

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет кораблебудування імені адмірала Макарова

Навчально-науковий інститут комп'ютерних наук та управління проектами

Кафедра програмного забезпечення автоматизованих систем

Спеціальність 121 «Інженерія програмного забезпечення»

Освітня програма «Інженерія програмного забезпечення»

«Допущений до захисту»

Завідувач кафедри

«___» _____ 2021 р.

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття ступеня вищої освіти «магістр»

на тему: Нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, та розробка програми для її реалізації

Виконав: студент 6151м групи

Приходько А.С.

(підпис, ПІБ)

Керівник роботи:

доцент кафедри ПЗАС, к.т.н,

доцент Макарова Л.М.

(посада, науковий ступень вчене звання)

(підпис, ПІБ)

Миколаїв – 2021 р.

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет кораблебудування імені адмірала Макарова

Навчально-науковий інститут комп'ютерних наук та управління проектами
 Кафедра програмного забезпечення автоматизованих систем
 Спеціальність 121 «Інженерія програмного забезпечення»
 Освітня програма «Інженерія програмного забезпечення»

«ЗАТВЕРДЖУЮ»

Гарант освітньої програми

_____ проф. С.Б.Приходько

(підпис)

«___» _____ 2021 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ
на здобуття ступеня вищої освіти «магістр»

Студенту Приходьку Андрію Сергійовичу _____
 (Прізвище, ім'я, по батькові)

1. Тема роботи: Нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, та розробка програми для її реалізації

Керівник роботи Макарова Л.М

Затверджені наказом ректора № 1239уч від «13» _____ 10 _____ 2021 р.

2. Термін подання роботи: 06.12.2021 р.

3. Вихідні дані по роботі: _____

4. Перелік питань, що належать до розробки (найменування розділів) _____

Вступ (Актуальність теми. Зв'язок роботи з науковими програмами, планами, темами. Мета і завдання дослідження. Об'єкт дослідження. Предмет дослідження. Методи дослідження. Наукова новизна одержаних результатів. Практичне значення одержаних результатів. Особистий внесок здобувача. Апробація результатів досліджень. Публікації.); Огляд літератури та обґрунтування необхідності проведення досліджень за обраною темою; Викладення результатів власних досліджень з висвітленням того нового, що пропонується; Проект програмного забезпечення; Організаційно-економічний розділ; Розділи з охорони праці та охорони навколишнього середовища; Висновки; Список використаних джерел; Додатки (технічне завдання, текст програми, опис програми, інструкція користувача, програма і методика випробувань програмного забезпечення)

5. Перелік презентаційних матеріалів _____

6. Консультанти розділів роботи

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

7. Дата видачі завдання 13.10.2021 р. _____

КАЛЕНДАРНИЙ ПЛАН

Номер	Назва етапів роботи	Терміни виконання	Примітка
1.	Підготовка розділу Вступ	18.10.2021	НС*
2.	Підготовка розділу з огляду літератури та обґрунтування необхідності проведення досліджень за обраною темою	19.10.2021	НС*
3.	Підготовка розділу (ів) з результатів власних досліджень	22.10.2021	НС*
4.	Підготовка розділу з проекту програмного забезпечення	16.11.2021	
5.	Підготовка організаційно-економічного розділу	19.11.2021	
6.	Підготовка розділу з охорони праці	22.11.2021	
7.	Підготовка розділу з охорони навколишнього середовища	24.11.2021	
8.	Підготовка розділу Висновки	25.11.2021	
9.	Оформлення списку використаних джерел та додатків	29.11.2021	
10.	Подання на кафедру ПЗАС тексту остаточного варіанту роботи, підписаного її керівником, у роздрукованому та електронному форматі разом із заявами щодо самостійності виконання роботи та ідентичності друкованої та електронної версії роботи (Додатки 1 і 2 «Порядку здійснення заходів з перевірки робіт на наявність текстових збігів/ідентичності/схожості із використанням програмно-технічних засобів», який введений в дію наказом ректора НУК за №20 від 20.01.2020 р.)	06.12.2021	не пізніше, ніж за 14 діб до захисту (згідно п.4.1 зазначеного Порядку
11.	Підготовка презентації та доповіді	09.12.2021	
12.	Попередній захист роботи на засіданні кафедри ПЗАС	10.12.2021	
13.	Подання на кафедру ПЗАС електронних версії наступних документів у форматі pdf: кваліфікаційної роботи; файлу-опису кваліфікаційної роботи (згідно Додатку до наказу ректора НУК за №287-уч від 19.05.2020 р.); презентації доповіді	20.12.2021	

* - за результатами наукового стажування (НС), яке було з 01.09.2021 до 10.10.2021 р.

Студент

_____ (підпис)

_____ (ПІБ)

Керівник роботи

_____ (підпис)

_____ (ПІБ)

РЕФЕРАТ

Приходько Андрій Сергійович

«Нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, та розробка програми для її реалізації»

Кваліфікаційна робота на здобуття освітнього рівня магістра зі спеціальності 121 – «Інженерія програмного забезпечення». Національний університет кораблебудування імені адмірала Макарова. Миколаїв, 2021 р.

Обсяг роботи: 103 стор., 6 табл., 13 рис., 32 використаних джерела, 5 додатків.

Актуальність теми роботи: полягає в тому, що оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, є актуальною задачею, виходячи з популярності та поширеності цього фреймворку.

Метою роботи є підвищення достовірності оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, за допомогою нелінійної регресійної моделі.

Об'єктом дослідження є процес оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Предметом дослідження є нелінійні регресійні моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Методи дослідження: методи теорії ймовірностей та математичної статистики, математичного програмування, регресійного аналізу.

Наукова новизна одержаних результатів. Удосконалено нелінійну регресійну модель для оцінювання розміру web застосунків, що створюються

з використанням фреймворку CakePhp, на основі десяткового логарифму. Це дає змогу здійснити підвищення достовірності відповідного оцінювання у порівнянні з лінійною регресійною моделлю.

Практичне значення одержаних результатів. Розроблено програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp. Програму створено на мові програмування Python. Програму, що розроблено, можна використовувати для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на ранніх стадіях їх проектування.

Апробація результатів досліджень. Результати досліджень, викладені у роботі, були оприлюднені на чотирьох конференціях

Публікації. Результати роботи викладені у 3 публікаціях.

Ключові слова: оцінка кількості строк коду програмного забезпечення; web-застосунок; нелінійна регресійна модель, CakePhp.

ABSTRACT

Prykhodko Andrii Sergiyovich

«A nonlinear regression model for estimating the size of web applications created using CakePhp framework and developing the software for its implementation»

The qualification work in obtaining a master's degree in specialty 121 – "Software Engineering". Admiral Makarov National University of Shipbuilding. Mykolayiv, 2021.

Work contains 103 pages, 6 tables, 13 figures, 32 references, 5 appendices.

Relevance of the theme: is that estimating the size of web applications developed using the CakePhp framework is an urgent task based on the popularity and prevalence of this framework.

The object of the study is the process of estimating the size of web applications created using the CakePhp framework.

The subject of the study is nonlinear regression models for estimating the size of web applications created using the CakePhp framework.

The aim is to increase the confidence of estimating the size of web applications created using the CakePhp framework using a nonlinear regression model.

Research methods: methods of probability theory and mathematical statistics, mathematical programming, regression analysis.

Scientific novelty of the obtained results: The nonlinear regression model for estimating the size of web applications created using the CakePhp framework, based on the decimal logarithm, has been improved. This makes it possible to increase the reliability of the corresponding estimate compared to the linear regression model.

The practical significance of the results obtained: A program for estimating the size of web applications created using the CakePhp framework has been developed. The program is created in the Python programming language. The software can be used to estimate the size of web applications created using the CakePhp framework in the early stages of their design.

Approbation of research results: The results of the research presented in the paper were published at four conferences.

Publications: The results of the work are presented in 3 publications.

Keywords: LOC estimation; Web app; nonlinear regression model, CakePhp.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	9
ВСТУП.....	10
1 АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ ТА ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ ПРОВЕДЕННЯ ДОСЛІДЖЕНЬ ЗА ОБРАНОЮ ТЕМОЮ.....	14
1.1 АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ	14
1.2 ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ ПРОВЕДЕННЯ ДОСЛІДЖЕНЬ ЗА ОБРАНОЮ ТЕМОЮ	19
1.3 ВИСНОВКИ ЗА РОЗДІЛОМ 1	21
2 НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР	23
2.1 ПЕРЕВІРКА ДАНИХ ДЛЯ ПОБУДОВИ НЕЛІНІЙНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР, НА НАЯВНІСТЬ МУЛЬТИКОЛІНЕАРНОСТІ ТА ВИКИДІВ	23
2.2 ПОБУДОВА НЕЛІНІЙНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР	32
2.3 ВИСНОВКИ ЗА РОЗДІЛОМ 2	37
3 РОЗРОБКА ПРОЕКТУ ПРОГРАМИ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР	39
3.1 ПОСТАНОВКА ЗАДАЧІ НА РОЗРОБКУ ПРОЕКТУ ПРОГРАМИ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР..	39
3.2 РОЗРОБКА ЕСКІЗНОГО ПРОЕКТУ ПРОГРАМИ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР	42

3.3 РОЗРОБКА ТЕХНІЧНОГО ПРОЕКТУ ПРОГРАМИ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР	47
3.4 РОЗРОБКА РОБОЧОГО ПРОЕКТУ ТА ТЕСТУВАННЯ ПРОГРАМИ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ САКЕРНР..	54
3.5 ВИСНОВКИ ЗА РОЗДІЛОМ 3	58
4 РОЗРАХУНОК ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ ВІД РОЗРОБКИ І ВПРОВАДЖЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ.....	59
4.1 РОЗРАХУНОК ВИТРАТ НА СТВОРЕННЯ Й ЕКСПЛУАТАЦІЮ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ	59
5 ОХОРОНА ПРАЦІ В ОФІСІ ФІРМИ	64
5.1 ВСТУПНА ЧАСТИНА	64
5.2 АНАЛІЗ ШКІДЛИВИХ ТА НЕБЕЗПЕЧНИХ ФАКТОРІВ В ОФІСІ ФІРМИ	66
5.3 РОЗРАХУНОК РІВНЯ ОСВІТЛЕННЯ В ОФІСІ ФІРМИ.....	72
5.4 РОЗРОБКА ЗАХОДІВ ЩОДО ЗМЕНШЕННЯ ВПЛИВУ ШКІДЛИВИХ ТА НЕБЕЗПЕЧНИХ ФАКТОРІВ.....	75
6 ОХОРОНА НАВКОЛИШНЬОГО СЕРЕДОВИЩА	77
6.1 ЗАБРУДНЕННЯ НАВКОЛИШНЬОГО СЕРЕДОВИЩА В ПРОЦЕСІ ВИКОРИСТАННЯ КОМП'ЮТЕРНОЇ ТЕХНІКИ.....	78
6.2 ВИКОРИСТАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ВИРІШЕННЯ ЕКОЛОГІЧНИХ ПРОБЛЕМ	79
ВИСНОВКИ.....	83
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	85
ДОДАТОК А – ТЕХНІЧНЕ ЗАВДАННЯ	90
ДОДАТОК Б – ТЕКСТ ПРОГРАМИ	94
ДОДАТОК В – ОПИС ПРОГРАМИ.....	97
ДОДАТОК Г – ІНСТРУКЦІЯ КОРИСТУВАЧА	98
ДОДАТОК Д - ПРОГРАМА ТА МЕТОДИКА ВИПРОБОВУВАНЬ	100

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ПЗ – програмне забезпечення;

Classes – the number of classes;

COCOMO – the COConstructive COst MOdel;

DIT – the Depth of Inheritance Tree metric provides for each class a measure of the inheritance levels from the object hierarchy top. The minimum value of DIT for a class is 1. A DIT value of 0 indicates a root while a value of 2 and 3 indicates a higher degree of reuse. If there is a majority of DIT values below 2, it may represent poor exploitation of the advantages of OO design and inheritance. It is recommended a maximum DIT value of 5 since deeper trees constitute greater design complexity as more methods and classes are involved (<https://www.cachequality.com/docs/metrics/depth-inheritance-tree>);

KLOC – the thousand lines of code;

LB – lower bound;

LOC – the lines of code;

MD – the Mahalanobis distance;

Methods – the number of methods;

MMRE – a mean magnitude of relative error;

MRE – a magnitude of relative error;

PHP – hypertext preprocessor;

PRED – percentage of prediction;

SMD – the squared Mahalanobis distance;

UB – upper bound;

VIF – variance influence factor;

ВСТУП

Актуальність теми. Оцінювання розміру web застосунків як і іншого програмного забезпечення (ПЗ) є важливою задачею на ранній стадії розробки програмного забезпечення, оскільки ця інформація використовується для прогнозування трудомісткості створення ПЗ за допомогою математичних моделей, наприклад, СОСОМО II. Зараз багато web застосунків створюються за допомогою різних фреймворків, які прискорюють розробку ПЗ. Фреймворки визначають структуру і архітектуру майбутнього ПЗ. Найвідомішими з них є Symfony, Laravel, CakePhp, Yii.

В якості PHP-фреймворку для дослідження було обрано CakePhp через декілька причин:

1. Цей фреймворк є одним з найбільш використовуваних у розробці web застосунків.
2. Для цього фреймворку не було створено регресійної моделі для оцінювання кількості строк коду web застосунків, що створюються з його використанням.

На теперішній час відомі як лінійне регресійне рівняння [1], так і нелінійна регресійна модель [2] для оцінювання кількості строк коду інформаційних PHP-систем з відкритим кодом в залежності від трьох метрик концептуальної моделі даних у вигляді діаграми класів: загальна кількість класів, загальна кількість зв'язків та середня кількість атрибутів на клас. Але розмір web застосунків, що створюються з використанням різних фреймворків, може залежати від інших метрик [3, 4]. На сьогоднішній день відома модель для веб-застосунків, що розробляються з використанням фреймворку Laravel [3]. Також не завжди можливо використати лінійну регресійну модель. Існують чотири основні припущення, які виправдовують використання моделей лінійної регресії, одним з яких є нормальність

розподілу помилок. Але це припущення справедливе лише в окремих випадках. Це призводить до необхідності використання нелінійних регресійних моделей та рівнянь. Усе це в цілому потребує побудови нелінійної моделі для оцінювання кількості строк коду web застосунків, що створюються з використанням PHP-фреймворку CakePhp.

