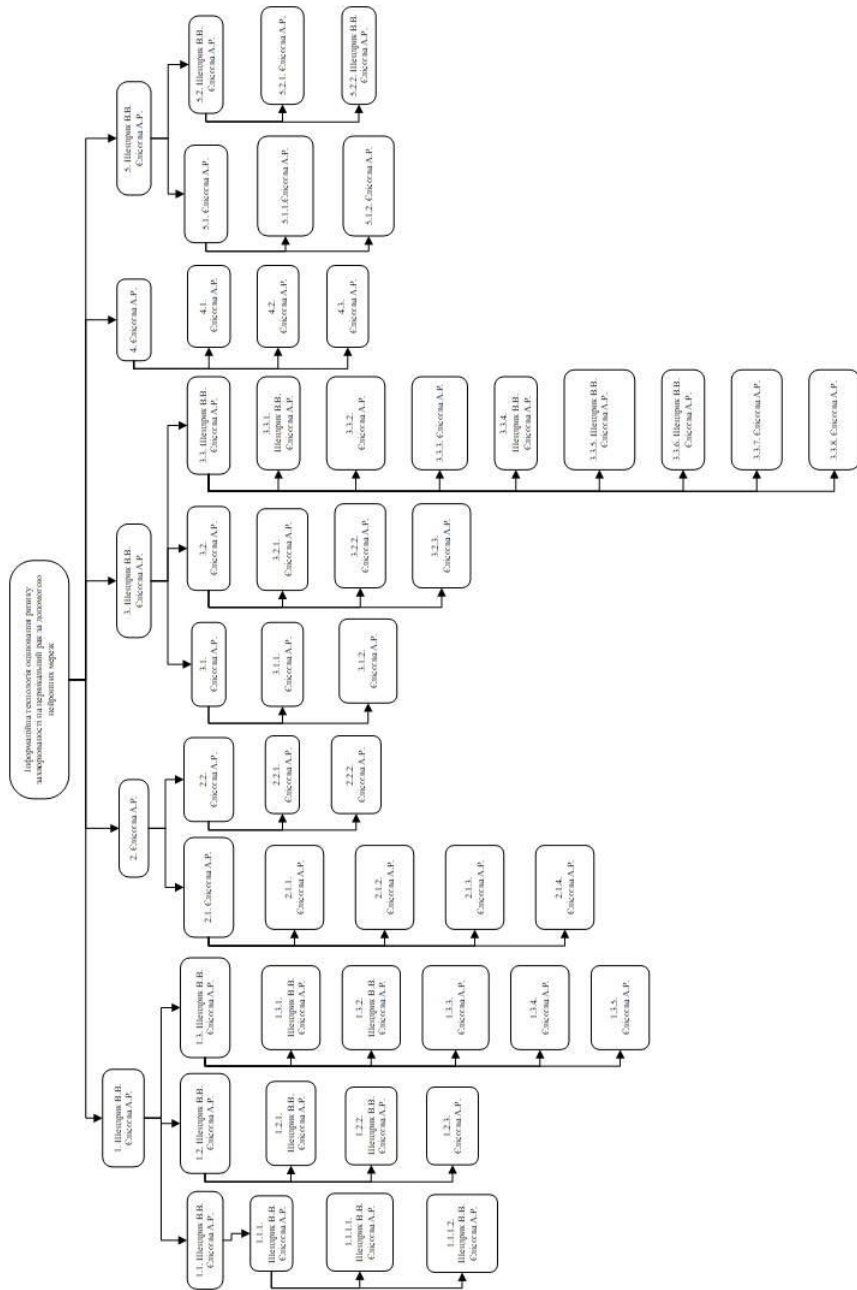
№		Ризик	Імовірність виникнення	Втрати	Ступінь впливу	Рівень ризику		
1	1.1 Процеси проекту	1.1.1	Некоректно складене ТЗ	3	4	12	Істотний	Недопустимий
		1.1.2	Неоптимальний розподіл часу виконання робіт	2	5	10	Помірний	Виправданий
		1.1.3	Зрив плану робіт проекту	1	5	5	Незначний	Виправданий
	1.2 Людські ресурси	1.2.1	Непорозуміння між замовником та розробниками	1	1	1	Ігноруючий	Прийнятний
		1.2.2	Захворювання розробників	2	5	10	Помірний	Виправданий
		1.2.3	Сімейні обставини розробників	1	4	4	Ігноруючий	Прийнятний
		1.2.4	Низька мотивація розробників	3	4	12	Істотний	Недопустимий
	1.3 Фінансування	1.3.1	Обмеженість бюджетних коштів	4	5	20	Критичний	Недопустимий
1.3.2		Зниження заробітної плати ІТ-працівникам	2	4	8	Незначний	Виправданий	
2	Управління проектом	2.1	Невиконання учасниками проекту своїх зобов'язань	1	5	5	Незначний	Виправданий
		2.2	Неефективне планування кошторисної вартості проекту	2	3	6	Незначний	Виправданий
		2.3	Затримка введення проекту в експлуатацію	1	4	4	Ігноруючий	Прийнятний
		2.4	Неоптимальне планування управління проектом	3	5	15	Істотний	Недопустимий
		2.5	Низький рівень кваліфікації виконавців проекту	2	5	10	Помірний	Виправданий
3	Технічні	3.1	Проблеми з технічним обладнанням	3	5	15	Істотний	Недопустимий
		3.2	Несумісність СУБД розробника та замовника	1	5	5	Незначний	Виправданий
4	Зовнішні	4.1	Підвищення цін на апаратне та програмне забезпечення	4	1	4	Ігноруючий	Прийнятний
		4.2	Ризик негативних соціальних наслідків	2	1	2	Ігноруючий	Прийнятний
		4.3	Зміна вимог замовника у процесі реалізації проекту	5	5	25	Критичний	Недопустимий
		4.4	Природні катастрофи	1	5	5	Незначний	Виправданий

Рисунок А.3 – Проаналізовані ризики проекту

ДОДАТОК Б

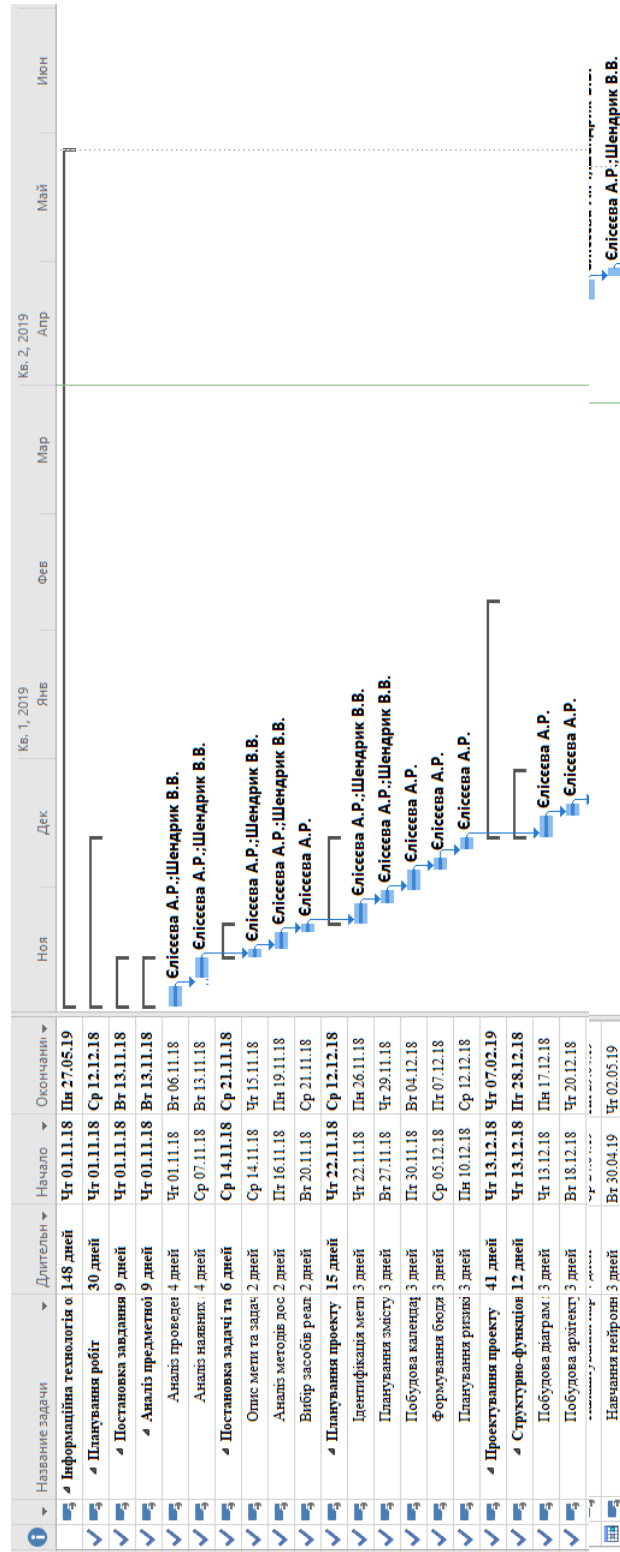


Ри
су
но
к
Б.
1
–
Ст
ру
кт
ур
а
ро
біт
пр
ое
кт
у
(
W



Ри
су
но
к
Б.
2
–
Ор
га
ніз
ац
ій
на
ст
ру
кт
ур
а
пр
ое

Ри
су
но
к
Б.
3
—
Ді
аг
ра
ма
Га
нт
а



Елісеєва А.Р.; Шендрик В.В.



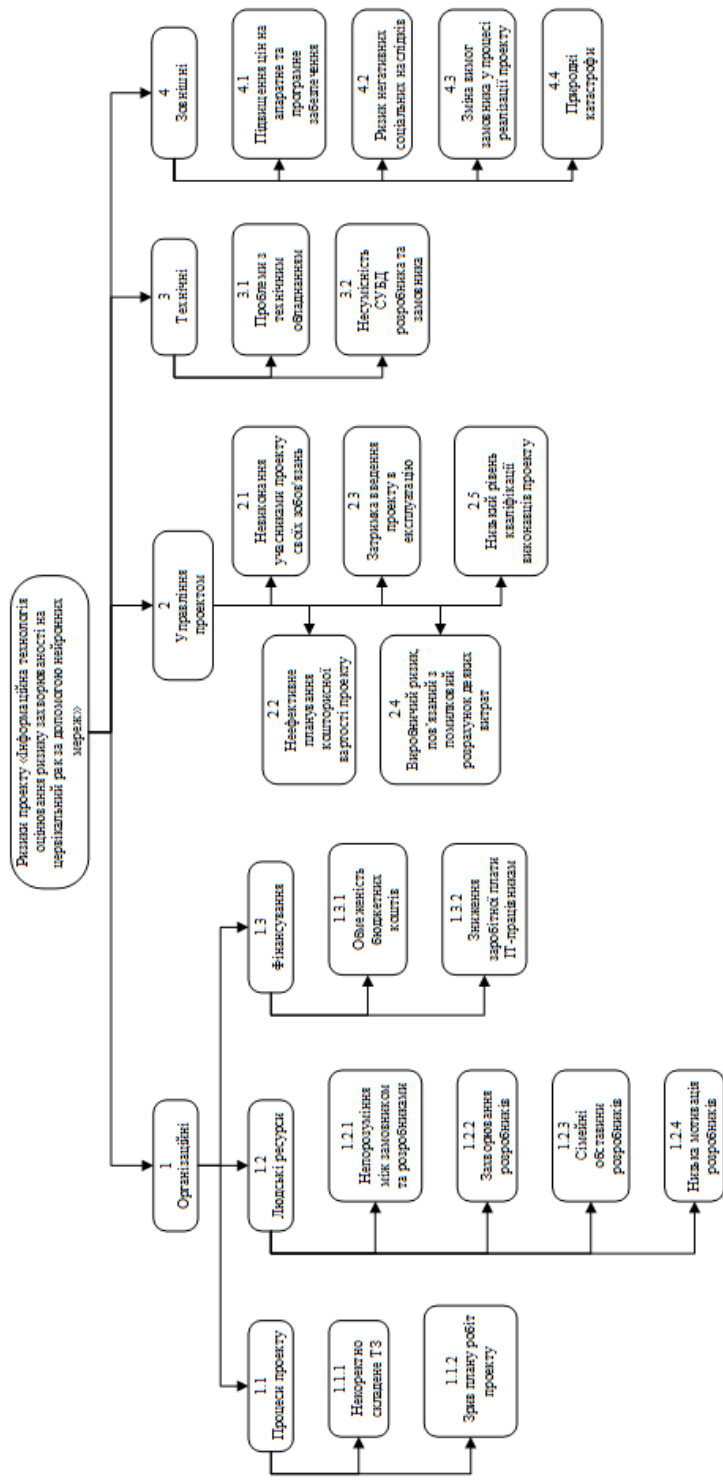
Ри
су
но
к
Б.
4
–
Пр
од
ов
же
нн
я
діа
гр
ам
и
Га
нт
а