Зазвичай у якості нормалізуючого перетворення використовують десятковий логарифм. Наприклад, такі відомі моделі як COCOMO, ISBSG, COCOMO II були створені саме із використанням нормалізуючого перетворення у вигляді десяткового логарифму. Тому удосконалювати регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, ми будемо саме на основі нормалізуючого перетворення у вигляді десяткового логарифму.

Мета і завдання дослідження. Метою роботи є підвищення достовірності оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, за допомогою нелінійної регресійної моделі.

Для досягнення поставленої мети потрібно вирішити такі завдання.

- Проаналізувати існуючі регресійні моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp;
- Зібрати набір даних, за якими буде будуватися нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, та перевірити ці дані на наявність багатовимірних викидів і мультиколінеарності;
- Удосконалити регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на основі нормалізуючого перетворення;

- На підставі запропонованої нелінійної регресійної моделі розробити програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Об'єктом дослідження є процес оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Предметом дослідження є нелінійні регресійні моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Методи дослідження. У вирішенні поставлених завдань використано методи теорії ймовірності, математичної статистики, лінійного і нелінійного регресійного аналізу.

Наукова новизна одержаних результатів. Удосконалено нелінійну регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на основі десяткового логарифму. Це дає змогу здійснити підвищення достовірності відповідного оцінювання у порівнянні з лінійною регресійною моделлю.

Практичне значення одержаних результатів. Розроблено програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp. Програму створено на мові програмування Python. Програму, що розроблено, можна використовувати для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на ранніх стадіях їх проектування.

Особистий внесок здобувача. У тезах конференції [5-7], що опубліковані у співавторстві, автору належать нелінійні регресійні моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на основі нормалізуючого перетворення у вигляді десяткового логарифму.

Апробація результатів досліджень. Результати досліджень, викладені у роботі, були оприлюднені на чотирьох конференціях: Вісімнадцятій всеукраїнській конференції студентів і молодих науковців (Одеса, 23 квітня 2021 року); П'ятій науково-практичній інтернет-конференції за тематикою «Проблеми моделювання та розроблення інформаційних систем» (Дрогобич, 23 квітня 2021 року); IEEE 16th International Conference on Computer Science and Information Technologies (Львів, 22-25 вересня 2021 року); II науково-практичній інтернет-конференції «Інформаційні технології, моделі, алгоритми, системи» (Миколаїв, 26-28 жовтня 2021 року).

Публікації. Результати роботи викладені у 3 публікаціях.

1 АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ ТА ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ ПРОВЕДЕННЯ ДОСЛІДЖЕНЬ ЗА ОБРАНОЮ ТЕМОЮ

1.1 Аналіз існуючих методів та математичних моделей для оцінювання розміру web застосунків

При оцінюванні трудомісткості розробки різноманітного ПЗ, в тому числі і web застосунків, часто використовуються методи нелінійного регресійного аналізу, нелінійні регресійні моделі та рівняння, зокрема СОСОМО II [8, 9], основним фактором в яких є розмір ПЗ. Але, як правило, розмір програми на початковому етапі її розробки не відомий, що потребує його оцінювання. В [1, 10] для оцінювання кількості строк коду інформаційних РНР-систем з відкритим кодом Hee Beng Kuan Tan, Yuan Zhao та Hongyu Zhang запропонували лінійне регресійне рівняння в залежності від трьох метрик концептуальної моделі даних у вигляді діаграми класів. Це рівняння було побудоване на основі методів множинного лінійного регресійного аналізу. Але, як відомо, при побудові лінійних регресійних моделей необхідно виконання певних умов, зокрема залишки (residuals) або залежна змінна повинні бути розподілені за нормальним законом, що має місце лише в поодиноких випадках. А це веде до необхідності побудови нелінійних регресійних моделей для оцінювання кількості строк ПЗ та застосування відповідних методів множинного нелінійного регресійного аналізу.

Нормалізуючі перетворення часто є добрим способом побудови моделей, рівнянь, довірчих інтервалів та інтервалів прогнозування нелінійних регресій [11-21]. В роботах [20-24] запропоновано рівняння нелінійної регресії для оцінювання розміру ПЗ інформаційних системи, що розробляються на базі мови програмування РНР. Запропоноване рівняння побудовано за допомогою множинного нелінійного регресійного аналізу із застосуванням чотиривимірного перетворення Джонсона сім'ї S_B на основі таких же метрик діаграми класів, що і в [1, 10]: загальна кількість класів, загальна кількість зв'язків та середня кількість атрибутів на клас. Але для РНР-застосунків з відкритим кодом, що не є інформаційними системами, наприклад, таких як різноманітні фреймворки, включаючи фреймворк CakePhp, конвертори, НТТР-клієнти та інші, регресійні моделі можуть залежати в тому числі від інших метрик. Наприклад, існує регресійна модель для веб-застосунків, що розроблюються з використанням фреймворку Laravel, яка залежить від інших метрик [3].

Зазвичай при побудові нелінійних регресійних моделей у якості нормалізуючого перетворення використовують десятковий логарифм. Наприклад, такі відомі моделі як COCOMO, ISBSG, COCOMO II були створені саме із використанням нормалізуючого перетворення у вигляді десятичного логарифму. Тому удосконалювати регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, ми будемо саме на основі нормалізуючого перетворення у вигляді десятичного логарифму.

Існуюче ПЗ, яке застосовується для оцінювання залежної змінної нелінійної регресії (наприклад, LabView, MatLab, Scilab), не має у своєму складі потрібних функцій, які дозволяли виконувати відповідне оцінювання. Тому виникає потреба у створенні оригінальної програми для оцінювання

розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Якщо випадкові величини Z_1 і Z_Y мають нормальний розподіл, то проблем з побудовою рівняння регресії немає тому, що воно буде лінійним

$$\hat{Z}_Y = b_0 + b_1 Z_1. \quad (1.1)$$

В загальному випадку k незалежних змінних (регресорів) лінійна регресійна модель є такою:

$$Z_Y = \hat{Z}_Y + \varepsilon = \bar{Z}_Y + (\mathbf{Z}_X^+)^T \hat{\mathbf{b}} + \varepsilon, \quad (1.2)$$

де $\hat{Z}_Y$ – результат передбачення рівняння лінійної регресії для значень компонент вектору $\mathbf{z}_X = \{Z_1, Z_2, \dots, Z_k\}$; ε – випадкова величина з розподілом Гаусу, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$; $\mathbf{Z}_X^+$ – матриця центрованих регресорів, яка містить значення $Z_{1_i} - \bar{Z}_1, Z_{2_i} - \bar{Z}_2, \dots, Z_{k_i} - \bar{Z}_k$; $\hat{\mathbf{b}}$ – оцінка для вектору параметрів рівняння лінійної регресії, $\mathbf{b} = \{b_1, b_2, \dots, b_k\}^T$.

У разі нормального розподілу не виникає проблем і з побудовою довірчого інтервалу лінійної регресії. Цей інтервал визначається із використанням квантилів розподілу Стюдента [15]

$$\hat{Z}_{Y_i} \pm t_{\alpha/2, N-2} S_{Z_Y} \sqrt{\frac{1}{N} + \frac{(Z_{1_i} - \bar{Z}_1)^2}{S_{Z_1 Z_1}}},$$

де $\hat{Z}_{Y_i}$ – значення $\hat{Z}_Y$ за рівнянням (1.1) за значенням Z_{1_i} ; Z_{1_i} – i -те значення Z_1 ; $t_{\alpha/2, N-2}$ – квантиль t -розподілу Стюдента з $N - 2$ степенями вільності та

$$\alpha/2 \quad \text{рівнем} \quad \text{значимості}; \quad S_{Z_1 Z_1} = \sum_{i=1}^N (Z_{1_i} - \bar{Z}_1)^2; \quad \bar{Z}_1 = \frac{1}{N} \sum_{i=1}^N Z_{1_i};$$

$$S_{Z_Y}^2 = \frac{1}{N-1} \sum_{i=1}^N (Z_{Y_i} - \bar{Z}_Y)^2; \quad \bar{Z}_Y = \frac{1}{N} \sum_{i=1}^N Z_{Y_i}.$$

Для випадку k незалежних змінних довірчий інтервал лінійної регресії може бути записаний таким чином [15]:

$$\hat{Z}_Y \pm t_{\alpha/2, \nu} S_{Z_Y} \left\{ \frac{1}{N} + (\mathbf{z}_X^+)^T [(\mathbf{Z}_X^+)^T \mathbf{Z}_X^+]^{-1} (\mathbf{z}_X^+) \right\}^{1/2}, \quad (1.3)$$

а інтервал передбачення лінійної регресії як

$$\hat{Z}_Y \pm t_{\alpha/2, \nu} S_{Z_Y} \left\{ 1 + \frac{1}{N} + (\mathbf{z}_X^+)^T [(\mathbf{Z}_X^+)^T \mathbf{Z}_X^+]^{-1} (\mathbf{z}_X^+) \right\}^{1/2}, \quad (1.4)$$

де $S_{Z_Y}^2 = \frac{1}{\nu} \sum_{i=1}^N (Z_{Y_i} - \hat{Z}_{Y_i})^2$, $\nu = N - k - 1$; $(\mathbf{Z}_X^+)^T \mathbf{Z}_X^+ - k \times k$ матриця

$$(\mathbf{Z}_X^+)^T \mathbf{Z}_X^+ = \begin{pmatrix} S_{Z_1 Z_1} & S_{Z_1 Z_2} & \dots & S_{Z_1 Z_k} \\ S_{Z_2 Z_1} & S_{Z_2 Z_2} & \dots & S_{Z_2 Z_k} \\ \dots & \dots & \dots & \dots \\ S_{Z_k Z_1} & S_{Z_k Z_2} & \dots & S_{Z_k Z_k} \end{pmatrix}, \quad (1.5)$$

де $S_{Z_q Z_r} = \sum_{i=1}^N [Z_{q_i} - \bar{Z}_q][Z_{r_i} - \bar{Z}_r]$, $q, r = 1, 2, \dots, k$.

У разі, якщо випадкові величини будуть мати негаусівський розподіл, рівняння та модель регресії будуть нелінійними. Але яку саме не лінійність потрібно застосовувати заздалегідь, як правило, не відомо. Тому на сьогодні для побудови нелінійного рівняння регресії та нелінійної регресійної моделі в основному використовують метод простого перебору. За цим методом беруть різні не лінійності, будують сукупність нелінійних моделей регресії і з цієї сукупності за певним критерієм обирають найкраще рівняння чи найкращу модель. Але застосування методу простого перебору потребує багато часу і не гарантує знаходження найкращого рівняння нелінійної регресії чи нелінійної регресійної моделі тому, що вони в цю сукупність можуть не потрапити. Із-за цього для побудови нелінійного рівняння регресії чи нелінійної регресійної моделі прагнуть використовувати інші методи, які не потребують простого перебору. До цих методів відносять методи, що базуються на лінеарізуючих та

нормалізуючих перетвореннях. Методи, що використовують лінеаризацію, можуть застосовуватися лише в поодиноких випадках, коли вдається знайти лінеарізуюче перетворення. І навпаки, нормалізуюче перетворення, як правило, завжди вдається підібрати, що дає можливість застосовувати для побудови нелінійної регресійної моделі саме метод на основі нормалізуючих перетворень. Суть цього методу полягає у наступному. Спочатку потрібно нормалізувати емпіричні дані, для чого треба здійснити вибір нормалізуючого перетворення і оцінювання його параметрів. Далі за нормалізованими даними отримати лінійну регресійну модель. І на останок, використовуючи обране нормалізуюче перетворення та побудоване лінійне рівняння регресії для нормалізованих даних, здійснити перехід до нелінійної регресійної моделі [12].

Отже за [11] нелінійна регресійна модель може бути побудована за допомогою нормалізуючого перетворення і записана як

$$Y = \psi_Y^{-1} \left[\bar{Z}_Y + (\mathbf{Z}_X^+) \hat{\mathbf{b}} + \varepsilon \right], \quad (1.6)$$

де ψ_Y – перша компонента вектору Ψ , $\Psi = \{\psi_Y, \psi_1, \psi_2, \dots, \psi_k\}^T$, нормалізуючого перетворення

$$\mathbf{T} = \Psi(\mathbf{P}), \quad (1.7)$$

яке має зворотне перетворення

$$\mathbf{P} = \Psi^{-1}(\mathbf{T}). \quad (1.8)$$

В (1.7) та (1.8) $\mathbf{P} = \{Y, X_1, X_2, \dots, X_k\}^T$ – це негаусівський вектор, а $\mathbf{T} = \{Z_Y, Z_1, Z_2, \dots, Z_k\}^T$ – гаусівський вектор.

Аналогічно із застосуванням нормалізуючих перетворень в роботі [10] було запропоновано визначати інтервал передбачення нелінійної регресії за (1.4) як

$$\psi_Y^{-1} \left(\hat{Z}_Y \pm t_{\alpha/2, \nu} S_{Z_Y} \left\{ 1 + \frac{1}{N} + (\mathbf{z}_X^+)^T \left[(\mathbf{Z}_X^+)^T \mathbf{Z}_X^+ \right]^{-1} (\mathbf{z}_X^+) \right\}^{1/2} \right). \quad (1.9)$$

В (1.9) знаки мінус та плюс відповідають рівнянням нижньої та верхньої границь інтервалів передбачення нелінійних регресій; $t_{\alpha/2, \nu}$ – квантіль t -розподілу Стюдента з кількістю ступенів вільності ν та рівнем значущості $\alpha/2$; $\mathbf{Z}_X^+$ – матриця центрованих регресорів, яка містить значення $Z_{1_i} - \bar{Z}_1$, $Z_{2_i} - \bar{Z}_2$, $Z_{3_i} - \bar{Z}_3$; $\mathbf{z}_X^+ = \{Z_{1_i} - \bar{Z}_1, Z_{2_i} - \bar{Z}_2, Z_{3_i} - \bar{Z}_3\}^T$; $S_{Z_Y}^2 = \frac{1}{\nu} \sum_{i=1}^N (Z_{Y_i} - \hat{Z}_{Y_i})^2$, $\nu = N - k - 1$; $(\mathbf{Z}_X^+)^T \mathbf{Z}_X^+ - k \times k$ матриця (1.5).

Практичне застосування нормалізуючих перетворень для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, потребує пошуку відповідного перетворення. У якості нормалізуючого перетворення використовується перетворення у вигляді десяткового логарифму.

1.2 Обґрунтування необхідності проведення досліджень за обраною темою

Як було зазначено в попередньому підрозділі, для оцінювання розміру ПЗ інформаційних системи з відкритим кодом на РНР існують рівняння лінійної регресії [10, 11], які побудовано за допомогою множинного лінійного регресійного аналізу на основі метрик діаграми класів.

Проте існують чотири основні припущення, які виправдовують використання моделей лінійної регресії, одним з яких є нормальність розподілу помилок. Але це припущення справедливе лише в окремих випадках. Це призводить до необхідності використання нелінійних

регресійних моделей та моделей, в тому числі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Також було з'ясовано, що нормалізуючі перетворення часто є добрим способом побудови моделей, довірчих інтервалів та інтервалів передбачення нелінійних регресій [11-21]. В роботах [20-24] запропоновано рівняння нелінійної регресії для оцінювання розміру ПЗ інформаційних системи, що розробляються на базі мови програмування PHP. Запропоноване рівняння побудовано за допомогою множинного нелінійного регресійного аналізу із застосуванням чотиривимірного перетворення Джонсона сім'ї S_B на основі таких же метрик діаграми класів, що і в [10, 11]: загальна кількість класів, загальна кількість зв'язків та середня кількість атрибутів на клас. Але для PHP-застосунків, що не є інформаційними системами, наприклад, таких як різноманітні фреймворки, включаючи фреймворк CakePhp, конвертори, HTTP-клієнти та інші, регресійні моделі можуть залежати в тому числі від інших метрик. Наприклад, існує регресійна модель для веб-застосунків, що розроблюються з використанням фреймворку Laravel, яка залежить від інших метрик [3].

Зазвичай при побудові нелінійних регресійних моделей у якості нормалізуючого перетворення використовують десятковий логарифм. Наприклад, такі відомі моделі як COCOMO, ISBSG, COCOMO II були створені із використанням нормалізуючого перетворення у вигляді десятичного логарифму. Тому створювати регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, доцільно саме на основі нормалізуючого перетворення у вигляді десятичного логарифму.

Таким чином, виникає потреба у проведенні досліджень за обраною темою.

1.3 Висновки за розділом 1

У даному розділі наведено огляд існуючих методів та моделей для оцінювання розміру web застосунків з відкритим кодом, здійснено обґрунтування необхідності проведення досліджень за обраною темою.

1) Зараз для оцінювання розміру web застосунків використовуються методи лінійного та нелінійного регресійного аналізу.

2) Існуючі регресійні рівняння для оцінювання розміру web застосунків, як правило, є лінійними. Використання моделей лінійної регресії потребує виконання певних чотирьох припущень, одним з яких є нормальність розподілу помилок. Але це припущення справедливе лише в окремих випадках. Це призводить до необхідності розробки нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

3) Існуючі нелінійні регресійні рівняння та моделі для оцінювання розміру web застосунків побудовані за метриками концептуальної моделі даних у вигляді діаграми класів для інформаційних систем: загальна кількість класів, загальна кількість зв'язків та середня кількість атрибутів на клас. Але для РНР-застосунків з відкритим кодом, що не є інформаційними системами, наприклад, таких як різноманітні фреймворки, в тому числі фреймворк CakePhp, конвертори, HTTP-клієнти та інші, регресійні моделі можуть залежати в тому числі від інших метрик. А це веде до необхідності створення нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

4) Нормалізуючі перетворення часто є добрим способом побудови моделей, моделей, довірчих інтервалів та інтервалів передбачення нелінійних регресій. Причому зазвичай у якості нормалізуючого перетворення використовують десятковий логарифм. Наприклад, такі відомі моделі як

COCOMO, ISBSG, COCOMO II були створені саме із використанням нормалізуючого перетворення у вигляді десяткового логарифму. Але використання одновимірних нормалізуючих перетворень для побудови моделей нелінійної регресії приводить до гірших результатів оцінювання, ніж використання багатовимірних нормалізуючих перетворень. Таким чином, виникає потреба у проведенні досліджень за обраною темою.

Результати досліджень поточного розділу автором опубліковано в роботі [6].

2 НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ CAKEPHP

2.1 Перевірка даних для побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на наявність мультиколінеарності та викидів

Для побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, були зібрані дані з метрик 38 web застосунків, що були створені з використанням фреймворку CakePhp та розташованих на сайті GitHub (<https://github.com>): кількість строк коду у тисячах KLOC; кількість класів Classes; середня кількість методів на кожен клас; глибина дерева успадкування DIT. Ці дані були отримані за допомогою відомого інструменту PhpMetrics (<https://phpmetrics.org/>) та наведені у таблиці 2.1.

Таблиця 2.1

Набір даних та квадрат відстані Махаланобіса (SMD)

№	Назва	KLOC	Classes	Methods by class	DIT	SMD
1	2	3	4	5	6	7
1	Slug	0,448	4,00	4,75	1,40	2,07

Продовження табл. 2.1

1	2	3	4	5	6	7
2	cakephp-api	7,846	90,00	3,86	1,71	0,08
3	migrations	4,345	42,00	4,19	1,81	0,11
4	cakephp-realworld-example-app	2,717	60,00	2,20	2,11	3,84
5	webadmin	2,954	45,00	3,13	2,13	2,34
6	audit-stash	1,717	26,00	2,54	1,44	2,09
7	cakephp-jwt-auth	0,212	1,00	7,00	2,00	4,31
8	cakephp-file-storage	1,149	14,00	3,86	1,73	0,30
9	app	0,477	8,00	2,63	2,20	3,84
10	infinitas	61,269	487,00	4,93	1,79	19,41
11	cakephp-social-share	0,124	2,00	2,50	2,00	2,60
12	cakephp-csvview	0,358	2,00	3,50	2,00	1,22
13	FlyPie	0,349	8,00	2,25	1,38	3,15
14	cakephp3-bootstrap-helpers	3,486	29,00	3,62	2,00	1,05
15	CakePdf	1,538	11,00	5,82	1,80	1,51
16	Footprint	0,365	4,00	4,00	1,50	1,08
17	bootstrap-ui	2,332	23,00	4,00	1,79	0,21
18	passbolt_api	24,347	328,00	3,59	1,24	7,78
19	cakephp-tools	12,433	66,00	6,52	1,79	3,08
20	asset_compress	0,948	10,00	4,30	1,78	0,28
21	search	1,826	20,00	4,15	1,60	0,47
22	CakePHP-Proffer	0,882	9,00	3,56	1,60	0,72

Продовження табл. 2.1

1	2	3	4	5	6	7
23	Webservice	3,567	23,00	6,83	1,24	7,30
24	debug_kit	3,735	49,00	3,57	1,70	0,22
25	cms	47,61	289,00	5,33	1,33	18,86
26	users	4,029	58,00	3,38	1,54	0,61
27	crud	3,5	40,00	4,18	1,63	0,23
28	cakephp-new-relic	0,131	3,00	2,00	1,67	2,73
29	cakephp-tinyauth	2,227	23,00	4,22	1,46	1,11
30	AclManager	0,521	2,00	8,50	3,00	20,51
31	authorization	1,494	22,00	3,41	1,35	1,94
32	cakephp-imagine-plugin	0,906	5,00	7,80	1,60	7,59
33	CakePHP-HybridAuth	0,471	5,00	3,80	2,00	0,97
34	bedita	24,04	303,00	3,67	1,75	5,34
35	acl	2,681	23,00	4,78	1,71	0,39
36	vendor	20,327	308,00	3,94	1,20	9,35
37	croogo	24,66	367,00	3,36	1,64	11,55
38	cakephp-adminlte-theme	0,142	3,00	3,00	2,00	1,76

В таблиці 2.1 також наведені значення квадрату відстані Махаланобіса SMD. Значення SMD для кожної багатовимірної точки даних i , $i=1,2,\dots,N$, визначається як

$$d_i^2 = (\mathbf{X}_i - \bar{\mathbf{X}})^T \mathbf{S}_N^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}), \quad (2.1)$$

де $\bar{\mathbf{X}}$ – вектор вибірових середніх; $\mathbf{S}_N$ – вибірова коваріаційна матриця

$$\mathbf{S}_N = \frac{1}{N} \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T. \quad (2.2)$$

Тут $\mathbf{X}_i$ – випадковий вектор $\mathbf{X} = \{X_1, X_2, \dots, X_m\}^T$ для кожної багатовимірної точки даних i , $i = 1, 2, \dots, N$; m – кількість компонентів вектору $\mathbf{X}$. В нашому випадку m дорівнює 4.

Зазначимо, що для даних з таблиці 2.1 вектор $\mathbf{X}$ є негаусівським випадковим вектором тому, що для застосунка 30 значення SMD більше ніж величина квантіля розподілу χ^2 для рівня значимості 0,005. Також про негаусівський розподіл чотиривимірних даних з таблиці 2.1 свідчить значення багатовимірного ексцесу β_2 , оцінка якого визначається як [25]

$$\hat{\beta}_2 = \frac{1}{N} \sum_{i=1}^N \left\{ (\mathbf{X}_i - \bar{\mathbf{X}})^T \mathbf{S}_N^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}) \right\}^2. \quad (2.3)$$

Відомо, що для багатовимірного нормального розподілу $\beta_2 = m(m+2)$. У нашому випадку $\beta_2 = 24$. Для чотиривимірних даних з таблиці 2.1 оцінка β_2 за (2.3) дорівнює 44,38.

Враховуючи на те, що нам необхідно побудувати регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, потрібно перевірити майбутні фактори (предиктори) на наявність мультиколінеарності.

Під мультиколінеарністю розуміють наявність лінійної залежності між двома або більше незалежними змінними (факторами) у регресійній моделі. Мультиколінеарність – це негативне явище множинного регресійного аналізу, яке не дозволяє оцінити окремий вплив кожного фактору (регресора) на залежну змінну. Наявність мультиколінеарності свідчить про те, що в множинній регресійній моделі два або більше факторів (незалежних змінні) пов'язані між собою або мають високий ступінь кореляції.

Наявність мультиколінеарності будемо визначати за коефіцієнтами впливу дисперсії (VIFs) серед майбутніх предикторів (факторів) в моделі множинної лінійної регресії. Для лінійної моделі множинної регресії з k -предикторами X_i , $i = 1, \dots, k$, VIFs – це діагональні елементи оберненої коваріаційної матриці $k \times k$ k -предикторів [26]. Значення VIFs більше за 10 часто сприймаються як сигнал, що дані мають проблеми з мультиколінеарністю. У разі, якщо значення VIFs знаходяться у межах від 1 до 5, то мультиколінеарності немає.

Коефіцієнти впливу дисперсії для 3 факторів (Classes, Methods, DIT) для даних таблиці 2.1 мають значення 1.08, 1.03, 1.11 відповідно. Значення VIFs, менші за 5. Це вказує на те, що дані не мають проблеми з мультиколінеарністю.

Враховуючи негаусівський розподіл чотиривимірних даних з таблиці 2.1 в подальшому для виявлення викидів у цих даних з таблиці 2.2 ми використовуємо метод, оснований на нормалізуючих перетвореннях і квадраті відстані Махаланобіса SMD [27, 28], який для кожної багатовимірної точки нормалізованих даних i , $i = 1, 2, \dots, N$, визначається як

$$d_i^2 = (\mathbf{Z}_i - \bar{\mathbf{Z}})^T \mathbf{S}_N^{-1} (\mathbf{Z}_i - \bar{\mathbf{Z}}), \quad (2.4)$$

де $\bar{\mathbf{Z}}$ – вектор вибірових середніх нормалізованих даних; $\mathbf{S}_N$ – вибіркова коваріаційна матриця нормалізованих даних

$$\mathbf{S}_N = \frac{1}{N} \sum_{i=1}^N (\mathbf{Z}_i - \bar{\mathbf{Z}})(\mathbf{Z}_i - \bar{\mathbf{Z}})^T. \quad (2.5)$$

Тут $\mathbf{Z}_i$ – гаусівський випадковий вектор $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_m\}^T$ для кожної багатовимірної точки даних i , $i = 1, 2, \dots, N$; m – кількість компонентів вектору $\mathbf{Z}$. В нашому випадку m дорівнює 4.

Вектор $\mathbf{Z}$ отримують шляхом перетворення негаусівського випадкового вектору $\mathbf{X} = \{X_1, X_2, \dots, X_m\}^T$ за допомогою багатовимірною нормалізуючого перетворення (1.6)

В нашому випадку, що розглядається, в перетворенні (1.6) $\Psi = \{\lg Y, \lg X_1, \lg X_2, \lg X_3\}^T$.

Відомо [17], якщо дані мають багатовимірний нормальний розподіл, то розподіл SMD поводить ся як розподіл χ^2 . Для великих $N - m$ (принаймні більше 25), відстань d_i^2 повинна вести себе приблизно як незалежні випадкові величини $\chi_{m,\alpha}^2$. Тут $\chi_{m,\alpha}^2$ – це квантиль розподілу χ^2 з рівнем значимості α . Також, якщо дані мають багатовимірний нормальний розподіл, то статистика

$$\frac{N(N-m)d_i^2}{m(N^2-1)} \quad (2.6)$$

поводиться як F -розподіл. При чому статистика (2.6) повинна вести себе приблизно як незалежні випадкові величини $F_{m,N-m,\alpha}$. Тут $F_{m,N-m,\alpha}$ – це квантиль F -розподілу з рівнем значимості α .

Рядок даних, для якого величина SMD більша, ніж квантиль розподілу χ^2 , та величина статистики (2.6) більша, ніж квантиль F -розподілу, розглядається як багатомірний викид, і відкидається. Після відкидання зменшене число багатомірних точок даних знову нормалізується, використовуючи перетворення (1.6), доки всі рядки даних не будуть мати величину SMD більшу, ніж квантиль розподілу χ^2 , та величину статистики (2.6) більшу, ніж квантиль F -розподілу.

Для рівня значимості 0,005 та перетворення у вигляді десяткового логарифму викиди відсутні тому, що для всіх рядків даних значення статистики (2.6) є меншим ніж величина квантіля $F_{m,N-m,\alpha}$, яка становить 4,50 для рівня значимості 0,005. Хоча застосування критерія на основі SMD до

даних з таблиці 2.1 без їх нормалізації показало, що дані застосунку 30 є чотиривимірним викидом, оскільки для цього рядка даних SMD дорівнює 20,51 і є більшим ніж величина квантіля розподілу χ^2 , яка становить 14,86 для рівня значимості 0,005, та значення статистики (2.6), що дорівнює 4,39, є меншим ніж величина квантіля $F_{m,N-m,\alpha}$, яка становить 4,50 для рівня значимості 0,005.