№	Назва задачі	Длительн	Начало	Окончани	П	Названия ресурсов	Тру,	Затраты
1	Інформаційна технологія о	148 днів	Чт 01.11.18	Пн 27.05.19			176 ч	10 350,00€
2	Планування робіт	30 днів	Чт 01.11.18	Ср 12.12.18			47 ч	3 250,00€
3	Постановка завдання	9 днів	Чт 01.11.18	Вт 13.11.18			16 ч	1 200,00€
4	Аналіз предметної	9 днів	Чт 01.11.18	Вт 13.11.18			16 ч	1 200,00€
5	Аналіз проведе	4 днів	Чт 01.11.18	Вт 06.11.18		Єлісеєва А.Р.;Шендрік В.В.	8 ч	600,00€
6	Аналіз наявних	4 днів	Ср 07.11.18	Вт 13.11.18	5	Єлісеєва А.Р.;Шендрік В.В.	8 ч	600,00€
7	Постановка задачі та	6 днів	Ср 14.11.18	Ср 21.11.18			10 ч	700,00€
8	Опис мети та задач	2 днів	Ср 14.11.18	Чт 15.11.18	6	Єлісеєва А.Р.;Шендрік В.В.	4 ч	300,00€
9	Аналіз методів дос	2 днів	Пт 16.11.18	Пн 19.11.18	8	Єлісеєва А.Р.;Шендрік В.В.	4 ч	300,00€
10	Вибір засобів реал	2 днів	Вт 20.11.18	Ср 21.11.18	9	Єлісеєва А.Р.	2 ч	100,00€
11	Планування проекту	15 днів	Чт 22.11.18	Ср 12.12.18			21 ч	1 350,00€
12	Ідентифікація мети	3 днів	Чт 22.11.18	Пн 26.11.18	10	Єлісеєва А.Р.;Шендрік В.В.	6 ч	450,00€
13	Планування змісту	3 днів	Вт 27.11.18	Чт 29.11.18	12	Єлісеєва А.Р.;Шендрік В.В.	6 ч	450,00€
14	Побудова календа	3 днів	Пт 30.11.18	Вт 04.12.18	13	Єлісеєва А.Р.	3 ч	150,00€
15	Формування бюдж	3 днів	Ср 05.12.18	Пт 07.12.18	14	Єлісеєва А.Р.	3 ч	150,00€
16	Планування ризику	3 днів	Пн 10.12.18	Ср 12.12.18	15	Єлісеєва А.Р.	3 ч	150,00€
17	Проектування проекту	41 днів	Чт 13.12.18	Чт 07.02.19			39 ч	1 950,00€
18	Структурно-функціо	12 днів	Чт 13.12.18	Пт 28.12.18			12 ч	600,00€
19	Побудова діаграм	3 днів	Чт 13.12.18	Пн 17.12.18	16	Єлісеєва А.Р.	3 ч	150,00€
20	Побудова архітекту	3 днів	Вт 18.12.18	Чт 20.12.18	19	Єлісеєва А.Р.	3 ч	150,00€
21	Побудова UML-діа	3 днів	Пт 21.12.18	Вт 25.12.18	20	Єлісеєва А.Р.	3 ч	150,00€
22	Розробка моделей	3 днів	Ср 26.12.18	Пт 28.12.18	21	Єлісеєва А.Р.	3 ч	150,00€
23	Проектування інтерф	27 днів	Ср 02.01.19	Чт 07.02.19			27 ч	1 350,00€
24	Розробка макету	12 днів	Ср 02.01.19	Чт 17.01.19	22	Єлісеєва А.Р.	12 ч	600,00€
25	Створення протот	15 днів	Пт 18.01.19	Чт 07.02.19	24	Єлісеєва А.Р.	15 ч	750,00€
26	Розробка проекту	62 днів	Пт 08.02.19	Пн 06.05.19			73 ч	4 200,00€
27	Розробка веб-додатк	35 днів	Пт 08.02.19	Чт 28.03.19			35 ч	1 750,00€
28	Створення дизайну	15 днів	Пт 08.02.19	Чт 28.02.19	25	Єлісеєва А.Р.	15 ч	750,00€
29	Програмування ос	20 днів	Пт 01.03.19	Чт 28.03.19	28	Єлісеєва А.Р.	20 ч	1 000,00€
30	Створення бази дан	7 днів	Пт 29.03.19	Пн 08.04.19			7 ч	350,00€
31	Створення міграції	3 днів	Пт 29.03.19	Вт 02.04.19	29	Єлісеєва А.Р.	3 ч	150,00€
32	Заповнення бази д	2 днів	Ср 03.04.19	Чт 04.04.19	31	Єлісеєва А.Р.	2 ч	100,00€
33	Заповнення бази д	2 днів	Пт 05.04.19	Пн 08.04.19	32	Єлісеєва А.Р.	2 ч	100,00€
34	Розробка нейронної м	20 днів	Вт 09.04.19	Пн 06.05.19			31 ч	2 100,00€
35	Формування вихід	3 днів	Вт 09.04.19	Чт 11.04.19	33	Єлісеєва А.Р.;Шендрік В.В.	4 ч	250,00€
36	Попередня обробк	3 днів	Пт 12.04.19	Вт 16.04.19	35	Єлісеєва А.Р.	3 ч	150,00€
37	Поділ вихідної вибі	2 днів	Ср 17.04.19	Чт 18.04.19	36	Єлісеєва А.Р.	2 ч	100,00€
38	Визначення структ	3 днів	Ср 17.04.19	Вт 23.04.19	37	Єлісеєва А.Р.;Шендрік В.В.	6 ч	450,00€
39	Налаштування пар	4 днів	Ср 24.04.19	Пн 29.04.19	38	Єлісеєва А.Р.;Шендрік В.В.	8 ч	600,00€
40	Навчання нейронн	3 днів	Вт 30.04.19	Чт 02.05.19	39	Єлісеєва А.Р.;Шендрік В.В.	6 ч	450,00€

Рисунок Б.5 – Бюджет проекту

	Тестування нейрог	1 день	Пт 03.05.19	Пт 03.05.19	40	Єлісеєва А.Р.	1 ч	50,00₴
	Оцінка точності на	1 день	Пн 06.05.19	Пн 06.05.19	41	Єлісеєва А.Р.	1 ч	50,00₴
	▲ Тестування проекту	6 днів	Вт 07.05.19	Вт 14.05.19			6 ч	300,00₴
	Модульне тестування	2 днів	Вт 07.05.19	Ср 08.05.19	42	Єлісеєва А.Р.	2 ч	100,00₴
	Інтеграційне тестуван	2 днів	Чт 09.05.19	Пт 10.05.19	44	Єлісеєва А.Р.	2 ч	100,00₴
	Системне тестування	2 днів	Пн 13.05.19	Вт 14.05.19	45	Єлісеєва А.Р.	2 ч	100,00₴
	▲ Задача проекту	9 днів	Ср 15.05.19	Пн 27.05.19			11 ч	650,00₴
	▲ Оформлення докуме	6 днів	Ср 15.05.19	Ср 22.05.19			6 ч	300,00₴
	Створення інструк	3 днів	Ср 15.05.19	Пт 17.05.19	46	Єлісеєва А.Р.	3 ч	150,00₴
	Створення інструк	3 днів	Пн 20.05.19	Ср 22.05.19	49	Єлісеєва А.Р.	3 ч	150,00₴
	▲ Задача проекту	3 днів	Чт 23.05.19	Пн 27.05.19			5 ч	350,00₴
	Представлення пре	1 день	Чт 23.05.19	Чт 23.05.19	50	Єлісеєва А.Р.	1 ч	50,00₴
	Дослідна експлуат:	2 днів	Пт 24.05.19	Пн 27.05.19	52	Єлісеєва А.Р.;Шендрик В.В.	4 ч	300,00₴

Рисунок Б.6 – Продовження бюджету проекту



Ри
су
но
к
Б.
7
—
Ри
зи
ки
пр
ое
кт
у

ДОДАТОК В

Скрипт для створення таблиць у базі даних:

```
--
-- PostgreSQL database dump
--

-- Dumped from database version 11.2
-- Dumped by pg_dump version 11.2

SET statement_timeout = 0;
SET lock_timeout = 0;
SET idle_in_transaction_session_timeout = 0;
SET client_encoding = 'UTF8';
SET standard_conforming_strings = on;
SELECT pg_catalog.set_config('search_path', '', false);
SET check_function_bodies = false;
SET client_min_messages = warning;
SET row_security = off;

SET default_tablespace = '';

SET default_with_oids = false;

CREATE TABLE public."AspNetUsers" (
  "Id" text NOT NULL,
  "UserName" character varying(256),
  "NormalizedUserName" character varying(256),
  "Email" character varying(256),
  "NormalizedEmail" character varying(256),
  "EmailConfirmed" boolean NOT NULL,
  "PasswordHash" text,
  "SecurityStamp" text,
  "ConcurrencyStamp" text,
  "PhoneNumber" text,
  "PhoneNumberConfirmed" boolean NOT NULL,
  "TwoFactorEnabled" boolean NOT NULL,
  "LockoutEnd" timestamp with time zone,
  "LockoutEnabled" boolean NOT NULL,
  "AccessFailedCount" integer NOT NULL
);

ALTER TABLE public."AspNetUsers" OWNER TO postgres;

CREATE TABLE public."DatasetModels" (
  "IDDataset" integer NOT NULL,
  "Age" integer NOT NULL,
  "NumberOfSexualPartners" integer NOT NULL,
  "AgeOfFirstSexualIntercourse" integer NOT NULL,
  "NumberOfPregnancies" integer NOT NULL,
  "Smokes" boolean NOT NULL,
  "AgeSmokes" integer NOT NULL,
  "PacksYearSmokes" integer NOT NULL,
  "UsingHormonalContraceptives" boolean NOT NULL,
  "AgeUsingHormonalContraceptives" integer NOT NULL,
  "IUD" boolean NOT NULL,
```

```

"AgeIUD" integer NOT NULL,
"STDs" boolean NOT NULL,
"NumberOfSTDs" integer NOT NULL,
"STDs_Condylomatosis" boolean NOT NULL,
"STDs_CervicalCondylomatosis" boolean NOT NULL,
"STDs_VaginalCondylomatosis" boolean NOT NULL,
"STDs_VulvoPerinealCondylomatosis" boolean NOT NULL,
"STDs_Syphilis" boolean NOT NULL,
"STDs_PelvicInflammatoryDisease" boolean NOT NULL,
"STDs_GenitalHerpes" boolean NOT NULL,
"STDs_MolluscumContagiosum" boolean NOT NULL,
"STDs_AIDS" boolean NOT NULL,
"STDs_HIV" boolean NOT NULL,
"STDs_HepatitisB" boolean NOT NULL,
"STDs_HPVP" boolean NOT NULL,
"STDs_NumberOfDiagnosis" boolean NOT NULL,
"STDs_TimeSinceFirstDiagnosis" boolean NOT NULL,
"STDs_TimeSinceLastDiagnosis" boolean NOT NULL,
"Dx_Cancer" boolean NOT NULL,
"Dx_CIN" boolean NOT NULL,
"Dx_HPVP" boolean NOT NULL,
"Dx" boolean NOT NULL,
"HinselmannTest" boolean NOT NULL,
"SchillerTest" boolean NOT NULL,
"Citology" boolean NOT NULL,
"FKIDUsers" text
);

ALTER TABLE public."DatasetModels" OWNER TO postgres;

CREATE TABLE public."PredictionBiopsyModels" (
  "IDPredictionBiopsy" integer NOT NULL,
  "PredictionBiopsy" integer NOT NULL,
  "DateTime" timestamp without time zone NOT NULL,
  "FKIDUsers" text
);

ALTER TABLE public."PredictionBiopsyModels" OWNER TO postgres;

CREATE TABLE public."TestDatasetModels" (
  "IDTestDataset" bigint,
  "Age" bigint,
  "NumberOfSexualPartners" bigint,
  "AgeOfFirstSexualIntercourse" bigint,
  "NumberOfPregnancies" bigint,
  "AgeSmokes" double precision,
  "PacksYearSmokes" double precision,
  "AgeUsingHormonalContraceptives" double precision,
  "AgeIUD" double precision,
  "NumberOfSTDs" bigint,
  "STDs_Condylomatosis" bigint,
  "STDs_CervicalCondylomatosis" bigint,
  "STDs_VaginalCondylomatosis" bigint,
  "STDs_VulvoPerinealCondylomatosis" bigint,
  "STDs_Syphilis" bigint,
  "STDs_PelvicInflammatoryDisease" bigint,
  "STDs_GenitalHerpes" bigint,
  "STDs_MolluscumContagiosum" bigint,
  "STDs_AIDS" bigint,
  "STDs_HIV" bigint,
  "STDs_HepatitisB" bigint,
  "STDs_HPVP" bigint,