В подальшому чотиривимірні негаусові дані, що наведені у таблиці 2.1, використовуються для побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Але спочатку для остаточного обґрунтування необхідності її побудови отримуємо лінійну регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, яка представляється у вигляді

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + \hat{b}_3 X_3 + \varepsilon, \quad (2.7)$$

де оцінки параметрів такі: $\hat{b}_0 = -4.4155$, $\hat{b}_1 = 0.9974$, $\hat{b}_2 = 1.02683$, $\hat{b}_3 = -0.05258$; ε – випадкова величина з розподілом Гаусу, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$; $\hat{\sigma}_\varepsilon$ – це оцінка середньо квадратичного відхилення, значення якої дорівнює 4,7637.

Перевіримо нульову гіпотезу про нормальність закону розподілу випадкової величини ε для моделі (2.7). Для цього будемо використовувати критерій Пірсона.

У разі великої вибірки значень випадкової величини (коли $n > 30$) перевірку нормальності закону розподілу результатів спостережень в тому числі виконують за критерієм Пірсона (критерієм χ^2). Відповідно за цим критерієм спочатку обчислюють значення χ^2 [29]

$$\chi^2 = \sum_{j=1}^m \frac{(n_j - np_j)^2}{np_j}, \quad (2.8)$$

де m – кількість підінтервалів (часток, кліток), на які розбивається інтервал $[x_{\min}, x_{\max}]$; $x_{\min}$ і $x_{\max}$ – відповідно мінімальне і максимальне значення випадкової величини X ; n_j – абсолютна частота в j -му підінтервалі (кількість значень випадкової величини, які попадають у j -ий підінтервал); p_j – ймовірність того, що значення випадкової величини X попадають у j -ий підінтервал.

Кількість підінтервалів (кліток), на які розбивається інтервал $[x_{\min}, x_{\max}]$ може бути визначений за наступними формулами: $m = \log_2 n + 1 = 3,3 \lg n + 1$ (формула Старджеса) або $m = 5 \lg n$ (формула Брукса і Карузера). Ми використовуємо формулу Старджеса, за якою $m = 6$.

Ймовірність того, що значення випадкової величини X попадають у j -ий підінтервал можуть бути визначені як

$$p_j = \int_{x_{j-1}}^{x_j} f(x) dx,$$

де x_{j-1} і x_j – ліва і права границі j -го підінтервалу; $f(x)$ – функція щільності ймовірності.

Функція щільності ймовірності нормального розподілу (розподілу Гауса) є такою:

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}},$$

де m_x – математичне сподівання випадкової величини X .

Після цього в залежності від рівня значимості α та кількості ступенів вільності ν за таблицею верхніх 100α %-вих точок розподілу χ^2 (знаходять значення $\chi_{кр}^2$). Кількість ступенів вільності ν визначається як

$$\nu = m - k - 1,$$

де k – кількість параметрів, від яких залежить закон розподілу. Для нормального розподілу $k = 2$.

Якщо $\chi^2 \leq \chi_{кр}^2$, то з ймовірністю $1 - \alpha$ можна прийняти гіпотезу про те, що закон розподілу результатів спостережень є нормальним, якщо $\chi^2 > \chi_{кр}^2$ – цю гіпотезу потрібно відкинути.

Для вибірки значень випадкової величини ε значення χ^2 дорівнює 30,06 за формулою (2.8), а $\chi_{кр}^2$ для 3 ступенів вільності та 0,05 рівня значимості дорівнює 7,81. Тобто $\chi^2 > \chi_{кр}^2$, і цю гіпотезу про нормальність розподілу випадкової величини ε потрібно відкинути.

У свою чергу це свідчить про відсутність теоретичного обґрунтування використання моделі лінійної регресії для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp. Це призводить до необхідності побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, за даними, що наведені у таблиці 2.1.

2.2 Побудова нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp

Для побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, спочатку чотиривимірні негаусівські дані з таблиці 2.1 ми нормалізуємо за допомогою одновимірного перетворення у вигляді десяткового логарифму. Далі для нормалізованих даних будуюмо лінійну регресійну модель

$$Z_Y = \hat{Z}_Y + \varepsilon = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2 + \hat{b}_3 Z_3 + \varepsilon, \quad (2.9)$$

де $Z_Y = \lg Y$; $Z_j = \lg X_j$, $j = 1, 2, 3$.

Параметри моделі (2.9) оцінювалися методом найменших квадратів. Оцінки параметрів такі: $\hat{b}_0 = -1.77226$, $\hat{b}_1 = 0.980278$, $\hat{b}_2 = 1.178674$, $\hat{b}_3 = 0.285567$.

На останок будуюмо нелінійну регресійну модель за лінійним рівнянням (2.9) та перетворенням у вигляді десяткового логарифму

$$Y = 10^{\varepsilon + \hat{b}_0} X_1^{\hat{b}_1} X_2^{\hat{b}_2} X_3^{\hat{b}_3}. \quad (2.10)$$

Рівняння (2.10) – це і є нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Побудована модель (2.10) була перевірена за множинним коефіцієнтом детермінації R^2 , середньою величиною відносної помилки MMRE і відсотком прогнозованих результатів, для яких величини відносної помилки MMRE менші за 0,25, PRED(0,25). Ці показники зазвичай використовуються для оцінювання якості прогнозування за допомогою регресійних моделей і в інженерії програмного забезпечення [30, 31]. Допустимі значення MMRE і

PRED(0,25) складають не більше 0,25 і не менше 0,75 відповідно. Допустиме значення множинного коефіцієнта детермінації R^2 приблизно таке ж, як для PRED(0,25).

Для моделі (2.10), що була побудована за даними метрик 38 web застосунків, що створюються з використанням фреймворку CakePhp з таблиці 2.1 значення R^2 , MMRE і PRED(0,25) що дорівнюють 0,964, 0,109 і 0,895 відповідно, вказують на задовільну якість моделі (2.10) з оцінками параметрів, що були отримані за даними з таблиці 2.1 для 38 web застосунків, що створюються з використанням фреймворку CakePhp.

Для покращення якості нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, було застосовано метод покращення нелінійних регресійних моделей на основі нормалізуючих перетворень із застосуванням квадрату відстані Махаланобіса та інтервалів передбачення [32]. Суть цього методу [32] полягає у наступному. Спочатку на першому етапі, як це зазвичай робиться, початкові негаусівські дані перевіряються на наявність викидів і, якщо останні знайдено, то вони відкидаються. Для цього використовується квадрат відстані Махаланобіса для нормалізованих даних. На першому етапі рівень значущості дорівнює 0,005. Далі на другому етапі будується нелінійна регресійна модель із застосуванням відповідного методу на основі нормалізуючих перетворень [11]. Після цього на третьому етапі для рівня значущості, що дорівнює 0,05, визначаються границі інтервалу передбачення нелінійної регресії за методом, наведеним в [11]. І на завершення, на четвертому етапі перевіряють, чи є серед даних, за якими будувалася нелінійна регресійна модель такі, що виходять за визначені границі інтервалу передбачення. Та, якщо останні знайдено, то вони відкидаються і ми повторюємо знову всі етапи, починаючи з першого, для нових даних. Якщо таких викидів не було, то повторення етапів завершується, відповідна нелінійна регресійна модель побудована.

Для визначення нижньої і верхньої границь інтервалів передбачення нелінійних регресій використовуємо відповідні рівняння (1.8).

Нижні і верхні границі інтервалів передбачення нелінійних регресій для нелінійної регресійної моделі (1.8) наведені в таблиці 2.2. Для нелінійної регресійної моделі (2.10) з оцінками параметрів, що були отримані за даними з таблиці 2.1 з 38 web застосунків, що створюються з використанням фреймворку CakePhp, виявилось: значення Y для двох застосунків 12 і 38 виходить за визначені межі інтервалу передбачення, що визначалися за (1.8). В таблиці 2.2 ліва границя інтервалу передбачення, що отриманий на першій ітерації за (1.8), позначена як LB_1 , а права – як UB_1 .

На другій ітерації дані з застосунків 14 та 25 були відкинуті і для аналізу використовувалися дані із залишених 35 застосунків. Для моделі (2.10) з оцінками параметрів, що були отримані за даними з таблиці 2.2 з 38 застосунків виявилось: значення Y для двох застосунків (6 і 22) виходять за визначені межі інтервалу передбачення. В таблиці 2.2 ліва границя інтервалу передбачення, що отриманий на третій ітерації за (1.8), позначена як LB_3 , а права – як UB_3 .

Таблиця 2.2

Нижні і верхні границі інтервалів передбачення нелінійних регресій

№	Y	Межі інтервалів передбачення нелінійних регресій					
		LB_1	UB_1	LB_2	UB_2	LB_3	UB_3
1	2	3	4	5	6	7	8
1	0,448	0,304	0,679	0,341	0,580	0,369	0,507
2	7,846	5,409	11,738	6,192	10,337	6,586	8,924
3	4,345	2,875	6,212	3,264	5,435	3,501	4,735
4	2,717	1,943	4,422	2,234	3,854	2,403	3,323
5	2,954	2,256	5,000	2,573	4,359	2,777	3,797

Продовження таблиці 2.2

1	2	3	4	5	6	7	8
6	1,717	0,924	2,035	1,053	1,776	-	-
7	0,212	0,135	0,308	0,149	0,258	0,167	0,232
8	1,149	0,880	1,892	0,989	1,643	1,065	1,439
9	0,477	0,340	0,758	0,382	0,652	0,418	0,574
10	61,269	37,464	84,841	43,732	75,163	46,450	64,136
11	0,124	0,080	0,179	0,089	0,152	0,098	0,135
12	0,358	0,120	0,264	-	-	-	-
13	0,349	0,246	0,556	0,279	0,479	0,298	0,412
14	3,486	1,726	3,755	1,954	3,271	-	-
15	1,538	1,132	2,469	1,270	2,129	1,383	1,877
16	0,365	0,255	0,561	0,285	0,481	0,309	0,421
17	2,332	1,507	3,243	1,701	2,826	1,829	2,470
18	24,347	15,854	35,462	18,498	31,536	19,271	26,495
19	12,433	7,423	16,447	8,480	14,356	9,140	12,496
20	0,948	0,723	1,559	0,811	1,349	0,878	1,187
21	1,826	1,328	2,860	1,500	2,491	1,607	2,171
22	0,882	0,506	1,094	0,567	0,947	-	-
23	3,567	2,475	5,650	2,826	4,880	3,031	4,197
24	3,735	2,721	5,873	3,097	5,153	3,302	4,464
25	47,61	22,768	50,906	26,460	45,130	-	-
26	4,029	2,923	6,323	3,337	5,560	3,536	4,787
27	3,5	2,657	5,722	3,017	5,012	3,220	4,350
28	0,131	0,086	0,196	0,097	0,167	0,105	0,146
29	2,227	1,509	3,270	1,709	2,850	1,824	2,470
30	0,521	0,363	0,889	0,407	0,737	0,467	0,666
31	1,494	1,093	2,395	1,241	2,086	1,319	1,795
32	0,906	0,700	1,585	0,786	1,349	0,863	1,188
33	0,471	0,326	0,709	0,364	0,609	0,397	0,540
34	24,04	16,664	37,036	19,349	32,823	20,478	28,028
35	2,681	1,833	3,953	2,069	3,441	2,226	3,009
36	20,327	16,448	36,924	19,189	32,800	20,001	27,551
37	24,66	17,783	39,546	20,705	35,139	21,797	29,846
38	0,142	0,149	0,327	-	-	-	-

Всього було 5 таких ітерації, після яких залишилося 32 застосунка. На п'ятій ітерації викидів не було, повторення етапів завершується, нелінійна

регресійна модель остаточно побудована за даними з 32 застосунків. В таблиці 2.3 ліва границя інтервалу передбачення, що отриманий на п'ятій ітерації за (1.8), позначена як LB_5 , а права – як UB_5 . Остаточно на п'ятій ітерації для даних з 32 застосунків оцінки параметрів такі: $\hat{b}_0 = -1.80484$, $\hat{b}_1 = 0.981865$, $\hat{b}_2 = 1.186738$, $\hat{b}_3 = 0.319789$.

Сума квадратів відхилень для моделі (2.10) у цьому випадку склала 0,02976, що майже у 2,6 рази менше за відповідну суму на першій ітерації. На четвертій ітерації для нормалізованих за перетворенням у вигляді десяткового логарифму даних з 32 застосунків матриця S_z для визначення нижньої і верхньої границь інтервалів передбачення нелінійної регресії за (1.8) є такою:

$$S_z = \begin{pmatrix} 15.9 & -0.45 & -0.61 \\ -0.45 & 0.73 & 0.03 \\ -0.61 & 0.03 & 0.21 \end{pmatrix}.$$

Остаточно після п'ятої ітерації модель (2.10) виявилася кращою ніж була після першої ітерації за всіма трьома показниками: множинним коефіцієнтом детермінації R^2 , середньою величиною відносної помилки MMRE і відсотком прогнозованих результатів PRED(0,25). Значення R^2 , MMRE і PRED(0,25) для моделі (2.15) після п'ятої ітерації складають 0,987, 0,055 і 1,000 відповідно, а після першої ітерації є відповідно такими: 0,964, 0,109 і 0,895. Різниця між цими значеннями R^2 складає 2,33%. Різниця між значеннями MMRE складає 98,18%. Різниця між значеннями PRED(0,25) складає 10,5%.

Також після четвертої ітерації нелінійну регресійну модель (2.10) ми порівняли з лінійною моделлю (2.9) з оцінками параметрів $\hat{b}_0 = -6.027$, $\hat{b}_1 = 0.0935$, $\hat{b}_2 = 0.7326$, $\hat{b}_3 = 1.4636$, що були знайдені за методом найменших квадратів для 32 застосунків, та з оцінкою $\hat{\sigma}_e$, значення якої дорівнює 3,871. Нелінійна регресійна модель (2.10) виявилася кращою за лінійну модель (2.9) за всіма трьома показниками: R^2 , MMRE і PRED(0,25). Значення R^2 , MMRE і

PRED(0,25) для лінійної регресійної моделі значно гірші: 0,926, 1,900 і 0,344 відповідно. Якщо різниця між значеннями R^2 складає всього 6,18%, то значення MMRE та PRED(0,25) для лінійної регресійної моделі гірші на 3354,54% та 65,6% відповідно.

2.3 Висновки за розділом 2

У даному розділі виконано перевірку даних для побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на наявність викидів та побудовано нелінійну регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

1) Перевірку даних для побудови нелінійної регресійної моделі для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на наявність викидів виконано за методом, який оснований на нормалізуючих перетвореннях і квадраті відстані Махаланобіса.

2) Удосконалено регресійну модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, на основі нормалізуючого перетворення у вигляді десяткового логарифму.

3) Отримані результати свідчать про те, що нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, має високу якість оцінювання. Ця нелінійна модель в порівнянні з лінійною регресійною моделлю має більші значення множинного коефіцієнту детермінації R^2 і відсотку прогнозованих результатів PRED(0,25), менше значення середньої величини відносної похибки MMRE та менші ширини довірчого інтервалу і інтервалу передбачення для більшої кількості значень незалежних змінних.

Результати досліджень поточного розділу автором опубліковано в роботі [6].

3 РОЗРОБКА ПРОЕКТУ ПРОГРАМИ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ WEB ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ CAKEPHP

3.1 Постановка задачі на розробку проекту програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Як правило сучасні програмні системи обробки інформації є орієнтованими на кінцевого користувача, тому зазвичай розробляються з зручним графічним інтерфейсом та отриманням (візуалізацією) результату обробки даних. Для реалізації програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp необхідно розглянути проектні рішення з вибору мови програмування.

Тобто це повинна бути програма, яка буде надавати можливість вводити вхідні дані, а також в якій будуть реалізовані алгоритми обробки цих даних за допомогою обраних методів.

Тому для розробки програми необхідно визначитися з мовою програмування, якою можливо реалізувати це ПЗ.

Для вирішення цього можна використовувати будь-які мови програмування, але оскільки програма виконує складні математичні обчислення, бажано щоб мова програмування дозволяло легко працювати з ними. Цим умовам відповідає мова Python.

Мова Python (<https://www.python.org/>) є високорівневою інтерпретованою мовою програмування. Включає засновані на матрицях структури даних, широкий спектр функцій, інтегроване середовище розробки,

об'єктно-орієнтовані можливості та інтерфейси до програм, написаних на інших мовах програмування. У зв'язці з бібліотеками для математичних обчислень NumPy та SciPy є альтернативою комерційним пакетам моделювання.

Переваги – безкоштовне для комерційного використання, достатній вбудований функціонал, не потребує багато обчислювальних ресурсів, працює на більшості ОС. Для розробки програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp недоліків не виявлено.

Отже будемо використовувати мову Python з бібліотеками NumPy та SciPy для розробки програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Для реалізації віконного інтерфейсу користувача та математичних алгоритмів обробки інформації добре підійде бібліотека Qt (<https://www.qt.io>), тому що вона добре документована, широко розповсюджена, а також має обгортку для мови Python.

Виходячи з цього можна зробити висновок, що мова Python добре підійде для розробки програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Виходячи з аналізу існуючих методів та моделей для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, сформулюємо постановку задачі. Потрібно створити програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, яка би задовольняла таким вимогам.

Вимоги до програми.

Вимоги функціональних характеристик.

Вимоги до складу функцій, що виконуються.

Програма має виконувати такі функції:

1) Вводити з інтерфейсу програми довірчу ймовірність, три метрики діаграми класів, за якими потрібно оцінити розмір ПЗ web застосунків, що створюються з використанням фреймворку CakePhp в тисячах строк коду: друга X_1 , третя X_2 та четверта X_3 метрики визначають відповідно загальну кількість класів, середню кількість методів на кожен клас та глибину дерева успадкування DIT.

2) Виконувати обчислення розміру web застосунка, що створюється з використанням фреймворку CakePhp.

3) Виконувати обчислення довірчого інтервалу нелінійної регресії.

4) Виконувати обчислення інтервалу передбачення нелінійної регресії.

Вимоги до складу й параметрів технічних засобів.

Програма має експлуатуватись на обладнанні з параметрами не нижчими за наступні: 64-бітний процесор на архітектурі x86 або ARM, оперативна пам'ять 4Гб (2 Гб мінімум), простір на жорсткому диску 100 Мб мінімум.

Вимоги до інформаційної та програмної сумісності.

Для запуску та роботи з програмним забезпеченням необхідна операційна система Windows 8/10/11, Ubuntu 18/19/20 або Mac OS 10.12/10.13/10.14

На підставі цієї постановки задачі розроблено технічне завдання, яке наведено у додатку А.

3.2 Розробка ескізного проекту програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp

В ескізному проекті було виявлено варіанти використання та побудовано їх сценарії, розроблено зовнішню специфікацію програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Даний проект передбачає наявність лише одного актора – користувача програми, який буде повноцінно використовувати всі функції програми. Виявлені варіанти використання знаходяться в таблиці 3.1.

Таблиця 3.1 Виявлені варіанти використання

Основні актори	Найменування	Формулювання
1	2	3
Користувач	Ввести три метрики діаграми класів	Вводить три метрики діаграми класів, за якими потрібно оцінити розмір web застосунку, що створюється з використанням фреймворку CakePhp: загальну кількість класів, загальну кількість зв'язків між класами та загальну кількість методів.
Користувач	Ввести довірчу ймовірність	Вводить довірчу ймовірність, для якої потрібно оцінити довірчий інтервал та інтервал передбачення нелінійної регресії розміру web застосунку, що створюється з використанням фреймворку CakePhp.
Користувач	Отримати розмір web застосунка	Отримує розмір web застосунка, що створюється з використанням фреймворку CakePhp.

Продовження таблиці 3.1

1	2	3
Користувач	Отримати довірчий інтервал	Отримує довірчий інтервал нелінійної регресії для інтервального оцінювання вибіркового середнього розміру web застосунку, що створюється з використанням фреймворку CakePhp
Користувач	Отримати інтервал передбачення	Отримує інтервал передбачення нелінійної регресії для інтервального оцінювання вибіркового середнього розміру web застосунку, що створюється з використанням фреймворку CakePhp
Користувач	Ввести фактичний розмір web застосунку	Вводить розмір web застосунку що був розроблений з використанням фреймворку CakePhp, для якого буде оцінено відносну похибку
Користувач	Отримати відносну похибку	Отримує відносну похибку для веб застосунку, що був розроблений з використанням фреймворку CakePhp

Діаграма варіантів використання програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp зображена на рисунку 3.1.

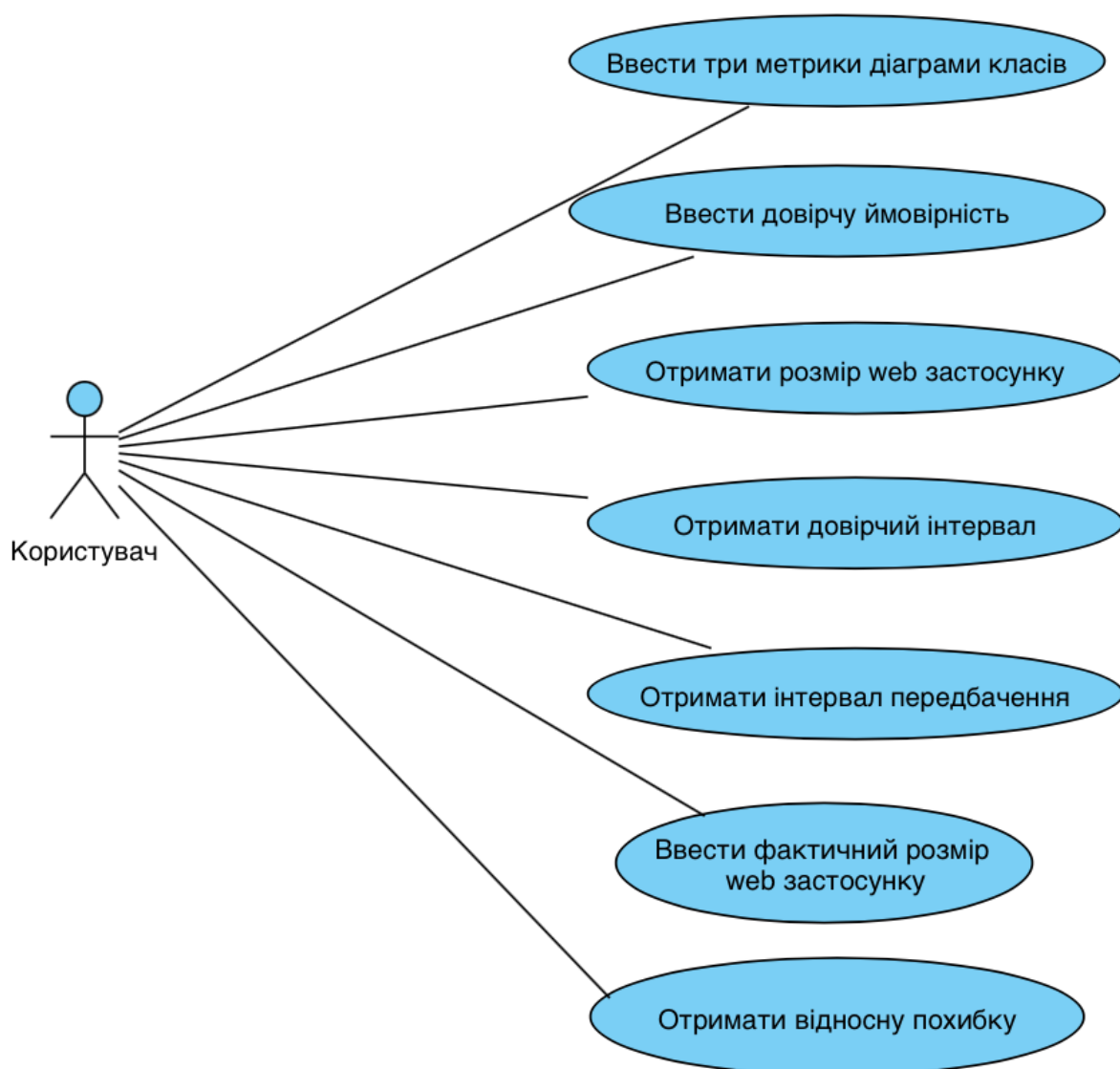


Рисунок 3.1 – Діаграма варіантів використання ПЗ для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp

Сценарії варіантів використання зображені у вигляді діаграми діяльності на рисунку 3.2.

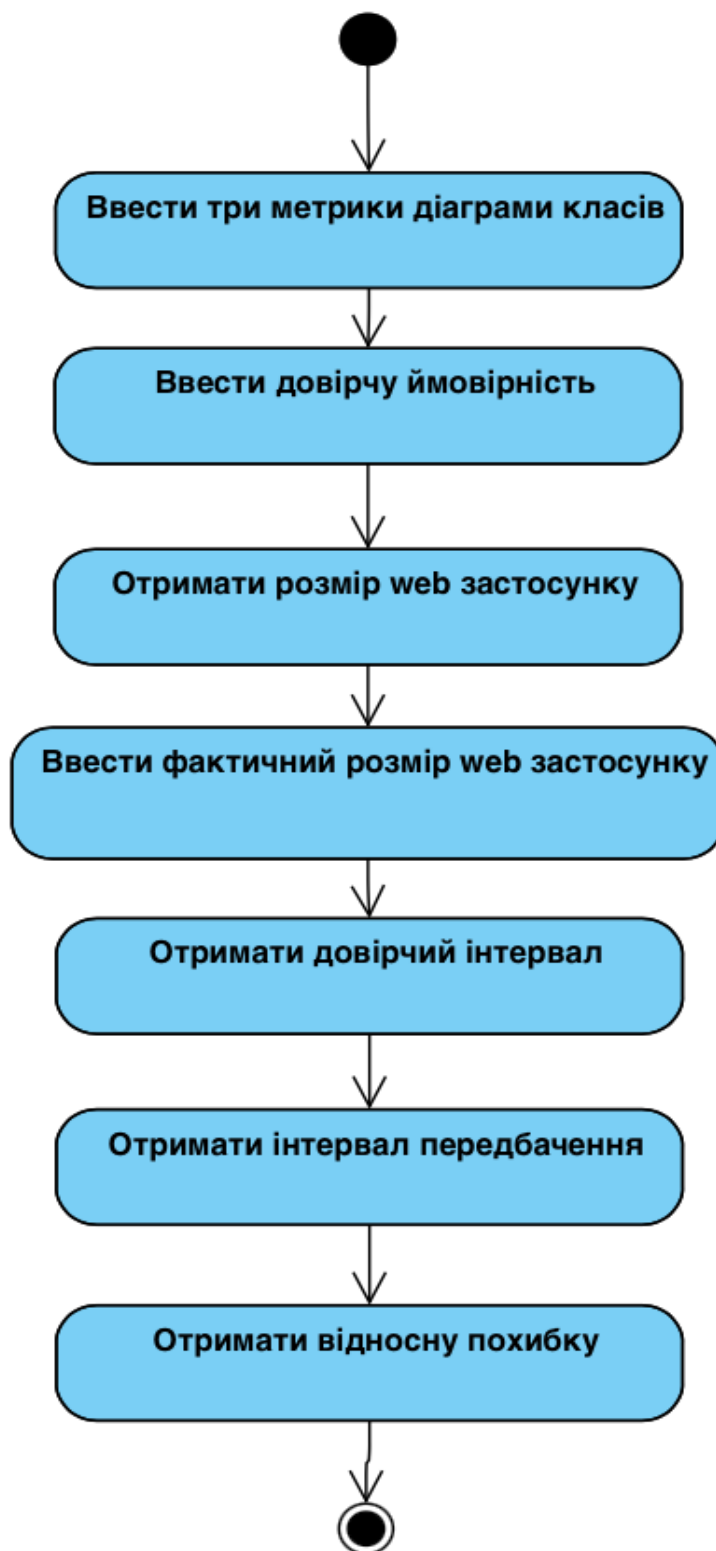


Рисунок 3.2 – Діаграма діяльності ПЗ для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp

Зовнішня специфікація програми.

Назва програми: CakePhpSizeEst.

Функції, які виконує програма:

1) Вводити з інтерфейсу програми довірчу ймовірність, три метрики діаграми класів, за якими потрібно оцінити розмір ПЗ web застосунків, що створюються з використанням фреймворку CakePhp в тисячах строк коду: друга X_1 , третя X_2 та четверта X_3 метрики визначають відповідно загальну кількість класів, середню кількість методів на кожен клас та глибину дерева успадкування DIT.

2) Виконувати обчислення розміру web застосунка, що створюється з використанням фреймворку CakePhp.

3) Виконувати обчислення довірчого інтервалу нелінійної регресії.

4) Виконувати обчислення інтервалу передбачення нелінійної регресії.

5) Вводити з інтерфейсу програми фактичний розмір веб застосунку, що був розроблений з використанням фреймворку CakePhp.

6) Виконувати обчислення відносної похибки.