```

```

"STDs_NumberOfDiagnosis" bigint,
"STDs_TimeSinceFirstDiagnosis" bigint,
"STDs_TimeSinceLastDiagnosis" bigint,
"Biopsy" bigint,
"Smokes_0" bigint,
"Smokes_1" bigint,
"UsingHormonalContraceptives_0" bigint,
"UsingHormonalContraceptives_1" bigint,
"IUD_0" bigint,
"IUD_1" bigint,
"STDs_0" bigint,
"STDs_1" bigint,
"Dx_Cancer_0" bigint,
"Dx_Cancer_1" bigint,
"Dx_CIN_0" bigint,
"Dx_CIN_1" bigint,
"Dx_HPVC_0" bigint,
"Dx_HPVC_1" bigint,
"Dx_0" bigint,
"Dx_1" bigint,
"HinselmannTest_0" bigint,
"HinselmannTest_1" bigint,
"Citology_0" bigint,
"Citology_1" bigint,
"SchillerTest_0" bigint,
"SchillerTest_1" bigint
);

```

```
ALTER TABLE public."TestDatasetModels" OWNER TO postgres;
```

```

CREATE TABLE public."TrainDatasetModels" (
  "IDTrainDataset" bigint,
  "Age" bigint,
  "NumberOfSexualPartners" bigint,
  "AgeOfFirstSexualIntercourse" bigint,
  "NumberOfPregnancies" bigint,
  "AgeSmokes" double precision,
  "PacksYearSmokes" double precision,
  "AgeUsingHormonalContraceptives" double precision,
  "AgeIUD" double precision,
  "NumberOfSTDs" bigint,
  "STDs_Condylomatosis" bigint,
  "STDs_CervicalCondylomatosis" bigint,
  "STDs_VaginalCondylomatosis" bigint,
  "STDs_VulvoPerinealCondylomatosis" bigint,
  "STDs_Syphilis" bigint,
  "STDs_PelvicInflammatoryDisease" bigint,
  "STDs_GenitalHerpes" bigint,
  "STDs_MolluscumContagiosum" bigint,
  "STDs_AIDS" bigint,
  "STDs_HIV" bigint,
  "STDs_HepatitisB" bigint,
  "STDs_HPVC" bigint,
  "STDs_NumberOfDiagnosis" bigint,
  "STDs_TimeSinceFirstDiagnosis" bigint,
  "STDs_TimeSinceLastDiagnosis" bigint,
  "Biopsy" bigint,
  "Smokes_0" bigint,
  "Smokes_1" bigint,
  "UsingHormonalContraceptives_0" bigint,
  "UsingHormonalContraceptives_1" bigint,
  "IUD_0" bigint,

```

```

"IUD_1" bigint,
"STDs_0" bigint,
"STDs_1" bigint,
"Dx_Cancer_0" bigint,
"Dx_Cancer_1" bigint,
"Dx_CIN_0" bigint,
"Dx_CIN_1" bigint,
"Dx_HPVI_0" bigint,
"Dx_HPVI_1" bigint,
"Dx_0" bigint,
"Dx_1" bigint,
"HinselmannTest_0" bigint,
"HinselmannTest_1" bigint,
"Citology_0" bigint,
"Citology_1" bigint,
"SchillerTest_0" bigint,
"SchillerTest_1" bigint
);

ALTER TABLE public."TrainDatasetModels" OWNER TO postgres;

CREATE TABLE public."__EFMigrationsHistory" (
  "MigrationId" character varying(150) NOT NULL,
  "ProductVersion" character varying(32) NOT NULL
);

ALTER TABLE public."__EFMigrationsHistory" OWNER TO postgres;

ALTER TABLE ONLY public."AspNetUsers"
  ADD CONSTRAINT "PK_AspNetUsers" PRIMARY KEY ("Id");

ALTER TABLE ONLY public."DatasetModels"
  ADD CONSTRAINT "PK_DatasetModels" PRIMARY KEY ("IDDataset");

ALTER TABLE ONLY public."PredictionBiopsyModels"
  ADD CONSTRAINT "PK_PredictionBiopsyModels" PRIMARY KEY ("IDPredictionBiopsy");

ALTER TABLE ONLY public."__EFMigrationsHistory"
  ADD CONSTRAINT "PK__EFMigrationsHistory" PRIMARY KEY ("MigrationId");

ALTER TABLE ONLY public."DatasetModels"
  ADD CONSTRAINT "FK_DatasetModels_AspNetUsers_FKIDUsers" FOREIGN KEY ("FKIDUsers")
REFERENCES public."AspNetUsers"("Id") ON DELETE RESTRICT;

ALTER TABLE ONLY public."PredictionBiopsyModels"
  ADD CONSTRAINT "FK_PredictionBiopsyModels_AspNetUsers_FKIDUsers" FOREIGN KEY
("FKIDUsers") REFERENCES public."AspNetUsers"("Id") ON DELETE RESTRICT;

CREATE INDEX "IX_DatasetModels_FKIDUsers" ON public."DatasetModels" USING btree
("FKIDUsers");

CREATE INDEX "IX_PredictionBiopsyModels_FKIDUsers" ON public."PredictionBiopsyModels" USING
btree ("FKIDUsers");

CREATE UNIQUE INDEX "AspNetUsersIndex" ON public."AspNetUsers" USING btree ("Id");

CREATE UNIQUE INDEX "IDDatasetIndex" ON public."DatasetModels" USING btree ("IDDataset");

CREATE UNIQUE INDEX "IDPredictionBiopsyIndex" ON public."PredictionBiopsyModels" USING btree
("IDPredictionBiopsy");

```

```

CREATE UNIQUE INDEX "MigrationIdIndex" ON public."__EFMigrationsHistory" USING btree
("MigrationId");

CREATE UNIQUE INDEX "IDTrainDatasetIndex" ON public."TrainDatasetModels" USING btree
("IDTrainDataset");

CREATE UNIQUE INDEX "IDTestDatasetIndex" ON public."TestDatasetModels" USING btree
("IDTestDataset");

--
-- PostgreSQL database dump complete
--

```

Основной программный код

appsettings.json

```

{
  "ConnectionStrings": {
    "DefaultConnection": "Host=localhost;Port=5432;Database=ccp;User ID=postgres;Password=root"
  },
  "Logging": {
    "LogLevel": {
      "Default": "Warning"
    }
  },
  "AllowedHosts": "*"
}

```

Program.cs

```

using System;
using System.Collections.Generic;
using System.IO;
using System.Linq;
using System.Threading.Tasks;
using Microsoft.AspNetCore;
using Microsoft.AspNetCore.Hosting;
using Microsoft.Extensions.Configuration;
using Microsoft.Extensions.Logging;

namespace CCP_Project
{
    public class Program
    {
        public static void Main(string[] args)
        {
            CreateWebHostBuilder(args).Build().Run();
        }

        public static IWebHostBuilder CreateWebHostBuilder(string[] args) =>
            WebHost.CreateDefaultBuilder(args)
                .UseStartup<Startup>();
    }
}

```

Startup.cs

```

using System;
using System.Collections.Generic;
using System.Linq;
using System.Threading.Tasks;
using Microsoft.AspNetCore.Builder;
using Microsoft.AspNetCore.Identity;
using Microsoft.AspNetCore.Hosting;
using Microsoft.AspNetCore.Http;
using Microsoft.AspNetCore.HttpsPolicy;
using Microsoft.AspNetCore.Mvc;
using Microsoft.EntityFrameworkCore;
using CCP_Project.Data;
using Microsoft.Extensions.Configuration;
using Microsoft.Extensions.DependencyInjection;
using CCP_Project.Models;

namespace CCP_Project
{
    public class Startup
    {
        public Startup(IConfiguration configuration)
        {
            Configuration = configuration;
        }

        public IConfiguration Configuration { get; }

        // This method gets called by the runtime. Use this method to add services to the container.
        public void ConfigureServices(IServiceCollection services)
        {
            services.Configure<CookiePolicyOptions>(options =>
            {
                // This lambda determines whether user consent for non-essential cookies is needed for a given
request.
                options.CheckConsentNeeded = context => true;
                options.MinimumSameSitePolicy = SameSiteMode.None;
            });

            services.AddEntityFrameworkNpgsql().AddDbContext<ApplicationDbContext>(options =>
                options.UseNpgsql(Configuration.GetConnectionString("DefaultConnection"))
            );

            services.Configure<CookiePolicyOptions>(options =>
            {
                // This lambda determines whether user consent for non-essential cookies is needed for a given
request.
                options.CheckConsentNeeded = context => true;
                options.MinimumSameSitePolicy = SameSiteMode.None;
            });

            services.AddDefaultIdentity<IdentityUser>()
                .AddEntityFrameworkStores<ApplicationDbContext>();

            services.AddAuthentication()
                .AddFacebook(facebookOptions =>
                {
                    facebookOptions.AppId = Configuration["Authentication:Facebook:AppId"];
                    facebookOptions.AppSecret = Configuration["Authentication:Facebook:AppSecret"];
                }
            );
        }
    }
}

```



```

    })
    .AddGoogle(googleOptions =>
    {
        googleOptions.ClientId = Configuration["Authentication:Google:ClientId"];
        googleOptions.ClientSecret = Configuration["Authentication:Google:ClientSecret"];
    })
    .AddMicrosoftAccount(microsoftOptions =>
    {
        microsoftOptions.ClientId = Configuration["Authentication:Microsoft:ApplicationId"];
        microsoftOptions.ClientSecret = Configuration["Authentication:Microsoft:Password"];
    });

    services.AddMvc().SetCompatibilityVersion(CompatibilityVersion.Version_2_1);
}

// This method gets called by the runtime. Use this method to configure the HTTP request pipeline.
public void Configure(IApplicationBuilder app, IHostingEnvironment env)
{
    if (env.IsDevelopment())
    {
        app.UseDeveloperExceptionPage();
        app.UseDatabaseErrorPage();
    }
    else
    {
        app.UseExceptionHandler("/Home/Error");
        app.UseHsts();
    }

    app.UseHttpsRedirection();
    app.UseStaticFiles();
    app.UseCookiePolicy();

    app.UseAuthentication();

    app.UseMvc(routes =>
    {
        routes.MapRoute(
            name: "default",
            template: "{controller=Home}/{action=Index}/{id?}");
    });
}
}
}

```

HomeController.cs

```

using System;
using System.Collections.Generic;
using System.Diagnostics;
using System.Linq;
using System.Threading.Tasks;
using Microsoft.AspNetCore.Mvc;
using CCP_Project.Models;

namespace CCP_Project.Controllers
{
    public class HomeController : Controller
    {

```

```
public IActionResult Index()
{
    return View();
}

public IActionResult About()
{
    ViewData["Message"] = "Your application description page.";

    return View();
}

public IActionResult Contact()
{
    ViewData["Message"] = "Anna Eliseeva";

    return View();
}

public IActionResult Privacy()
{
    return View();
}

[ResponseCache(Duration = 0, Location = ResponseCacheLocation.None, NoStore = true)]
public IActionResult Error()
{
    return View(new ErrorViewModel { RequestId = Activity.Current?.Id ?? HttpContext.TraceIdentifier
});
}
}
}
```

DatasetModelsController.cs

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Threading.Tasks;
using Microsoft.AspNetCore.Mvc;
using Microsoft.AspNetCore.Mvc.Rendering;
using Microsoft.EntityFrameworkCore;
using CCP_Project.Data;
using CCP_Project.Models;
using Microsoft.AspNetCore.Authorization;
using System.Security.Claims;
using Microsoft.Scripting.Hosting;
using IronPython.Hosting;
using Microsoft.AspNetCore.Hosting;
using System.Diagnostics;

namespace CCP_Project.Controllers
{
    [Authorize]
    public class DatasetModelsController : Controller
    {
        private readonly ApplicationDbContext _context;
        private readonly IHostingEnvironment _hostingEnvironment;
```

```

public DatasetModelsController(ApplicationDbContext context, IHostingEnvironment
hostingEnvironment)
{
    _context = context;
    _hostingEnvironment = hostingEnvironment;
}

public IActionResult Loader()
{
    ViewData["Message"] = "Loader";
    return View();
    //return NotFound();
    //return RedirectToAction("Index", "PredictionBiopsyModels");
}

// GET: DatasetModels
public async Task<IActionResult> Index()
{
    var applicationDbContext = _context.DatasetModels.Include(d => d.User);
    return View(await applicationDbContext.ToListAsync());
}

// GET: DatasetModels/Details/5
public async Task<IActionResult> Details(int? id)
{
    if (id == null)
    {
        return NotFound();
    }

    var datasetModel = await _context.DatasetModels
        .Include(d => d.User)
        .FirstOrDefaultAsync(m => m.IDDataset == id);
    if (datasetModel == null)
    {
        return NotFound();
    }

    return View(datasetModel);
}

// GET: DatasetModels/Create
public IActionResult Create()
{
    ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id");
    return View();
}

// POST: DatasetModels/Create
// To protect from overposting attacks, please enable the specific properties you want to bind to, for
// more details see http://go.microsoft.com/fwlink/?LinkId=317598.
[HttpPost]
[ValidateAntiForgeryToken]
public async Task<IActionResult>
Create([Bind("IDDataset, Age, NumberOfSexualPartners, AgeOfFirstSexualIntercourse, NumberOfPregnancies, Smokes,
AgeSmokes, PacksYearSmokes, UsingHormonalContraceptives, AgeUsingHormonalContraceptives, IUD, AgeIUD, STDs,
NumberOfSTDs, STDs_Condylomatosis, STDs_CervicalCondylomatosis, STDs_VaginalCondylomatosis, STDs_VulvoP
erinealCondylomatosis, STDs_Syphilis, STDs_PelvicInflammatoryDisease, STDs_GenitalHerpes, STDs_MolluscumCont
agiosum, STDs_AIDS, STDs_HIV, STDs_HepatitisB, STDs_HPv, STDs_NumberOfDiagnosis, STDs_TimeSinceFirstDia
gnosis, STDs_TimeSinceLastDiagnosis, Dx_Cancer, Dx_CIN, Dx_HPv, Dx_HinselmannTest, SchillerTest, Citology, FKID
Users")] DatasetModel datasetModel)
{

```

```

if (ModelState.IsValid)
{
    var dt = DateTime.Now;
    datasetModel.FKIDUsers = User.FindFirst(ClaimTypes.NameIdentifier).Value;

    _context.Add(datasetModel);

    await _context.SaveChangesAsync();

ExecutePythonScript($"{_hostingEnvironment.ContentRootPath}\\Python\\scripts\\Cervical_cancer_prediction_new.py
");

    return RedirectToAction(nameof(Loader));
    //return RedirectToAction("Index", "PredictionBiopsyModels");
}

ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id", datasetModel.FKIDUsers);

return View(datasetModel);
}

private void ExecutePythonScript(string scriptPath)
{
    ProcessStartInfo startInfo = new ProcessStartInfo();
    startInfo.FileName = "C:\\Windows\\System32\\cmd.exe";
    startInfo.Arguments = $"/c C:\\ProgramData\\Anaconda3\\python.exe \"{scriptPath}\"";
    startInfo.WorkingDirectory = "C:\\Development\\ASP.Net
Core\\CCP_Project\\CCP_Project\\Python\\scripts\\";

    Process.Start(startInfo);
}

// GET: DatasetModels/Edit/5
public async Task<IActionResult> Edit(int? id)
{
    if (id == null)
    {
        return NotFound();
    }

    var datasetModel = await _context.DatasetModels.FindAsync(id);
    if (datasetModel == null)
    {
        return NotFound();
    }
    ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id", datasetModel.FKIDUsers);
    return View(datasetModel);
}

// POST: DatasetModels/Edit/5
// To protect from overposting attacks, please enable the specific properties you want to bind to, for
// more details see http://go.microsoft.com/fwlink/?LinkId=317598.
[HttpPost]
[ValidateAntiForgeryToken]
public async Task<IActionResult> Edit(int id,
[Bind("IDDataset, Age, NumberOfSexualPartners, AgeOfFirstSexualIntercourse, NumberOfPregnancies, Smokes, AgeSmokes, PacksYearSmokes, UsingHormonalContraceptives, AgeUsingHormonalContraceptives, IUD, AgeIUD, STDs, NumberOfSTDs, STDs_Condylomatosis, STDs_CervicalCondylomatosis, STDs_VaginalCondylomatosis, STDs_VulvoPerinealCondylomatosis, STDs_Syphilis, STDs_PelvicInflammatoryDisease, STDs_GenitalHerpes, STDs_MolluscumContagiosum

```

```

,STDs_AIDS,STDs_HIV,STDs_HepatitisB,STDs_HP V,STDs_NumberOfDiagnosis,STDs_TimeSinceFirstDiagnosis,STDs_TimeSinceLastDiagnosis,Dx_Cancer,Dx_CIN,Dx_HP V,Dx,HinselmannTest,SchillerTest,Citology,FKIDUsers"]
DatasetModel datasetModel)
{
    if (id != datasetModel.IDDataset)
    {
        return NotFound();
    }

    if (ModelState.IsValid)
    {
        try
        {
            _context.Update(datasetModel);
            await _context.SaveChangesAsync();
        }
        catch (DbUpdateConcurrencyException)
        {
            if (!DatasetModelExists(datasetModel.IDDataset))
            {
                return NotFound();
            }
            else
            {
                throw;
            }
        }
        return RedirectToAction(nameof(Index));
    }
    ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id", datasetModel.FKIDUsers);
    return View(datasetModel);
}

// GET: DatasetModels/Delete/5
public async Task<IActionResult> Delete(int? id)
{
    if (id == null)
    {
        return NotFound();
    }

    var datasetModel = await _context.DatasetModels
        .Include(d => d.User)
        .FirstOrDefaultAsync(m => m.IDDataset == id);
    if (datasetModel == null)
    {
        return NotFound();
    }

    return View(datasetModel);
}

// POST: DatasetModels/Delete/5
[HttpPost, ActionName("Delete")]
[ValidateAntiForgeryToken]
public async Task<IActionResult> DeleteConfirmed(int id)
{
    var datasetModel = await _context.DatasetModels.FindAsync(id);
    _context.DatasetModels.Remove(datasetModel);
    await _context.SaveChangesAsync();
    return RedirectToAction(nameof(Index));
}

```

```

private bool DatasetModelExists(int id)
{
    return _context.DatasetModels.Any(e => e.IDDataset == id);
}
}
}

```

PredictionBiopsyModelsController.cs

```

using System;
using System.Collections.Generic;
using System.Linq;
using System.Threading.Tasks;
using Microsoft.AspNetCore.Mvc;
using Microsoft.AspNetCore.Mvc.Rendering;
using Microsoft.EntityFrameworkCore;
using CCP_Project.Data;
using CCP_Project.Models;

namespace CCP_Project.Controllers
{
    public class PredictionBiopsyModelsController : Controller
    {
        private readonly ApplicationDbContext _context;

        public PredictionBiopsyModelsController(ApplicationDbContext context)
        {
            _context = context;
        }

        // GET: PredictionBiopsyModels
        public async Task<IActionResult> Index()
        {
            var applicationDbContext = _context.PredictionBiopsyModels.Include(p => p.User);
            var test = await applicationDbContext.ToListAsync();
            return View(await applicationDbContext.ToListAsync());
        }

        // GET: PredictionBiopsyModels/Details/5
        public async Task<IActionResult> Details(int? id)
        {
            if (id == null)
            {
                return NotFound();
            }

            var predictionBiopsyModel = await _context.PredictionBiopsyModels
                .Include(p => p.User)
                .FirstOrDefaultAsync(m => m.IDPredictionBiopsy == id);
            if (predictionBiopsyModel == null)
            {
                return NotFound();
            }

            return View(predictionBiopsyModel);
        }

        // GET: PredictionBiopsyModels/Create

```

```

public IActionResult Create()
{
    ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id");
    return View();
}

// POST: PredictionBiopsyModels/Create
// To protect from overposting attacks, please enable the specific properties you want to bind to, for
// more details see http://go.microsoft.com/fwlink/?LinkId=317598.
[HttpPost]
[ValidateAntiForgeryToken]
public async Task<IActionResult>
Create([Bind("IDPredictionBiopsy,PredictionBiopsy,DateTime,FKIDUsers")]
predictionBiopsyModel)
{
    if (ModelState.IsValid)
    {
        _context.Add(predictionBiopsyModel);
        await _context.SaveChangesAsync();
        return RedirectToAction(nameof(Index));
    }
    ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id",
predictionBiopsyModel.FKIDUsers);
    return View(predictionBiopsyModel);
}

// GET: PredictionBiopsyModels/Edit/5
public async Task<IActionResult> Edit(int? id)
{
    if (id == null)
    {
        return NotFound();
    }

    var predictionBiopsyModel = await _context.PredictionBiopsyModels.FindAsync(id);
    if (predictionBiopsyModel == null)
    {
        return NotFound();
    }
    ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id",
predictionBiopsyModel.FKIDUsers);
    return View(predictionBiopsyModel);
}

// POST: PredictionBiopsyModels/Edit/5
// To protect from overposting attacks, please enable the specific properties you want to bind to, for
// more details see http://go.microsoft.com/fwlink/?LinkId=317598.
[HttpPost]
[ValidateAntiForgeryToken]
public async Task<IActionResult> Edit(int id,
[Bind("IDPredictionBiopsy,PredictionBiopsy,DateTime,FKIDUsers")]
predictionBiopsyModel)
{
    if (id != predictionBiopsyModel.IDPredictionBiopsy)
    {
        return NotFound();
    }

    if (ModelState.IsValid)
    {
        try
        {

```

```

        _context.Update(predictionBiopsyModel);
        await _context.SaveChangesAsync();
    }
    catch (DbUpdateConcurrencyException)
    {
        if (!PredictionBiopsyModelExists(predictionBiopsyModel.IDPredictionBiopsy))
        {
            return NotFound();
        }
        else
        {
            throw;
        }
    }
    return RedirectToAction(nameof(Index));
}
ViewData["FKIDUsers"] = new SelectList(_context.Users, "Id", "Id",
predictionBiopsyModel.FKIDUsers);
return View(predictionBiopsyModel);
}

// GET: PredictionBiopsyModels/Delete/5
public async Task<IActionResult> Delete(int? id)
{
    if (id == null)
    {
        return NotFound();
    }

    var predictionBiopsyModel = await _context.PredictionBiopsyModels
        .Include(p => p.User)
        .FirstOrDefaultAsync(m => m.IDPredictionBiopsy == id);
    if (predictionBiopsyModel == null)
    {
        return NotFound();
    }

    return View(predictionBiopsyModel);
}

// POST: PredictionBiopsyModels/Delete/5
[HttpPost, ActionName("Delete")]
[ValidateAntiForgeryToken]
public async Task<IActionResult> DeleteConfirmed(int id)
{
    var predictionBiopsyModel = await _context.PredictionBiopsyModels.FindAsync(id);
    _context.PredictionBiopsyModels.Remove(predictionBiopsyModel);
    await _context.SaveChangesAsync();
    return RedirectToAction(nameof(Index));
}

private bool PredictionBiopsyModelExists(int id)
{
    return _context.PredictionBiopsyModels.Any(e => e.IDPredictionBiopsy == id);
}
}
}

```


DatasetModel.cs

```

using Microsoft.AspNetCore.Identity;
using System;
using System.Collections.Generic;
using System.ComponentModel.DataAnnotations;
using System.ComponentModel.DataAnnotations.Schema;
using System.Linq;
using System.Threading.Tasks;

namespace CCP_Project.Models
{
    public class DatasetModel
    {
        [Key]
        public int IDDataset { get; set; }

        [Required]
        [Range(12, 100)]
        public int Age { get; set; }

        [Required]
        [Range(0, 300)]
        [Display(Name = "Number of sexual partners")]
        public int NumberOfSexualPartners { get; set; }

        [Range(0, 100)]
        [Display(Name = "Age of first sexual intercourse")]
        public int AgeOfFirstSexualIntercourse { get; set; }

        [Required]
        [Range(0, 20)]
        [Display(Name = "Number of pregnancies")]
        public int NumberOfPregnancies { get; set; }

        [Display(Name = "Smoking")]
        public bool Smokes { get; set; }

        [Display(Name = "Smoking period (years)")]
        public int AgeSmokes { get; set; }

        [Display(Name = "Number of packs of cigarettes per year")]
        public int PacksYearSmokes { get; set; }

        [Display(Name = "Using hormonal contraceptives")]
        public bool UsingHormonalContraceptives { get; set; }

        [Display(Name = "Using hormonal contraceptives (years)")]
        public int AgeUsingHormonalContraceptives { get; set; }

        [Display(Name = "Using intrauterine device")]
        public bool IUD { get; set; }

        [Display(Name = "Using intrauterine device (years)")]
        public int AgeIUD { get; set; }

        [Display(Name = "Sexually transmitted diseases")]
        public bool STDs { get; set; }