Вхідні дані:

- значення довірчої ймовірності, трьох метрик діаграми класів, за якими потрібно оцінити розмір web-застосунку в тисячах строк коду: друга X_1 , третя X_2 та четверта X_3 метрики визначають відповідно загальну кількість класів, загальну кількість зв'язків між класами та загальну кількість методів.

- значення фактичного розміру веб застосунку, що був розроблений з використанням фреймворку CakePhp.

Вихідні дані:

- значення розміру web застосунка, що створюється з використанням фреймворку CakePhp, границі довірчого інтервалу та інтервалу передбачення.

- значення відносної похибки.

Зовнішні ефекти: поява графічного вікна для вводу та виводу даних, управління виводом даних, вихід з програми.

3.3 Розробка технічного проекту програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp

У технічному проекті програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, було розроблено діаграму класів, яка зображена на рисунку 3.3.

Для відображення динамічної моделі ПЗ було створено діаграми послідовностей, що зображені на рисунках 3.4 – 3.10.

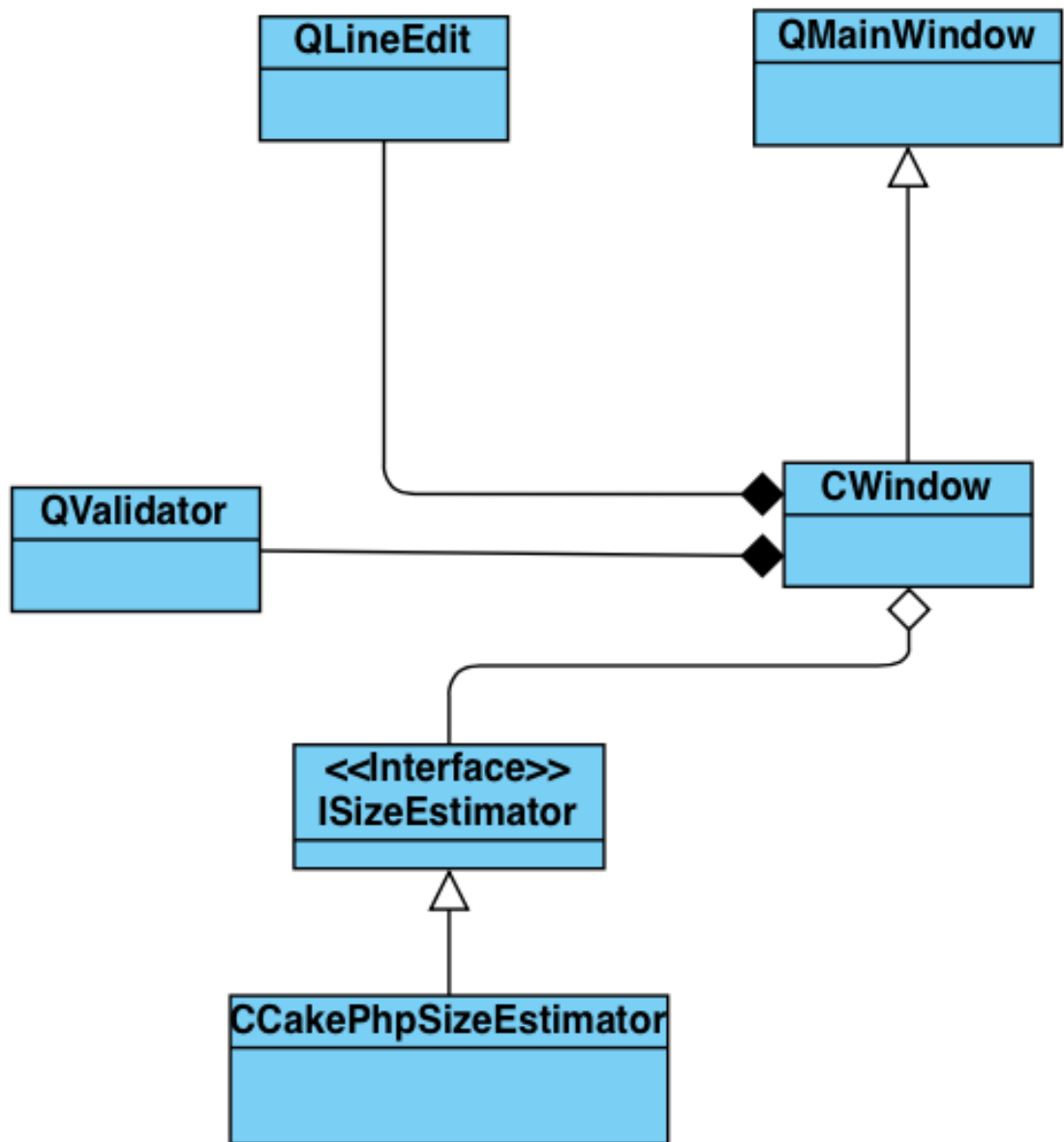


Рисунок 3.3 – Діаграма класів програми

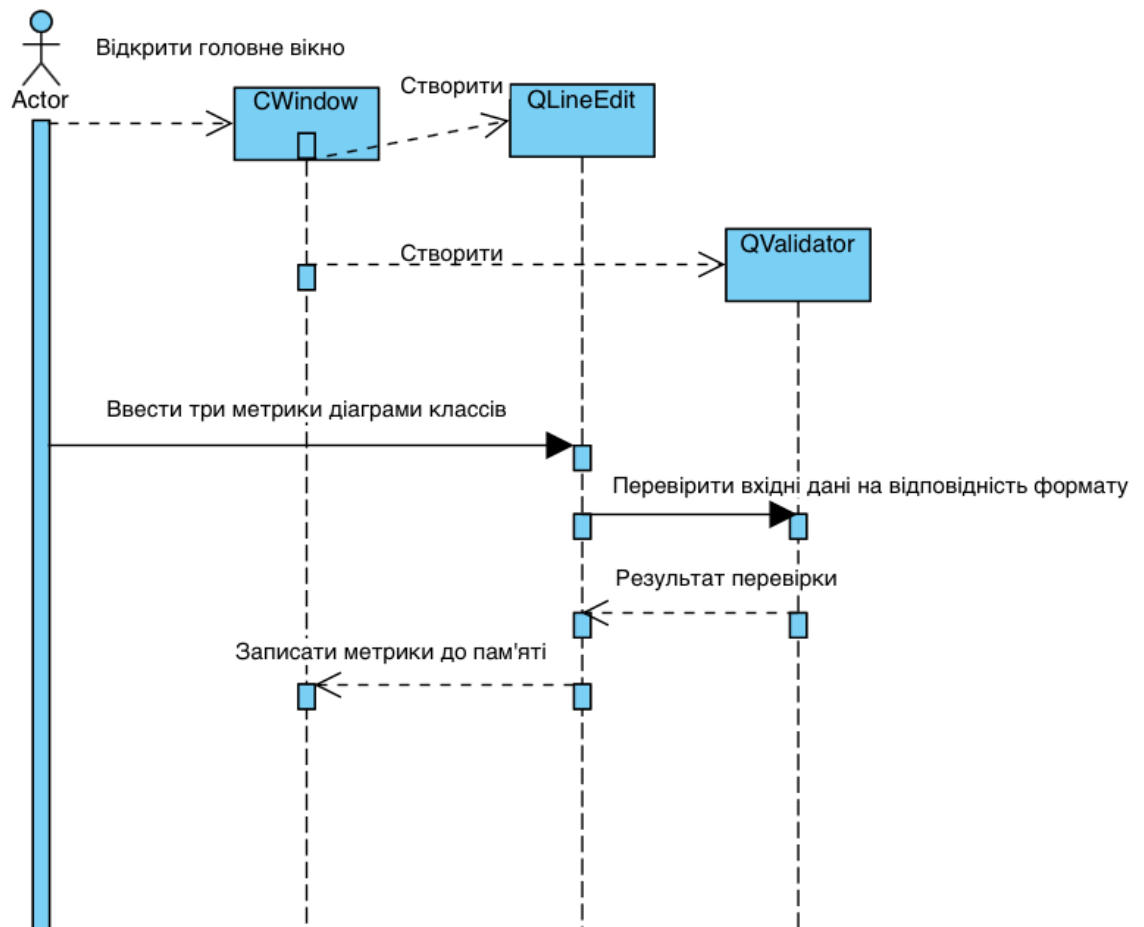


Рисунок 3.4 – Діаграма послідовності для варіанту використання «Ввести три метрики діаграми класів»

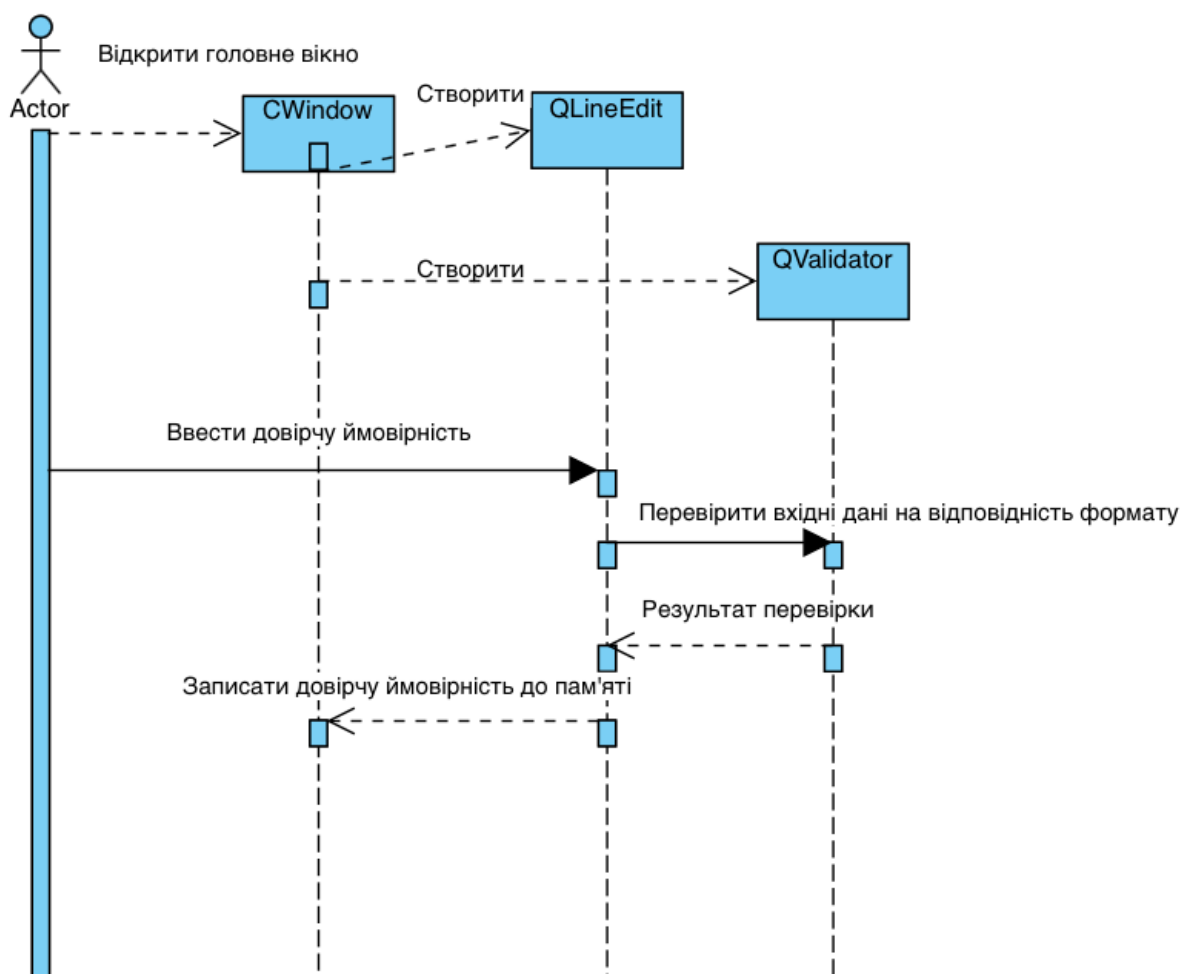


Рисунок 3.5 – Діаграма послідовності для варіанту використання «Ввести довірчу ймовірність»

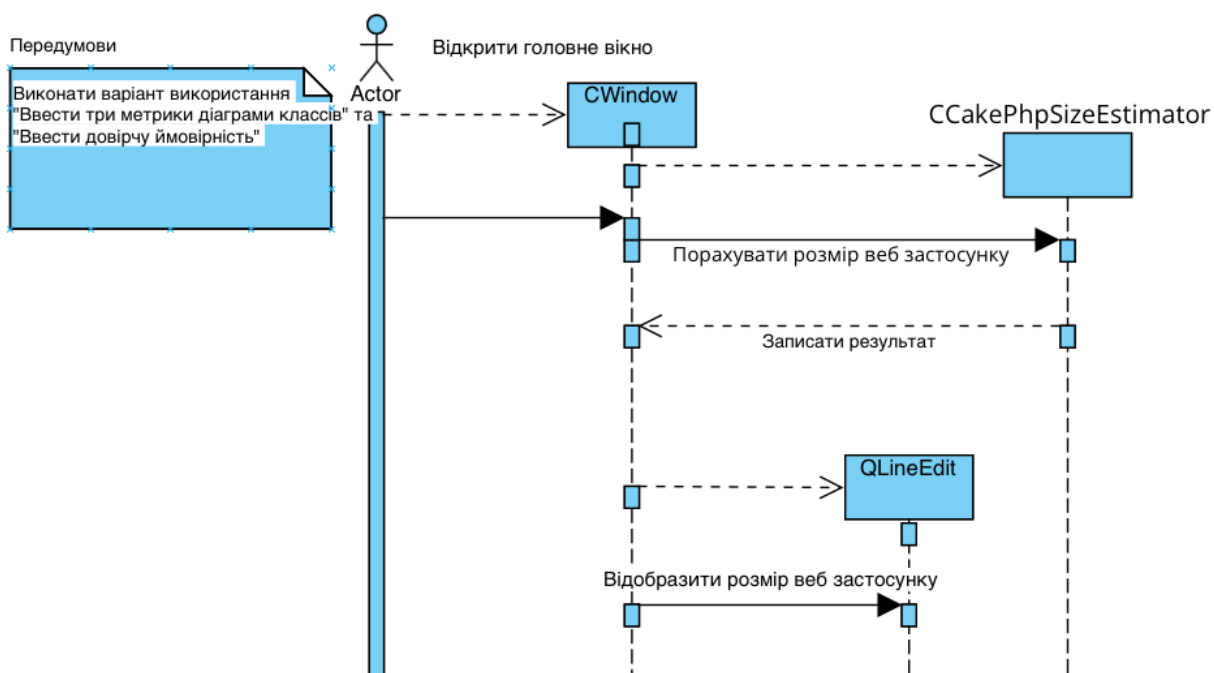


Рисунок 3.6 – Діаграма послідовності для варіанту використання «Отримати розмір web застосунку»