        [Display(Name = "Number of sexually transmitted diseases")]
    }
}

```

```

public int NumberOfSTDs { get; set; }

[Display(Name = "Sexually transmitted diseases: condylomatosis")]
public bool STDs_Condylomatosis { get; set; }

[Display(Name = "Sexually transmitted diseases: cervical condylomatosis")]
public bool STDs_CervicalCondylomatosis { get; set; }

[Display(Name = "Sexually transmitted diseases: vaginal condylomatosis")]
public bool STDs_VaginalCondylomatosis { get; set; }

[Display(Name = "Sexually transmitted diseases: vulvo perineal condylomatosis")]
public bool STDs_VulvoPerinealCondylomatosis { get; set; }

[Display(Name = "Sexually transmitted diseases: syphilis")]
public bool STDs_Syphilis { get; set; }

[Display(Name = "Sexually transmitted diseases: pelvic inflammatory disease")]
public bool STDs_PelvicInflammatoryDisease { get; set; }

[Display(Name = "Sexually transmitted diseases: genital herpes")]
public bool STDs_GenitalHerpes { get; set; }

[Display(Name = "Sexually transmitted diseases: molluscum contagiosum")]
public bool STDs_MolluscumContagiosum { get; set; }

[Display(Name = "AIDS (acquired immune deficiency syndrome)")]
public bool STDs_AIDS { get; set; }

[Display(Name = "HIV (human immunodeficiency viruses)")]
public bool STDs_HIV { get; set; }

[Display(Name = "Hepatitis B")]
public bool STDs_HepatitisB { get; set; }

[Display(Name = "HPV (human papillomavirus)")]
public bool STDs_HPVP { get; set; }

[Display(Name = "Sexually transmitted diseases: number of diagnosis")]
public bool STDs_NumberOfDiagnosis { get; set; }

[Display(Name = "Sexually transmitted diseases: time since first diagnosis")]
public bool STDs_TimeSinceFirstDiagnosis { get; set; }

[Display(Name = "Sexually transmitted diseases: time since last diagnosis")]
public bool STDs_TimeSinceLastDiagnosis { get; set; }

[Display(Name = "Cervical cancer diagnosis")]
public bool Dx_Cancer { get; set; }

[Display(Name = "CIN (cervical intraepithelial neoplasia) diagnosis")]
public bool Dx_CIN { get; set; }

[Display(Name = "HPV (human papillomavirus) diagnosis")]
public bool Dx_HPVP { get; set; }

[Display(Name = "Cancer diagnosis")]
public bool Dx { get; set; }

[Display(Name = "Hinselmann test")]
public bool HinselmannTest { get; set; }

```

```

    [Display(Name = "Schiller test")]
    public bool SchillerTest { get; set; }

    [Display(Name = "Citology")]
    public bool Citology { get; set; }

    public string FKIDUsers { get; set; }

    [ForeignKey("FKIDUsers")]
    public virtual IdentityUser User { get; set; }
}
}

```

PredictionBiopsyModel.cs

```

using Microsoft.AspNetCore.Identity;
using System;
using System.Collections.Generic;
using System.ComponentModel.DataAnnotations;
using System.ComponentModel.DataAnnotations.Schema;
using System.Linq;
using System.Threading.Tasks;

namespace CCP_Project.Models
{
    public class PredictionBiopsyModel
    {
        [Key]
        public int IDPredictionBiopsy { get; set; }

        public int PredictionBiopsy { get; set; }

        [Display(Name = "Date Time")]
        public DateTime DateTime { get; set; }

        public string FKIDUsers { get; set; }

        [ForeignKey("FKIDUsers")]
        public virtual IdentityUser User { get; set; }
    }
}

```

TrainDatasetModel.cs

```

using System;
using System.Collections.Generic;
using System.ComponentModel.DataAnnotations;
using System.Linq;
using System.Threading.Tasks;

namespace CCP_Project.Models
{
    public class TrainDatasetModel
    {
        [Key]

```

```
public int IDTrainDataset { get; set; }

public int Age { get; set; }

public int NumberOfSexualPartners { get; set; }

public int AgeOfFirstSexualIntercourse { get; set; }

public int NumberOfPregnancies { get; set; }

public int AgeSmokes { get; set; }

public int PacksYearSmokes { get; set; }

public int AgeUsingHormonalContraceptives { get; set; }

public int AgeIUD { get; set; }

public int NumberOfSTDs { get; set; }

public bool STDs_Condylomatosis { get; set; }

public bool STDs_CervicalCondylomatosis { get; set; }

public bool STDs_VaginalCondylomatosis { get; set; }

public bool STDs_VulvoPerinealCondylomatosis { get; set; }

public bool STDs_Syphilis { get; set; }

public bool STDs_PelvicInflammatoryDisease { get; set; }

public bool STDs_GenitalHerpes { get; set; }

public bool STDs_MolluscumContagiosum { get; set; }

public bool STDs_AIDS { get; set; }

public bool STDs_HIV { get; set; }

public bool STDs_HepatitisB { get; set; }

public bool STDs_HPVP { get; set; }

public bool STDs_NumberOfDiagnosis { get; set; }

public bool STDs_TimeSinceFirstDiagnosis { get; set; }

public bool STDs_TimeSinceLastDiagnosis { get; set; }

public bool Biopsy { get; set; }

public bool Smokes_0 { get; set; }

public bool Smokes_1 { get; set; }

public bool UsingHormonalContraceptives_0 { get; set; }

public bool UsingHormonalContraceptives_1 { get; set; }

public bool IUD_0 { get; set; }
```

```

public bool IUD_1 { get; set; }

public bool STDs_0 { get; set; }

public bool STDs_1 { get; set; }

public bool Dx_Cancer_0 { get; set; }

public bool Dx_Cancer_1 { get; set; }

public bool Dx_CIN_0 { get; set; }

public bool Dx_CIN_1 { get; set; }

public bool Dx_HPVS_0 { get; set; }

public bool Dx_HPVS_1 { get; set; }

public bool Dx_0 { get; set; }

public bool Dx_1 { get; set; }

public bool HinselmannTest_0 { get; set; }

public bool HinselmannTest_1 { get; set; }

public bool Citology_0 { get; set; }

public bool Citology_1 { get; set; }

public bool SchillerTest_0 { get; set; }

public bool SchillerTest_1 { get; set; }
}
}

```

TestDatasetModel.cs

```

using System;
using System.Collections.Generic;
using System.ComponentModel.DataAnnotations;
using System.Linq;
using System.Threading.Tasks;

namespace CCP_Project.Models
{
    public class TestDatasetModel
    {
        [Key]
        public int IDTestDataset { get; set; }

        public int Age { get; set; }

        public int NumberOfSexualPartners { get; set; }

        public int AgeOfFirstSexualIntercourse { get; set; }

        public int NumberOfPregnancies { get; set; }
    }
}

```

```
public int AgeSmokes { get; set; }

public int PacksYearSmokes { get; set; }

public int AgeUsingHormonalContraceptives { get; set; }

public int AgeIUD { get; set; }

public int NumberOfSTDs { get; set; }

public bool STDs_Condylomatosis { get; set; }

public bool STDs_CervicalCondylomatosis { get; set; }

public bool STDs_VaginalCondylomatosis { get; set; }

public bool STDs_VulvoPerinealCondylomatosis { get; set; }

public bool STDs_Syphilis { get; set; }

public bool STDs_PelvicInflammatoryDisease { get; set; }

public bool STDs_GenitalHerpes { get; set; }

public bool STDs_MolluscumContagiosum { get; set; }

public bool STDs_AIDS { get; set; }

public bool STDs_HIV { get; set; }

public bool STDs_HepatitisB { get; set; }

public bool STDs_HPVP { get; set; }

public bool STDs_NumberOfDiagnosis { get; set; }

public bool STDs_TimeSinceFirstDiagnosis { get; set; }

public bool STDs_TimeSinceLastDiagnosis { get; set; }

public bool Biopsy { get; set; }

public bool Smokes_0 { get; set; }

public bool Smokes_1 { get; set; }

public bool UsingHormonalContraceptives_0 { get; set; }

public bool UsingHormonalContraceptives_1 { get; set; }

public bool IUD_0 { get; set; }

public bool IUD_1 { get; set; }

public bool STDs_0 { get; set; }

public bool STDs_1 { get; set; }

public bool Dx_Cancer_0 { get; set; }

public bool Dx_Cancer_1 { get; set; }
```

```

public bool Dx_CIN_0 { get; set; }

public bool Dx_CIN_1 { get; set; }

public bool Dx_HPВ_0 { get; set; }

public bool Dx_HPВ_1 { get; set; }

public bool Dx_0 { get; set; }

public bool Dx_1 { get; set; }

public bool HinselmannTest_0 { get; set; }

public bool HinselmannTest_1 { get; set; }

public bool Citology_0 { get; set; }

public bool Citology_1 { get; set; }

public bool SchillerTest_0 { get; set; }

public bool SchillerTest_1 { get; set; }
}
}

```

Index.cshtml представления «Home»

```

@{
    ViewData["Title"] = "Home Page";
}

<div id="carouselExampleIndicators" class="carousel slide" data-ride="carousel">
    <ol class="carousel-indicators">
        <li data-target="#carouselExampleIndicators" data-slide-to="0" class="active"></li>
        <li data-target="#carouselExampleIndicators" data-slide-to="1"></li>
        <li data-target="#carouselExampleIndicators" data-slide-to="2"></li>
    </ol>
    <div class="carousel-inner">
        <div class="carousel-item active">
            
        </div>
        <div class="carousel-item">
            
        </div>
        <div class="carousel-item">
            
        </div>
    </div>
    <a class="carousel-control-prev" href="#carouselExampleIndicators" role="button" data-slide="prev">
        <span class="carousel-control-prev-icon" aria-hidden="true"></span>
        <span class="sr-only">Previous</span>
    </a>
    <a class="carousel-control-next" href="#carouselExampleIndicators" role="button" data-slide="next">
        <span class="carousel-control-next-icon" aria-hidden="true"></span>
        <span class="sr-only">Next</span>
    </a>
</div>

```

```

<div class="row">
  <div class="col-md-3">
    <h2>What is Cervical Cancer?</h2>
    <p>
      Cervical cancer is a cancer arising from the cervix.
      It is due to the abnormal growth of cells that have the ability to invade or spread to other parts of the
body.
    </p>
    <ul>
      <li><a href="https://www.medicalnewstoday.com/articles/159821.php">What you need to know
about cervical cancer</a></li>
      <li><a href="https://www.nccc-online.org/hpvcervical-cancer/cervical-cancer-overview/">Cervical
Cancer Overview</a></li>
    </ul>
  </div>
  <div class="col-md-3">
    <h2>Symptoms</h2>
    <p>
      Cervical cancer is a cancer arising from the cervix.
      It is due to the abnormal growth of cells that have the ability to invade or spread to other parts of the
body.
    </p>
    <ul>
      <li>bleeding between periods</li>
      <li>bleeding after sexual intercourse</li>
      <li>bleeding in post-menopausal women</li>
      <li>discomfort during sexual intercourse</li>
      <li>vaginal discharge with a strong odor</li>
      <li>vaginal discharge tinged with blood</li>
      <li>pelvic pain</li>
    </ul>
    <p>
      More detail information about symptoms of cervical cancer you can get follow links:
    </p>
    <ul>
      <li><a href="https://www.nhs.uk/conditions/cervical-cancer/symptoms/">Symptoms of cervical
cancer</a></li>
      <li><a
href="https://www.medicinenet.com/cancer/article.htm#what_are_cancer_symptoms_and_signs">What are cancer
symptoms and signs?</a></li>
    </ul>
  </div>
  <div class="col-md-3">
    <h2>Diagnosis</h2>
    <ul>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Pelvic
examination</a></li>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Pap test</a></li>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">HPV typing</a></li>
    </ul>
    <p>
      If the Pap test showed some abnormal cells and the HPV test is positive,
      then the doctor may suggest 1 or more of the following diagnostic tests:
    </p>
    <ul>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Colposcopy</a></li>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Biopsy</a></li>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Pelvic examination
under anesthesia</a></li>
      <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">X-ray</a></li>

```



```

        <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Computed tomography
(CT or CAT) scan</a></li>
        <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Magnetic resonance
imaging (MRI)</a></li>
        <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Positron emission
tomography (PET) or PET-CT scan</a></li>
        <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Cystoscopy</a></li>
        <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Proctoscopy</a></li>
        <li><a href="https://www.cancer.net/cancer-types/cervical-cancer/diagnosis">Laparoscopy</a></li>
    </ul>
</div>
<div class="col-md-3">
    <h2>Treatment</h2>
    <p>
        Treatment for cervical cancer depends on several factors, such as the stage of the cancer,
        other health problems you may have and your preferences.
    </p>
    <ul>
        <li><a href="https://www.mayoclinic.org/diseases-conditions/cervical-cancer/diagnosis-treatment/drc-
20352506">Surgery</a></li>
        <li><a href="https://www.mayoclinic.org/diseases-conditions/cervical-cancer/diagnosis-treatment/drc-
20352506">Radiation</a></li>
        <li><a href="https://www.mayoclinic.org/diseases-conditions/cervical-cancer/diagnosis-treatment/drc-
20352506">Chemotherapy</a></li>
    </ul>
</div>
</div>

```

Contact.cshtml представлення «Home»

```

@{
    ViewData["Title"] = "Contact";
}
<h2>@ViewData["Title"]</h2>
<h3>@ViewData["Message"]</h3>

<address>
    Location: Sumy, Ukraine<br />
    <strong>Phone:</strong> <a href="tel:+38 (066) 675 64 76">+38 (066) 675 64 76</a><br />
    @*<abbr title="Phone">P:</abbr>
    425.555.0100*@
</address>

<address>
    <strong>Support:</strong> <a href="mailto:anna.eliseeva.3000@gmail.com">anna.eliseeva.3000@gmail.com</a><br />
    @*<strong>Marketing:</strong> <a href="mailto:Marketing@example.com">Marketing@example.com</a>*@
</address>

```

Create.cshtml представлення «DatasetModels»

```
@model CCP_Project.Models.DatasetModel
```

```
@{
```

```

    ViewData["Title"] = "Create";
}

<script src="~/js/IfCheckboxIsTrue.js"></script>
<script src="~/js/NumberOfSexualPartners.js"></script>

@*<link rel="stylesheet" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.0/css/bootstrap.min.css">*@
<script src="https://ajax.googleapis.com/ajax/libs/jquery/3.4.0/jquery.min.js"></script>
<script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.0/js/bootstrap.min.js"></script>

<h2 align="center">Cervical cancer prediction</h2>

<h4 align="center">Individual characteristics</h4>
<hr />
<div class="row">
    @*<div class="col-md-4">*@
    <div>
        <form asp-action="Create">
            <div asp-validation-summary="ModelOnly" class="text-danger"></div>
            <div class="container">
                <fieldset>
                    <legend>Age</legend>
                    <div class="row">
                        @* First column *@
                        <div class="col-sm-4">
                            <div class="form-group">
                                <label asp-for="Age">Age</label>
                                <input asp-for="Age" class="form-control" />
                                <span asp-validation-for="Age" class="text-danger"></span>
                            </div>
                        </div>
                    </div>
                </fieldset>
                <fieldset>
                    <legend>Existing of sexual intercourse</legend>
                    <div class="row">
                        <div class="col-sm-4">
                            <div class="form-group">
                                <label asp-for="NumberOfSexualPartners">Number of sexual partners</label>
                                <input asp-for="NumberOfSexualPartners" id="sexpart" class="form-control"
onkeyup="Bar()" />
                                <span asp-validation-for="NumberOfSexualPartners" class="text-danger"></span>
                            </div>
                        </div>
                        <div class="col-sm-4">
                            <div class="form-group">
                                <label asp-for="AgeOfFirstSexualIntercourse" class="control-label">Age of first sexual
intercourse</label>
                                <input asp-for="AgeOfFirstSexualIntercourse" class="form-control" id="hidden_field1"
disabled />
                                <span asp-validation-for="AgeOfFirstSexualIntercourse" class="text-danger"></span>
                            </div>
                        </div>
                        <div class="col-sm-4">
                            <div class="form-group">
                                <label asp-for="NumberOfPregnancies" class="control-label">Number of
pregnancies</label>
                                <input asp-for="NumberOfPregnancies" class="form-control" id="hidden_field2" disabled
/>
                                <span asp-validation-for="NumberOfPregnancies" class="text-danger"></span>
                            </div>
                        </div>
                    </div>
                </fieldset>
            </div>
        </form>
    </div>

```

```

        </div>
    </fieldset>
    <fieldset>
        <legend>Smoking</legend>
        <div class="row">
            <div class="col-sm-4">
                <div class="form-group">
                    <div class="checkbox" style="padding-top: 25px">
                        <label>
                            Do you smoke? <input asp-for="Smokes" id="trigger1"
onlick="IfCheckboxIsTrue1()" /> @@Html.DisplayNameFor(model => model.Smokes)*@
                        </label>
                    </div>
                </div>
            </div>
            <div class="col-sm-4">
                <div class="form-group">
                    <label asp-for="AgeSmokes" class="control-label">How many years do you
smoke?</label>
                    <input asp-for="AgeSmokes" class="form-control" id="hidden_field3" disabled />
                    <span asp-validation-for="AgeSmokes" class="text-danger"></span>
                </div>
            </div>
            <div class="col-sm-4">
                <div class="form-group">
                    <label asp-for="PacksYearSmokes" class="control-label">How many packs of cigarettes
do you smoke per year?</label>
                    <input asp-for="PacksYearSmokes" class="form-control" id="hidden_field4" disabled />
                    <span asp-validation-for="PacksYearSmokes" class="text-danger"></span>
                </div>
            </div>
        </div>
    </fieldset>
    <fieldset>
        <legend>Using hormonal contraceptives</legend>
        <div class="row">
            <div class="col-sm-4">
                <div class="form-group">
                    <div class="checkbox" style="padding-top: 25px">
                        <label>
                            Do you use the hormonal contraceptives? <input asp-
for="UsingHormonalContraceptives" id="trigger2" onlick="IfCheckboxIsTrue2()" />
@@Html.DisplayNameFor(model => model.UsingHormonalContraceptives)*@
                        </label>
                    </div>
                </div>
            </div>
            <div class="col-sm-4">
                <div class="form-group">
                    <label asp-for="AgeUsingHormonalContraceptives" class="control-label">How many
years do you use the hormonal contraceptives?</label>
                    <input asp-for="AgeUsingHormonalContraceptives" class="form-control"
id="hidden_field5" disabled />
                    <span asp-validation-for="AgeUsingHormonalContraceptives" class="text-
danger"></span>
                </div>
            </div>
        </div>
    </fieldset>
    <fieldset>
        <legend>Intrauterine device</legend>
        <div class="row">

```

```

<div class="col-sm-4">
  <div class="form-group">
    <div class="checkbox" style="padding-top: 25px">
      <label>
        Do you use the intrauterine device? <input asp-for="IUD" id="trigger3"
onclick="IfCheckboxIsTrue3()" /> @*@Html.DisplayNameFor(model => model.IUD)*@
      </label>
    </div>
  </div>
</div>
<div class="col-sm-4">
  <div class="form-group">
    <label asp-for="AgeIUD" class="control-label">How many years do you use the
intrauterine device?</label>
    <input asp-for="AgeIUD" class="form-control" id="hidden_field6" disabled />
    <span asp-validation-for="AgeIUD" class="text-danger"></span>
  </div>
</div>
</div>
</fieldset>

@*<div class="container">*@

<fieldset>
  <legend>Sexually transmitted diseases</legend>
  <div class="row">
    <div class="col-sm-4">
      <div class="form-group">
        <div class="checkbox" style="padding-top: 25px">
          <label>
            Do you have sexually transmitted diseases? <input asp-for="STDs" id="trigger4"
onclick="IfCheckboxIsTrue4()" /> @*@Html.DisplayNameFor(model => model.STDs)*@
          </label>
        </div>
      </div>
      <div class="form-group">
        <div class="checkbox">
          <label>
            Do you have cervical condylomatosis? <input asp-
for="STDs_CervicalCondylomatosis" id="hidden_field7" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_CervicalCondylomatosis)*@
          </label>
        </div>
      </div>
      <div class="form-group">
        <div class="checkbox">
          <label>
            Do you have syphilis? <input asp-for="STDs_Syphilis" id="hidden_field8" disabled
/> @*@Html.DisplayNameFor(model => model.STDs_Syphilis)*@
          </label>
        </div>
      </div>
      <div class="form-group">
        <div class="checkbox">
          <label>
            Do you have molluscum contagiosum? <input asp-
for="STDs_MolluscumContagiosum" id="hidden_field9" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_MolluscumContagiosum)*@
          </label>
        </div>
      </div>
    </div>
  </div>
</fieldset>

```

```

        <div class="checkbox">
            <label>
                Do you have hepatitis B? <input asp-for="STDs_HepatitisB" id="hidden_field10"
disabled /> @*@Html.DisplayNameFor(model => model.STDs_HepatitisB)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                How much time has passed since the first diagnosis? <input asp-
for="STDs_TimeSinceFirstDiagnosis" id="hidden_field11" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_TimeSinceFirstDiagnosis)*@
            </label>
        </div>
    </div>
</div>

@* Second column *@
<div class="col-sm-4">
    <div class="form-group">
        <label asp-for="NumberOfSTDs" class="control-label">How many sexually transmitted
disease diagnoses have been determined?</label>
        <input asp-for="NumberOfSTDs" class="form-control" id="hidden_field12" disabled />
        <span asp-validation-for="NumberOfSTDs" class="text-danger"></span>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have vaginal condylomatosis? <input asp-
for="STDs_VaginalCondylomatosis" id="hidden_field13" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_VaginalCondylomatosis)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have pelvic inflammatory disease? <input asp-
for="STDs_PelvicInflammatoryDisease" id="hidden_field14" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_PelvicInflammatoryDisease)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have AIDS (acquired immune deficiency syndrome)? <input asp-
for="STDs_AIDS" id="hidden_field15" disabled /> @*@Html.DisplayNameFor(model => model.STDs_AIDS)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have HPV (human papillomavirus)? <input asp-for="STDs_HPV"
id="hidden_field16" disabled /> @*@Html.DisplayNameFor(model => model.STDs_HPV)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">

```

```

        <label>
            How much time has passed since the last diagnosis? <input asp-
for="STDs_TimeSinceLastDiagnosis" id="hidden_field17" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_TimeSinceLastDiagnosis)*@
        </label>
    </div>
</div>
</div>

@* Third column *@
<div class="col-sm-4">
    <div class="form-group">
        <div class="checkbox" style="padding-top: 25px">
            <label>
                Do you have condylomatosis? <input asp-for="STDs_Condylomatosis"
id="hidden_field18" disabled /> @*@Html.DisplayNameFor(model => model.STDs_Condylomatosis)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have vulvo perineal condylomatosis? <input asp-
for="STDs_VulvoPerinealCondylomatosis" id="hidden_field19" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_VulvoPerinealCondylomatosis)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have genital herpes? <input asp-for="STDs_GenitalHerpes"
id="hidden_field20" disabled /> @*@Html.DisplayNameFor(model => model.STDs_GenitalHerpes)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Do you have HIV (human immunodeficiency viruses)? <input asp-for="STDs_HIV"
id="hidden_field21" disabled /> @*@Html.DisplayNameFor(model => model.STDs_HIV)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Enter number of sexually transmitted diagnosis: <input asp-
for="STDs_NumberOfDiagnosis" id="hidden_field22" disabled /> @*@Html.DisplayNameFor(model =>
model.STDs_NumberOfDiagnosis)*@
            </label>
        </div>
    </div>
</div>
</div>
</div>
</div>
</div>

<fieldset>
<legend>Diagnosis</legend>
<div class="row">
    <div class="col-sm-4">
        <div class="form-group">

```

```

        <div class="checkbox">
            <label>
                Have you ever been diagnosed with cancer? <input asp-for="Dx_Cancer" />
@@@Html.DisplayNameFor(model => model.Dx_Cancer)*@
            </label>
        </div>
    </div>
    <div class="form-group">
        <div class="checkbox">
            <label>
                Have you ever been diagnosed with CIN (cervical intraepithelial neoplasia)? <input
asp-for="Dx_CIN" /> @@@Html.DisplayNameFor(model => model.Dx_CIN)*@
            </label>
        </div>
    </div>
    <div class="col-sm-4">
        <div class="form-group">
            <div class="checkbox">
                <label>
                    Have you ever been diagnosed with HPV (human papillomavirus)? <input asp-
for="Dx_HPV" /> @@@Html.DisplayNameFor(model => model.Dx_HPV)*@
                </label>
            </div>
        </div>
    </div>
    <div class="col-sm-4">
        <div class="form-group">
            <div class="checkbox">
                <label>
                    Have you ever been diagnosed with cervical cancer? <input asp-for="Dx" />
@@@Html.DisplayNameFor(model => model.Dx)*@
                </label>
            </div>
        </div>
    </div>
</div>
</fieldset>
<fieldset>
    <legend>Clinical tests</legend>
    <div class="row">
        <div class="col-sm-4">
            <div class="form-group">
                <div class="checkbox">
                    <label>
                        Do you have a positive Hinselmann test result? <input asp-for="HinselmannTest" />
@@@Html.DisplayNameFor(model => model.HinselmannTest)*@
                    </label>
                </div>
            </div>
        </div>
    </div>
    <div class="col-sm-4">
        <div class="form-group">
            <div class="checkbox">
                <label>
                    Do you have a positive Schiller test result? <input asp-for="SchillerTest" />
@@@Html.DisplayNameFor(model => model.SchillerTest)*@
                </label>
            </div>
        </div>
    </div>
</div>
</div class="col-sm-4">

```