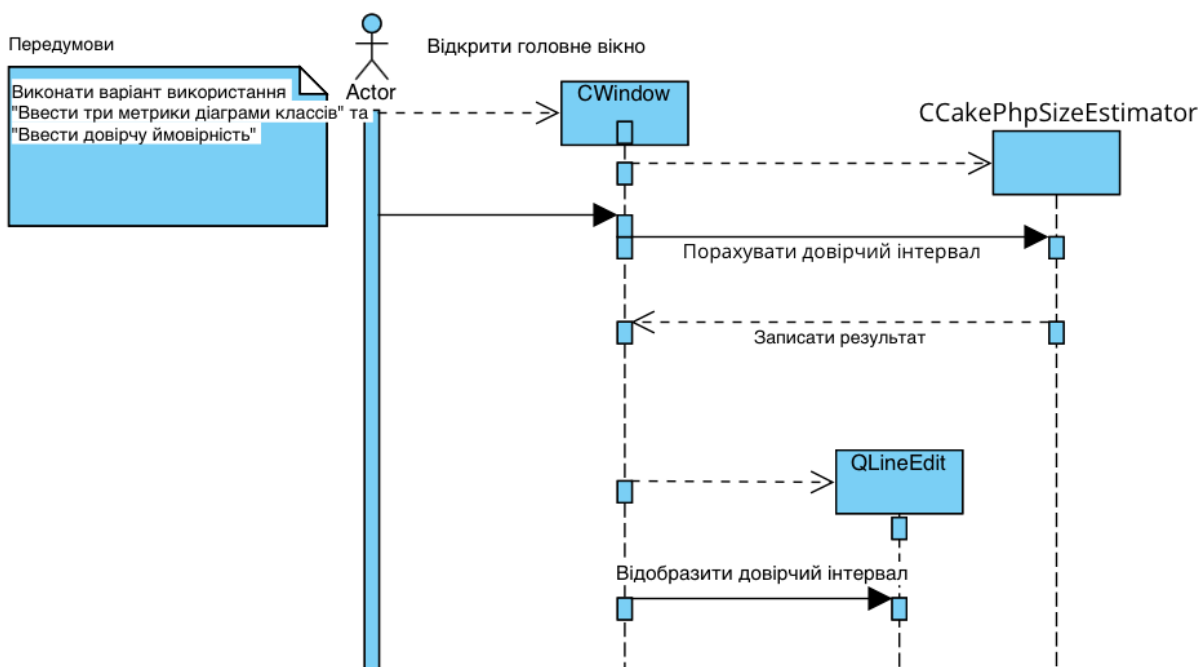


Рисунок 3.7 – Діаграма послідовності для варіанту використання «Отримати довірчий інтервал»

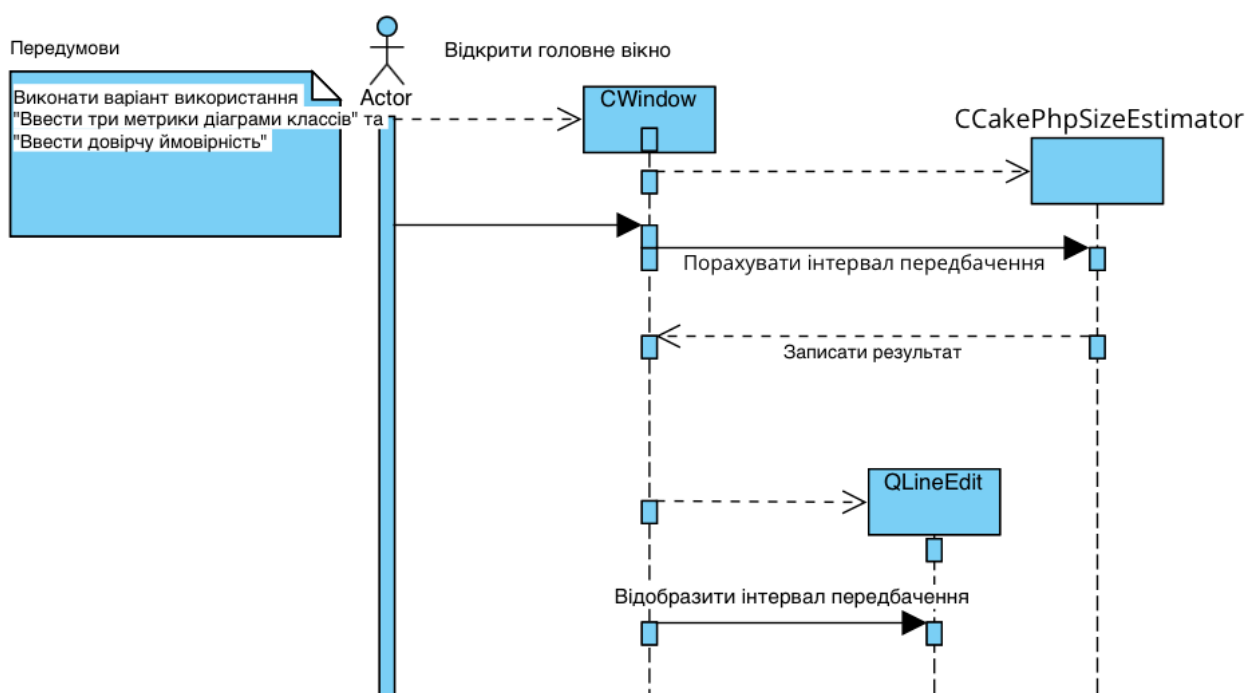


Рисунок 3.8 – Діаграма послідовності для варіанту використання «Отримати інтервал передбачення»

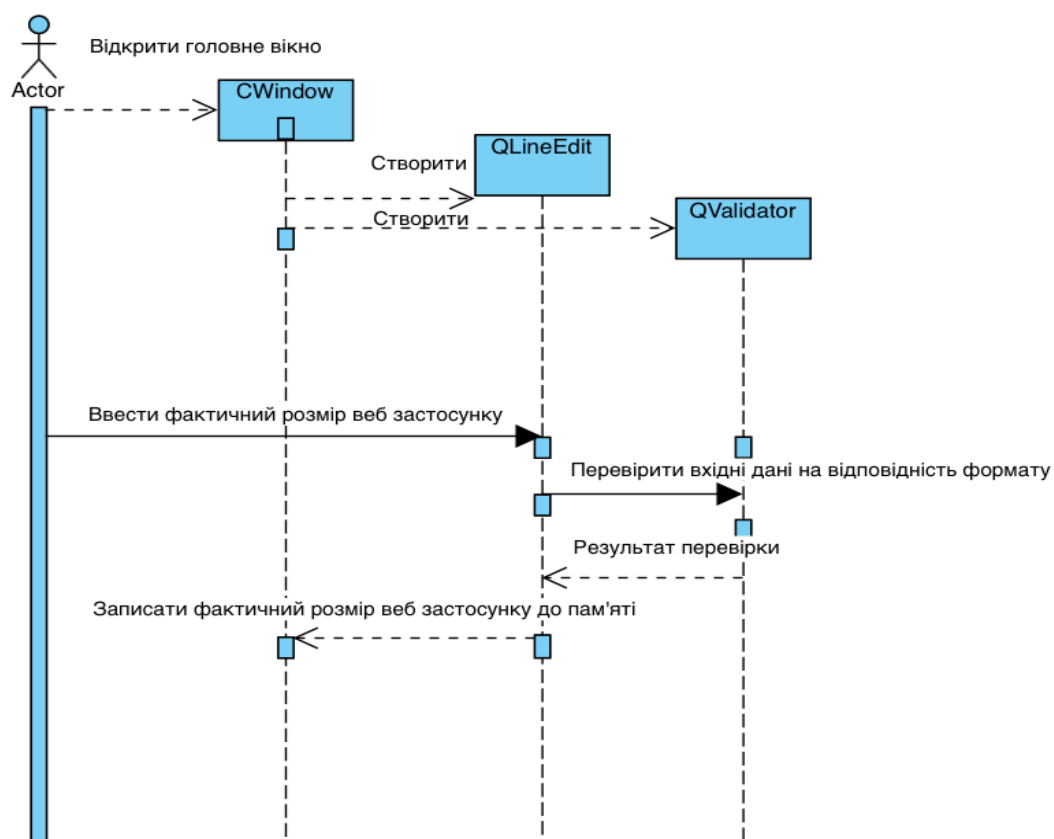


Рисунок 3.9 – Діаграма послідовності для варіанту використання «Ввести фактичний розмір web застосунку»

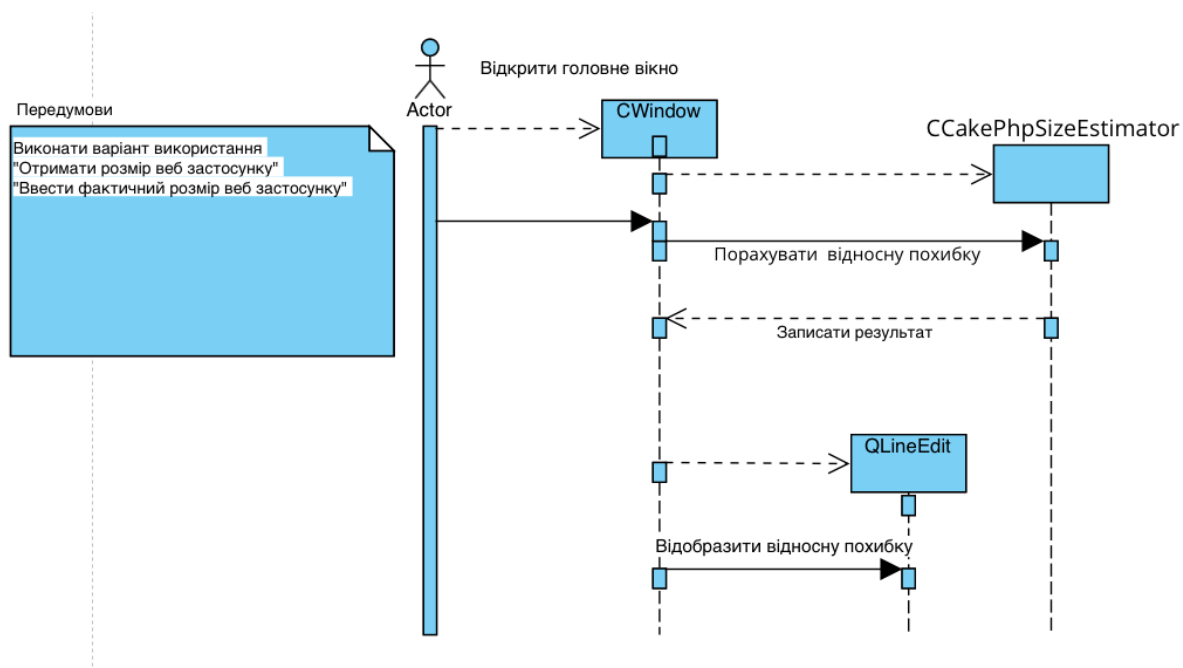


Рисунок 3.10 – Діаграма послідовності для варіанту використання «Отримати відносну похибку»

Наступним кроком після розробки технічного проекту є розробка робочого проекту.

3.4 Розробка робочого проекту та тестування програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp

В робочому проекті було розроблено програму для оцінювання розміру web застосунків, що створюється з використанням фреймворку CakePhp, здійснено тестування програми, розроблено програмну документацію, здійснено випробування програми.

Спочатку мовою Python було виконано кодування класів, що були розроблені на стадії технічного проекту. Текст програми наведений в додатку Б.

В програмі були використані такі інтерфейсні компоненти: рядки редагування та рядки для виводу тексту. Графічне вікно програми із зазначеними інтерфейсними компонентами наведено на рис.3.11.

Class count	4
Avg methods count	4,75
Depth of inheritance tree	1,4
Confidence level	0,95
Actual size	0,4
KLOC	0.4326
Confidence interval	0.4093 - 0.4573
Prediction interval	0.3695 - 0.5065
MRE	7.54%

Рисунок 3.11 – Графічне вікно програми з інтерфейсними компонентами

Було проведено функціональне тестування програми методом чорної скриньки. Для перевірки було використано дані з таблиць 2.1 та 2.2.

Програму та методику випробовувань наведено в Додатку Д.

Протокол тестування:

- 1) Виконана перевірка на відповідність технічному завданню:
 - перевірка функції вводу трьох метрик діаграми класів.
 - перевірка функції вводу довірчої ймовірності.

- перевірка функції отримання розміру web-застосунку.
- перевірка функції отримання довірчого інтервалу.
- перевірка функції отримання інтервалу передбачення.
- перевірка функції вводу фактичного розміру web-застосунку.
- перевірка функції отримання відносної похибки.

Всі функції були виконані успішно.

Для функції отримання відносної похибки було перевірено, яка буде поведінка, якщо не буде введено фактичного розміру web-застосунку. Для такого випадку програма не буде обчислювати відносну похибку.

2) Перевірка програми була виконана на операційних системах Mac OS 11.5, Windows 10 та Linux.

Інтерфейс програми для Windows та Linux зображений на рисунках 3.12 та 3.13, відповідно.

Class count	11
Avg methods count	5,82
Depth of inheritance tree	1,80
Confidence level	0,95
Actual size	1,538
KLOC	1.6109
Confidence interval	1.5492 - 1.6750
Prediction interval	1.3827 - 1.8768
MRE	4.52%

Рисунок 3.12 – Графічне вікно програми на платформі Windows 10

Class count	11
Avg methods count	5,82
Depth of inheritance tree	1,80
Confidence level	0,95
Actual size	1,538

KLOC	1.6109
Confidence interval	1.5492 - 1.6750
Prediction interval	1.3827 - 1.8768
MRE	4.52%

Рисунок 3.13 – Графічне вікно програми на платформі Linux

Було перевірено, що на різних платформах програма має однаковий функціонал за виключенням стилів інтерфейсу, який є специфічним для кожної платформи.