```

        <div class="form-group">
            <div class="checkbox">
                <label>
                    Do you have a positive citology result? <input asp-for="Citology" />
                </label>
            </div>
        </div>
    </div>
</div>
</div>
</div>
</fieldset>
@*</div>*@
@*<div class="form-group">
    <label asp-for="FKIDUsers" class="control-label"></label>
    <select asp-for="FKIDUsers" class ="form-control" asp-items="ViewBag.FKIDUsers"></select>
</div>*@

    <div class="text-center">
        <input type="submit" align="right" class="btn btn-info" value="Create" asp-action="Create" asp-
controller="DatasetModels" />
    </div>

</div>
</form>
</div>
</div>

<div>
    <a asp-action="Index" asp-controller="PredictionBiopsyModels">Back to List</a>
</div>

<hr />

```

Loader.cshtml представлення «DatasetModels»

```

@{
    ViewData["Title"] = "Loader";
}

@*<h2>Loader</h2>*@
@*<h3>@ViewData["Message"]</h3>*@

@*<script src="~/js/ProgressBar.js"></script>*@

<link rel="stylesheet" href="https://maxcdn.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css">
<script src="https://ajax.googleapis.com/ajax/libs/jquery/3.4.0/jquery.min.js"></script>
<script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.0/js/bootstrap.min.js"></script>

<div class="container">
    <div class="progress">
        <div id="dynamic" class="progress-bar progress-bar-striped progress-bar-animated" role="progressbar"
aria-valuenow="0" aria-valuemin="0" aria-valuemax="100" style="width: 0%">
            <span id="current-progress"></span>
        </div>
    </div>
</div>

<!-- jQuery Script -->

```



```

<script type="text/javascript">
  var i = 0;
  function makeProgress() {
    if (i < 100) {
      i = i + 1;
      $(".progress-bar").css("width", i + "%").text(i + " %");
    }
    // Wait for sometime before running this script again
    setTimeout("makeProgress()", 200);

    if (i == 99)
    {
      window.open("https://localhost:5001/PredictionBiopsyModels", "_self")
    }
  }
  makeProgress();
</script>

```

Details.cshtml представления «DatasetModels»

```

@model CCP_Project.Models.DatasetModel

@{
  ViewData["Title"] = "Details";
}

<h2>Details</h2>

<div>
  <h4>DatasetModel</h4>
  <hr />
  <dl class="dl-horizontal">
    <dt>
      @Html.DisplayNameFor(model => model.Age)
    </dt>
    <dd>
      @Html.DisplayFor(model => model.Age)
    </dd>
    <dt>
      @Html.DisplayNameFor(model => model.NumberOfSexualPartners)
    </dt>
    <dd>
      @Html.DisplayFor(model => model.NumberOfSexualPartners)
    </dd>
    <dt>
      @Html.DisplayNameFor(model => model.AgeOfFirstSexualIntercourse)
    </dt>
    <dd>
      @Html.DisplayFor(model => model.AgeOfFirstSexualIntercourse)
    </dd>
    <dt>
      @Html.DisplayNameFor(model => model.NumberOfPregnancies)
    </dt>
    <dd>
      @Html.DisplayFor(model => model.NumberOfPregnancies)
    </dd>
    <dt>
      @Html.DisplayNameFor(model => model.Smokes)
    </dt>

```

```

<dd>
  @Html.DisplayFor(model => model.Smokes)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.AgeSmokes)
</dt>
<dd>
  @Html.DisplayFor(model => model.AgeSmokes)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.PacksYearSmokes)
</dt>
<dd>
  @Html.DisplayFor(model => model.PacksYearSmokes)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.UsingHormonalContraceptives)
</dt>
<dd>
  @Html.DisplayFor(model => model.UsingHormonalContraceptives)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.AgeUsingHormonalContraceptives)
</dt>
<dd>
  @Html.DisplayFor(model => model.AgeUsingHormonalContraceptives)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.IUD)
</dt>
<dd>
  @Html.DisplayFor(model => model.IUD)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.AgeIUD)
</dt>
<dd>
  @Html.DisplayFor(model => model.AgeIUD)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.NumberOfSTDs)
</dt>
<dd>
  @Html.DisplayFor(model => model.NumberOfSTDs)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_Condylomatosis)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_Condylomatosis)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_CervicalCondylomatosis)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_CervicalCondylomatosis)

```

```

</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_VaginalCondylomatosis)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_VaginalCondylomatosis)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_VulvoPerinealCondylomatosis)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_VulvoPerinealCondylomatosis)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_Syphilis)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_Syphilis)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_PelvicInflammatoryDisease)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_PelvicInflammatoryDisease)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_GenitalHerpes)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_GenitalHerpes)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_MolluscumContagiosum)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_MolluscumContagiosum)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_AIDS)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_AIDS)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_HIV)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_HIV)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_HepatitisB)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_HepatitisB)
</dd>
<dt>
  @Html.DisplayNameFor(model => model.STDs_HPVP)
</dt>
<dd>
  @Html.DisplayFor(model => model.STDs_HPVP)
</dd>
<dt>

```

```

    @Html.DisplayNameFor(model => model.STDs_NumberOfDiagnosis)
</dt>
<dd>
    @Html.DisplayFor(model => model.STDs_NumberOfDiagnosis)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.STDs_TimeSinceFirstDiagnosis)
</dt>
<dd>
    @Html.DisplayFor(model => model.STDs_TimeSinceFirstDiagnosis)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.STDs_TimeSinceLastDiagnosis)
</dt>
<dd>
    @Html.DisplayFor(model => model.STDs_TimeSinceLastDiagnosis)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.Dx_Cancer)
</dt>
<dd>
    @Html.DisplayFor(model => model.Dx_Cancer)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.Dx_CIN)
</dt>
<dd>
    @Html.DisplayFor(model => model.Dx_CIN)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.Dx_HPVP)
</dt>
<dd>
    @Html.DisplayFor(model => model.Dx_HPVP)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.Dx)
</dt>
<dd>
    @Html.DisplayFor(model => model.Dx)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.HinselmannTest)
</dt>
<dd>
    @Html.DisplayFor(model => model.HinselmannTest)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.SchillerTest)
</dt>
<dd>
    @Html.DisplayFor(model => model.SchillerTest)
</dd>
<dt>
    @Html.DisplayNameFor(model => model.Citology)
</dt>
<dd>
    @Html.DisplayFor(model => model.Citology)
</dd>
@*<dt>
    @Html.DisplayNameFor(model => model.User)
</dt>

```

```

        <dd>
            @Html.DisplayFor(model => model.User.Id)
        </dd>*@
    </dl>
</div>

<div>
    <a asp-action="Index" asp-controller="PredictionBiopsyModels">Back to List</a>
</div>

```

Index.cshtml представления «PredictionBiopsyModels»

```

@model IEnumerable<CCP_Project.Models.PredictionBiopsyModel>

@{
    ViewData["Title"] = "Index";
}

<h2 align="center">History</h2>

@*<p>
    <a asp-action="Create">Create New</a>
</p>*@
<table class="table">
    <thead>
        <tr>
            <th>
                Result
                @*@Html.DisplayNameFor(model => model.PredictionBiopsy)*@
            </th>
            <th>
                @Html.DisplayNameFor(model => model.DateTime)
            </th>
            <th>
                @*@Html.DisplayNameFor(model => model.User)*@
                Actions
            </th>
        </tr>
    </thead>
    <tbody>
    @foreach (var item in Model) {
        <tr>
            <td>
                @Html.DisplayFor(modelItem => item.PredictionBiopsy)
            </td>
            <td>
                @Html.DisplayFor(modelItem => item.DateTime)
            </td>
            @*<td>
                @Html.DisplayFor(modelItem => item.User.Id)
            </td>*@
            <td>
                @*<a asp-action="Edit" asp-route-id="@item.IDPredictionBiopsy">Edit</a> |*@
                <a
                    asp-controller="DatasetModels"
                    asp-action="Details"
                    id="@item.IDPredictionBiopsy">Details</a> |
                <a asp-action="Delete" asp-route-id="@item.IDPredictionBiopsy">Delete</a>
            </td>
        </tr>
    }
    </tbody>
</table>

```

```

}
</tbody>
</table>

```

Delete.cshtml представлення «PredictionBiopsyModels»

```

@model CCP_Project.Models.PredictionBiopsyModel

@{
    ViewData["Title"] = "Delete";
}

<h2>Delete</h2>

<h3>Are you sure you want to delete this?</h3>
<div>
    <h4>PredictionBiopsyModel</h4>
    <hr />
    <dl class="dl-horizontal">
        <dt>
            @Html.DisplayNameFor(model => model.PredictionBiopsy)
        </dt>
        <dd>
            @Html.DisplayFor(model => model.PredictionBiopsy)
        </dd>
        <dt>
            @Html.DisplayNameFor(model => model.DateTime)
        </dt>
        <dd>
            @Html.DisplayFor(model => model.DateTime)
        </dd>
        @*<dt>
    </dl>

    <form asp-action="Delete">
        <input type="hidden" asp-for="IDPredictionBiopsy" />
        <input type="submit" value="Delete" class="btn btn-default" /> |
        <a asp-action="Index">Back to List</a>
    </form>
</div>

```

_Layout.cshtml представлення «Shared»

```

<!DOCTYPE html>
<html>
<head>
    <meta charset="utf-8" />
    <meta name="viewport" content="width=device-width, initial-scale=1.0" />
    <title>@ ViewData["Title"] - CCP</title>

    <script src="https://cdnjs.cloudflare.com/ajax/libs/twitter-bootstrap/4.0.0-alpha/js/umd/util.js"></script>

    <environment include="Development">
        <link rel="stylesheet" href="~/lib/bootstrap/dist/css/bootstrap.css" />
    </environment>

```

```

<environment exclude="Development">
  <link rel="stylesheet" href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css"
    asp-fallback-href="~/lib/bootstrap/dist/css/bootstrap.min.css"
    asp-fallback-test-class="sr-only"      asp-fallback-test-property="position"      asp-fallback-test-
value="absolute"
    crossorigin="anonymous"
    integrity="sha384-
ggOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQUOhcWr7x9JvoRxT2MZw1T" />
</environment>
<link rel="stylesheet" href="~/css/site.css" />
</head>
<body>
  <header>
    <nav class="navbar navbar-expand-sm navbar-toggleable-sm navbar-light bg-white border-bottom box-
shadow mb-3">
      <div class="container">
        <a class="navbar-brand" asp-area="" asp-controller="Home" asp-action="Index">Cervical cancer
prediction</a>
        <button class="navbar-toggler" type="button" data-toggle="collapse" data-target=".navbar-collapse"
aria-controls="navbarSupportedContent"
          aria-expanded="false" aria-label="Toggle navigation">
          <span class="navbar-toggler-icon"></span>
        </button>
        <div class="navbar-collapse collapse d-sm-inline-flex flex-sm-row-reverse">
          <partial name="_LoginPartial" />
          <ul class="navbar-nav flex-grow-1">

            @if (User.Identity.IsAuthenticated)
            {
              <li class="nav-item">
                <a class="nav-link text-dark" asp-area="" asp-controller="Home" asp-
action="Index">Home</a>
              </li>
              <li class="nav-item">
                <a class="nav-link text-dark" asp-area="" asp-controller="DatasetModels" asp-
action="Create">Prediction</a>
              </li>
              <li class="nav-item">
                <a class="nav-link text-dark" asp-area="" asp-controller="PredictionBiopsyModels" asp-
action="Index">History</a>
              </li>
              @*<li class="nav-item">
                <a class="nav-link text-dark" asp-area="" asp-controller="Home" asp-action="Privacy">Privacy</a>
              </li>*@
              <li class="nav-item">
                <a class="nav-link text-dark" asp-area="" asp-controller="Home" asp-
action="Contact">Contact</a>
              </li>
            }
          </ul>
        </div>
      </div>
    </nav>
  </header>
  <div class="container">
    <partial name="_CookieConsentPartial" />
    <main role="main" class="pb-3">
      @RenderBody()
    </main>
  </div>

```

```

<footer class="border-top footer text-muted">
  <div class="container">
    &copy; @DateTime.Now.Year - Created by Anna Eliseeva - <a asp-area="" asp-controller="Home"
asp-action="Privacy">Privacy</a>
  </div>
</footer>

<environment include="Development">
  <script src="~/lib/jquery/dist/jquery.js"></script>
  <script src="~/lib/bootstrap/dist/js/bootstrap.bundle.js"></script>

</environment>
<environment exclude="Development">
  <script src="https://cdn.jsdelivr.net/npm/jquery/3.3.1/jquery.min.js"
asp-fallback-src="~/lib/jquery/dist/jquery.min.js"
asp-fallback-test="window.jQuery"
crossorigin="anonymous"
integrity="sha256-FgpCb/KJQlLNfOu91ta32o/NMZxltwRo8QtmkMRdAu8=">
  </script>
  <script src="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/js/bootstrap.bundle.min.js"
asp-fallback-src="~/lib/bootstrap/dist/js/bootstrap.bundle.min.js"
asp-fallback-test="window.jQuery && window.jQuery.fn && window.jQuery.fn.modal"
crossorigin="anonymous"
integrity="sha384-
xrRywqdh3PHs8keKZN+8zcc5TX0GRTLcCmivcbNjWm2rs5C8PRhcEn3czEjhAO9o">
  </script>
</environment>
<script src="~/js/site.js" asp-append-version="true"></script>

  @RenderSection("Scripts", required: false)
</body>
</html>

```

IfCheckboxIsTrue.js

```

function IfCheckboxIsTrue1() {

var checkbox = document.getElementById("trigger1")
var hidden3 = document.getElementById("hidden_field3")
var hidden4 = document.getElementById("hidden_field4")

if (checkbox.checked == true) {
  hidden3.disabled = false;
  hidden4.disabled = false;
} else {
  hidden3.disabled = true;
  hidden4.disabled = true;
}
};

function IfCheckboxIsTrue2() {

var checkbox2 = document.getElementById("trigger2")
var hidden5 = document.getElementById("hidden_field5")

if (checkbox2.checked == true) {
  hidden5.disabled = false;
} else {
  hidden5.disabled = true;
}
};

```



```

    }
};

function IfCheckboxIsTrue3() {

    var checkbox3 = document.getElementById("trigger3")
    var hidden6 = document.getElementById("hidden_field6")

    if (checkbox3.checked == true) {
        hidden6.disabled = false;
    } else {
        hidden6.disabled = true;
    }
};

function IfCheckboxIsTrue4() {

    var checkbox4 = document.getElementById("trigger4")
    var hidden7 = document.getElementById("hidden_field7")
    var hidden8 = document.getElementById("hidden_field8")
    var hidden9 = document.getElementById("hidden_field9")
    var hidden10 = document.getElementById("hidden_field10")
    var hidden11 = document.getElementById("hidden_field11")
    var hidden12 = document.getElementById("hidden_field12")
    var hidden13 = document.getElementById("hidden_field13")
    var hidden14 = document.getElementById("hidden_field14")
    var hidden15 = document.getElementById("hidden_field15")
    var hidden16 = document.getElementById("hidden_field16")
    var hidden17 = document.getElementById("hidden_field17")
    var hidden18 = document.getElementById("hidden_field18")
    var hidden19 = document.getElementById("hidden_field19")
    var hidden20 = document.getElementById("hidden_field20")
    var hidden21 = document.getElementById("hidden_field21")
    var hidden22 = document.getElementById("hidden_field22")

    if (checkbox4.checked == true) {
        hidden7.disabled = false;
        hidden8.disabled = false;
        hidden9.disabled = false;
        hidden10.disabled = false;
        hidden11.disabled = false;
        hidden12.disabled = false;
        hidden13.disabled = false;
        hidden14.disabled = false;
        hidden15.disabled = false;
        hidden16.disabled = false;
        hidden17.disabled = false;
        hidden18.disabled = false;
        hidden19.disabled = false;
        hidden20.disabled = false;
        hidden21.disabled = false;
        hidden22.disabled = false;
    } else {
        hidden7.disabled = true;
        hidden8.disabled = true;
        hidden9.disabled = true;
        hidden10.disabled = true;
        hidden11.disabled = true;
        hidden12.disabled = true;
        hidden13.disabled = true;
        hidden14.disabled = true;
        hidden15.disabled = true;
    }
};

```

```

        hidden16.disabled = true;
        hidden17.disabled = true;
        hidden18.disabled = true;
        hidden19.disabled = true;
        hidden20.disabled = true;
        hidden21.disabled = true;
        hidden22.disabled = true;
    }
};

```

NumberOfSexualPartners.js

```

function Bar() {
    var sexpart = document.getElementById("sexpart").value;
    var hidden1 = document.getElementById("hidden_field1")
    var hidden2 = document.getElementById("hidden_field2")

    if (sexpart > 0) {
        hidden1.disabled = false;
        hidden2.disabled = false;
    } else {
        hidden1.disabled = true;
        hidden2.disabled = true;
    }
}

```

Cervical_cancer_prediction_new.py

```

#!/usr/bin/env python
# coding: utf-8

import numpy as np
import pandas as pd
from sqlalchemy import create_engine
import psycog2
# Розробка моделі
# TensorFlow - це бібліотека з відкритим вихідним кодом для чисельних обчислень,
# в яких використовуються графіки потоків даних.
import tensorflow as tf
# Експорт бібліотеки Keras для роботи з нейронною мережею
import keras
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Dropout

# З'єднання з базою даних PostgreSQL
engine = create_engine('postgres://postgres:root@localhost:5432/ccp')

# Тренувальна вибірка
data = pd.read_sql_query('SELECT "Age", "NumberOfSexualPartners", "AgeOfFirstSexualIntercourse",
"NumberOfPregnancies", ' "AgeSmokes", "PacksYearSmokes", "AgeUsingHormonalContraceptives",
"AgeIUD", ' "NumberOfSTDs", "STDs_Condylomatosis", "STDs_CervicalCondylomatosis",
"STDs_VaginalCondylomatosis", ' "STDs_VulvoPerinealCondylomatosis", "STDs_Syphilis",
"STDs_PelvicInflammatoryDisease", ' "STDs_GenitalHerpes", "STDs_MolluscumContagiosum",
"STDs_AIDS", "STDs_HIV", "STDs_HepatitisB", ' "STDs_HPV", "STDs_NumberOfDiagnosis",

```

```
"STDs_TimeSinceFirstDiagnosis", "STDs_TimeSinceLastDiagnosis", ' "Smokes",
"UsingHormonalContraceptives", "IUD", ' "STDs", "Dx_Cancer", "Dx_CIN", "Dx_HP", ' "Dx",
"HinselmannTest", "Citology", "SchillerTest" ' FROM "DatasetModels" ORDER BY "IDDataset" DESC LIMIT
1', con=engine)
```

```
# Тренувальна вибірка
user = pd.read_sql_query('SELECT "FKIDUsers" FROM "DatasetModels" ORDER BY "IDDataset" DESC
LIMIT 1', con=engine)
```

```
increment = pd.read_sql_query('SELECT "IDDataset" 'FROM "DatasetModels"
ORDER BY "IDDataset" DESC LIMIT 1', con=engine)
```

```
# Перетворимо усі дані в числовий тип даних
data = data.applymap(lambda x: 1 if x == True else x)
data = data.applymap(lambda x: 0 if x == False else x)
```

```
data['Smokes_0'] = np.where(data['Smokes']==0, 1, 0)
data['UsingHormonalContraceptives_0'] = np.where(data['UsingHormonalContraceptives']==0, 1, 0)
data['IUD_0'] = np.where(data['IUD']==0, 1, 0)
data['STDs_0'] = np.where(data['STDs']==0, 1, 0)
data['Dx_Cancer_0'] = np.where(data['Dx_Cancer']==0, 1, 0)
data['Dx_CIN_0'] = np.where(data['Dx_CIN']==0, 1, 0)
data['Dx_HP'] = np.where(data['Dx_HP']==0, 1, 0)
data['Dx_0'] = np.where(data['Dx']==0, 1, 0)
data['HinselmannTest_0'] = np.where(data['HinselmannTest']==0, 1, 0)
data['Citology_0'] = np.where(data['Citology']==0, 1, 0)
data['SchillerTest_0'] = np.where(data['SchillerTest']==0, 1, 0)
```

```
data['Smokes_1'] = np.where(data['Smokes']==1, 1, 0)
data['UsingHormonalContraceptives_1'] = np.where(data['UsingHormonalContraceptives']==1, 1, 0)
data['IUD_1'] = np.where(data['IUD']==1, 1, 0)
data['STDs_1'] = np.where(data['STDs']==1, 1, 0)
data['Dx_Cancer_1'] = np.where(data['Dx_Cancer']==1, 1, 0)
data['Dx_CIN_1'] = np.where(data['Dx_CIN']==1, 1, 0)
data['Dx_HP'] = np.where(data['Dx_HP']==1, 1, 0)
data['Dx_1'] = np.where(data['Dx']==1, 1, 0)
data['HinselmannTest_1'] = np.where(data['HinselmannTest']==1, 1, 0)
data['Citology_1'] = np.where(data['Citology']==1, 1, 0)
data['SchillerTest_1'] = np.where(data['SchillerTest']==1, 1, 0)
```

```
del data['Smokes']
del data['UsingHormonalContraceptives']
del data['IUD']
del data['STDs']
del data['Dx_Cancer']
del data['Dx_CIN']
del data['Dx_HP']
del data['Dx']
del data['HinselmannTest']
del data['Citology']
del data['SchillerTest']
```

```
from keras.models import model_from_json
```

```
# load json and create model
json_file = open('Model.json', 'r')
loaded_model_json = json_file.read()
json_file.close()
loaded_model = model_from_json(loaded_model_json)
# load weights into new model
loaded_model.load_weights("Model.h5")
```

```

# evaluate loaded model on test data
loaded_model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
biopsy = loaded_model.predict_classes(data)

intermediate1 = biopsy[0]
result = intermediate1[0]

from datetime import datetime

current_datetime = datetime.now().strftime('%Y-%m-%d %H:%M:%S')

userID = user['FKIDUsers'].iloc[0]

prediction = pd.DataFrame({"IDPredictionBiopsy":[increment.iloc[0].iloc[0]],
                          "PredictionBiopsy":[result],
                          "DateTime":[current_datetime],
                          "FKIDUsers":[userID]})

prediction.to_sql('PredictionBiopsyModels', con=engine, if_exists = 'append', index=False)

```

Model.json

```

{"class_name": "Sequential", "config": {"name": "sequential_1", "layers": [{"class_name": "Dense", "config": {"name": "dense_1", "trainable": true, "batch_input_shape": [null, 46], "dtype": "float32", "units": 100, "activation": "relu", "use_bias": true, "kernel_initializer": {"class_name": "RandomUniform", "config": {"minval": -0.05, "maxval": 0.05, "seed": null}}, "bias_initializer": {"class_name": "Zeros", "config": {}}, "kernel_regularizer": null, "bias_regularizer": null, "activity_regularizer": null, "kernel_constraint": null, "bias_constraint": null}, {"class_name": "Dense", "config": {"name": "dense_2", "trainable": true, "units": 100, "activation": "relu", "use_bias": true, "kernel_initializer": {"class_name": "RandomUniform", "config": {"minval": -0.05, "maxval": 0.05, "seed": null}}, "bias_initializer": {"class_name": "Zeros", "config": {}}, "kernel_regularizer": null, "bias_regularizer": null, "activity_regularizer": null, "kernel_constraint": null, "bias_constraint": null}, {"class_name": "Dense", "config": {"name": "dense_3", "trainable": true, "units": 100, "activation": "relu", "use_bias": true, "kernel_initializer": {"class_name": "RandomUniform", "config": {"minval": -0.05, "maxval": 0.05, "seed": null}}, "bias_initializer": {"class_name": "Zeros", "config": {}}, "kernel_regularizer": null, "bias_regularizer": null, "activity_regularizer": null, "kernel_constraint": null, "bias_constraint": null}, {"class_name": "Dense", "config": {"name": "dense_4", "trainable": true, "units": 100, "activation": "relu", "use_bias": true, "kernel_initializer": {"class_name": "RandomUniform", "config": {"minval": -0.05, "maxval": 0.05, "seed": null}}, "bias_initializer": {"class_name": "Zeros", "config": {}}, "kernel_regularizer": null, "bias_regularizer": null, "activity_regularizer": null, "kernel_constraint": null, "bias_constraint": null}, {"class_name": "Dense", "config": {"name": "dense_5", "trainable": true, "units": 1, "activation": "sigmoid", "use_bias": true, "kernel_initializer": {"class_name": "RandomUniform", "config": {"minval": -0.05, "maxval": 0.05, "seed": null}}, "bias_initializer": {"class_name": "Zeros", "config": {}}, "kernel_regularizer": null, "bias_regularizer": null, "activity_regularizer": null, "kernel_constraint": null, "bias_constraint": null}}], "keras_version": "2.2.4", "backend": "tensorflow"}

```