3.5 Висновки за розділом 3

У даному розділі було виконано постановку задачі на розробку проекту програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, розроблено ескізний, технічний та робочий проекту програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

1. В ескізному проекті було виявлено варіанти використання та побудовано їх сценарії та розроблено зовнішню специфікацію програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

2. В технічному проекті було спроектовано програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp на основі покрокової деталізації зверху вниз, розроблено специфікацію головного модулю програми, побудовано алгоритми головного модулю програми.

3. В робочому проекті було розроблено програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, здійснено тестування програми методом чорної скриньки, розроблено програмну документацію, здійснено випробування програми.

4 РОЗРАХУНОК ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ ВІД РОЗРОБКИ І ВПРОВАДЖЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Впровадження інформаційних технологій пов'язано з капітальними вкладеннями як на придбання техніки, так і на розробку проектів, виконання підготовчих робіт і підготовку кадрів. Тому впровадженню повинно передувати економічне обґрунтування доцільності впровадження інформаційної системи. Це означає, що повинна бути визначена ефективність використання автоматизованих комп'ютерних технологій.

4.1 Розрахунок витрат на створення й експлуатацію програмного забезпечення

Витрати на розробку ПЗ складаються з витрат на зарплату розробника, на амортизацію комп'ютера, на якому виконується розробка, на експлуатацію цього комп'ютера, на засоби розробки та витратні матеріали і комплектуючі.

Розробка ПЗ виконується програмістом, місячний оклад якого складає 10000 грн. Додаткова заробітна плата складає 20% від основної. Виходячи з цього, основна і додаткова заробітна плата розроблювача системи 12000 грн/міс, а вартість сучасного комп'ютера складає 15000 грн. (приведена вартість комп'ютера на базі Intel Core i9).

Вартість розробки ПЗ розраховується по формулі:

$$\text{Спр} = (\text{Ззп} + \text{Зсз} + \text{Ззг} + \text{Зе}) * \text{T} + \text{Зм},$$

де Т – тривалість розробки, міс.; Ззп – основна і додаткова заробітна плата обслуговуючого персоналу, грн.; Зсз – відрахування на соціальні заходи

(38% від основної і додаткової заробітної плати), грн.; Ззг – загальногосподарські витрати (10% від основної заробітної плати), грн.; Зе – витрати на електроенергію; Зм – витрати на основні і допоміжні матеріали.

Зе при споживаній потужності 0,5 кВт, тривалості роботи на місяць, рівної $22 \times 6 = 132$ годин, вартості кіловат-години електроенергії 1,68 грн. складає:

$$З_e = 132 * 1,68 * 0,5 = 110,88 \text{ грн.}$$

Витрати на допоміжні матеріали приведені в табл. 4.1.

Таблиця 4.1 – Витрати на допоміжні матеріали

Пункти витрат	Сума, грн.
Папір	90,00
Заправлення картриджа до принтера	120,00
Література	450,00
Непередбачені витрати	250,00
Разом	910,00

Витрати на розробку ПЗ приведені в табл. 4.2.

Таблиця 4.2 – Витрати на розробку системи

Найменування витрат	Одиниця виміру	Кількість
Тривалість розробки	Міс.	1
Основна і додаткова заробітна плата	Грн.	12000,00
Відрахування на соціальні заходи	Грн.	4560,00
Загальногосподарські витрати	Грн.	1000,00
Витрати на допоміжні матеріали	Грн.	910,00
Витрати на електроенергію	Грн.	110,88

Відповідно до формули, вартість розробки ПЗ складає:

$$C_{\text{пр}} = (12000 + 4560 + 1000 + 110,88) * 1 + 910 = 18580,88 \text{ грн.}$$

Амортизаційні відрахування на устаткування складають 60% балансової вартості в рік:

$$A_{\text{об}} = 15000 * 0,6 = 9000 \text{ грн.}$$

У масштабах підприємства річні витрати на основні і допоміжні матеріали визначаються у розмірі 5% вартості основного устаткування:

$$B_{\text{м}} = 15000 * 0,05 = 750 \text{ грн.}$$

Річний обсяг робіт комп'ютера у годинах визначається в такий спосіб:

$$F_{\text{м}} = 264,5 * T_{\text{з}}$$

де $T_{\text{з}}$ – це середнє місячне завантаження устаткування (близько 4 годин);

264,5 – середня кількість робочих днів у році.

Отже, річний обсяг роботи комп'ютера складе:

$$F_{\text{м}} = 264,5 * 4 = 1058 \text{ годин.}$$

Витрати на електроенергію $З_{\text{е}}$ при 1058 годинах роботи устаткування в рік складуть

$$З_{\text{е}} = 1058 * 1,68 * 0,5 = 888,72 \text{ грн.}$$

Експлуатаційні витрати для комп'ютера за рік складуть:

$$З_{\text{зр}} = 9000 + 750 + 888,72 = 10638,72 \text{ грн.}$$

Отже, у перший рік витрати на створення й експлуатацію ПЗ:

$$З_{\text{се}} = 18580,88 + 10638,72 = 29219,60 \text{ грн.}$$

4.2 Економічна ефективність розробки і впровадження програмного забезпечення

Основним показником економічної ефективності функціонування системи є підвищення ефективності керування інформацією у вигляді зниження витрат на керування при одночасному збільшенні швидкості і якості одержання потрібного результату.

Крім багатьох інших негативних ефектів, ручна обробка інформації спричиняє наступні негативні економічні ефекти:

- високі витрати на складування паперових документів (сейфи, шафи, папки й ін.);
- підвищені витрати на канцтовари;
- витрати, пов'язані з коригуванням раніше допущених помилок (людський фактор).

До числа основних факторів, що визначають приріст прибутку в зв'язку з впровадженням ПЗ, відносяться:

- підвищення продуктивності праці;
- вивільнення робочого часу.

Крім того, не піддається прямій грошовій оцінці підвищення оперативності керування, якість одержуваних результатів, поліпшення організації праці і т.д.

Обов'язковою умовою визначення економічної ефективності ПЗ є порівнянність усіх показників у часі, за цінами й іншими нормами, використовуваними для визначення показників, за змістом і колом елементів витрат.

Визначимо пряму економічну ефективність, ґрунтуючись на тому, що впровадження ПЗ вивільняє 0,6 працівника (за експертною оцінкою фахівців).

Зарплата 0,6 працівника в рік складає:

$$(8000 * 12) * 0,6 = 57600 \text{ грн.}$$

Річний економічний ефект розраховується по формулі:

$$\Delta C_{\text{год}} = \Delta C_{\text{п}} - E_{\text{п}} \cdot k$$

де $\Delta C_{\text{п}}$ – вивільнені кошти після впровадження ПЗ (57600 грн.) мінус експлуатаційні витрати (10638,72 грн.) = 46961,28 грн.

$E_{\text{п}}$ – коефіцієнт ефективності (дорівнює коефіцієнту амортизації(0,6));

k – одноразові витрати на впровадження продукту (29219,60);

$$\Delta C_{\text{год}} = 46961,28 - 29219,60 * 0,6 = 10645,01 \text{ (грн.)}$$

Строк окупності системи розраховується по формулі:

$$T = \frac{k}{\Delta C} = \frac{29219,60}{46961,28} \approx 0,62$$

Отже строк окупності системи складає приблизно 7 місяців.

5 ОХОРОНА ПРАЦІ В ОФІСІ ФІРМИ

5.1 Вступна частина

Охорона праці – ряд систем розрахованих на догляд за життям та здоров'ям робітників у приміщенні де працюють люди. До охорони праці стосуються процеси трудової діяльності, які включають багато заходів щодо захисту людської праці, а саме правові, лікувально-профілактичні, санітарно-гігієнічні, соціально-економічні та інші.

До функцій охорони праці входить саме проведення цих заходів, задля зниження впливу шкідливих факторів на організм робітників, шляхом застосування технік безпеки під час виконання роботи працівниками.

До техніки безпеки можна віднести ряд організаційних та технічних методів гарантування безпеки, що при виконанні необхідних для техніки безпеки заходів запобігають відсутністю небезпечних факторів, або зниження можливих випадків.

Робоче місце – місце постійного або тимчасового перебування працівника під час його трудової діяльності, в якому виконують виробничі завдання. Робоче місце є частиною виробничо-технологічної структури підприємства (організації), воно призначене для виконання частини технологічного (виробничого) процесу і визначається на основі трудових та інших діючих норм і нормативів.

Робоча зона – визначений простір, у якому розташовані робочі місця постійного або тимчасового перебування працівника під час його трудової діяльності, обмежений по висоті над рівнем підлоги або майданчика. До постійних відносяться робочі місця, на яких працює знаходиться більше 50%

робочого часу за зміну або більше двох годин безперервно. Якщо робота 80 здійснюється в різних пунктах робочої зони, то постійним робочим місцем вважається вся робоча зона.

Умови праці – сукупність факторів виробничого середовища, що впливає на здоров'я і працездатність людини в процесі праці. Дослідження умов праці показали, що чинниками виробничого середовища в процесі праці є:

- санітарно-гігієнічна обстановка, яка визначає зовнішню середу в робочій зоні - мікроклімат, механічні коливання, випромінювання, температуру, освітлення та інші;
- психофізіологічні елементи: робоча поза, фізичне навантаження, нервовопсихологічну напругу та інші, які обумовлені самим процесом праці;
- естетичні елементи: оформлення виробничих приміщень, обладнання, робочого місця, робочого інструменту та інші;
- соціально-психологічні елементи, складові характеристики так званого психологічного клімату.

Згідно Конституції України (ч. 4 ст. 43) кожна працездатна людина має право на належні, безпечні і здорові умови праці. А тому, у відповідності до вимог ст. 153 Кодексу законів про працю України, ст. 6 та ч. 1 ст. 13 Закону України «Про охорону праці», роботодавець зобов'язаний створити на робочому місці в кожному структурному підрозділі умови праці відповідно до нормативно-правових актів, а також забезпечити додержання вимог законодавства щодо прав працівників у галузі охорони праці.

Отже, треба провести аналіз та виявити шкідливі та небезпечні фактори, які можуть впливати на працездатність, загрожувати здоров'ю чи життю працівників на робочих місцях у офісі, провести аналіз їх впливу на здоров'я,

встановити граничнодопустимі рівні цих факторів на робочих місцях, тим самим визначити оптимальні умови необхідні для праці та навчання.

5.2 Аналіз шкідливих та небезпечних факторів в офісі фірми

Для аналізу шкідливих та небезпечних факторів необхідно розглянути характеристики умов праці робітників, а також можливі фактори, що можуть негативно вплинути на їхнє здоров'я.

Під умовами праці розуміють сукупність факторів та обставин трудового процесу і виробничого середовища, в якому здійснюється діяльність людини, що впливають на здоров'я та працездатність.

Під факторами трудового процесу розуміють основні його характеристики: важкість праці та напруженість праці. Важкість праці – це характеристика трудового процесу, яка відображає переважне навантаження на опорно-руховий апарат і функціональні системи організму (серцево-судинну, дихальну та інші, що забезпечують його діяльність). Напруженість праці – це характеристика трудового процесу, яка відображає навантаження переважно на центральну нервову систему, органи чуття, емоційну сферу працівника.

Під виробничим середовищем розуміють сукупність фізичних, хімічних, біологічних, психофізіологічних факторів на виробництві, що діють на людину. Усі ці фактори класифікують як небезпечні та шкідливі.

Небезпечні виробничі фактори – це ті, вплив яких на працівника призводить до травм, раптового погіршення здоров'я чи до смерті.

Шкідливі виробничі фактори – це ті, вплив яких на працівника може призвести до захворювання та зниження працездатності.

До фізичних небезпечних і шкідливих факторів належать:

- рухомі машини і механізми, рухомі частини виробничого обладнання, вироби, що пересуваються (матеріали, заготовки);
- підвищена запиленість і загазованість повітря робочої зони;
- підвищена чи знижена температура поверхонь обладнання, матеріалів, повітря робочої зони;
- підвищені рівні шуму, вібрації, ультразвуку, інфразвукових коливань;
- підвищений чи знижений барометричний тиск і його різкі зміни;
- підвищена чи знижена вологість, рухомість, іонізація повітря;
- підвищений рівень іонізуючих випромінювань, напруги в електромережі, статичної електрики, електромагнітних випромінювань, напруженості електричного і магнітного полів;
- відсутність чи брак природного світла, знижена контрастність, пряма і відбита блискотливість, підвищена пульсація світлового потоку;
- підвищені рівні ультрафіолетової та інфрачервоної радіації;
- гострі краї, шершавість на поверхні заготовок, інструментів і обладнання;
- розташування робочого місця на значній висоті відносно землі (підлоги).

До хімічних небезпечних і шкідливих виробничих факторів належать речовини, які за характером впливу на організм людини поділяються на токсичні, подразнюючі, сенсабілізуючі, канцерогенні і мутагенні. Ці хімічні речовини впливають на репродуктивну функцію людини. За шляхами проникнення в організм людини вони поділяються на ті, що проникають через органи дихання, шлунково-кишковий тракт, шкіряний покрив і слизові оболонки.

До біологічних небезпечних і шкідливих виробничих факторів належать патогенні мікроорганізми (бактерії, віруси, рикетсії, спірохети, грибки) та продукти їх життєдіяльності, а також рослини та живі істоти.

До психофізіологічних небезпечних і шкідливих виробничих факторів належать фізичні (статичні і динамічні) та нервово-психічні перевантаження (розумове перенапруження, перенапруження аналізаторів, монотонність праці, емоційні перевантаження).

До небезпечних факторів стосується і гігієнічні умови праці робочих місць, що негативно впливають на здоров'я людей, які там працюють.

Гігієнічне оцінювання умов і характеру праці на робочих місцях виконується на основі гігієнічної класифікації праці за показниками шкідливості та небезпечності факторів виробничого середовища, важкості та напруженості трудового процесу. Гігієнічна класифікація базується на принципі диференціації умов праці залежно від фактично визначених рівнів указаних факторів у порівнянні з санітарними нормами, правилами, гігієнічними нормами, а також з урахуванням можливого шкідливого впливу їх на стан здоров'я працівників.

Умови праці та принципи гігієнічної класифікації поділяються на чотири класи.

Оптимальні умови праці (1-й клас) – умови, за яких зберігається не лише здоров'я працівників, а й створюються передумови для підтримання високого рівня працездатності. Оптимальні гігієнічні нормативи виробничих факторів встановлені для мікроклімату і факторів трудового процесу. Для інших факторів за оптимальні умовно приймаються такі умови праці, за яких несприятливі фактори виробничого середовища не перевищують рівнів, прийнятих за безпечні для населення.

Допустимі умови праці (2-й клас) – характеризуються такими рівнями шкідливих виробничих факторів виробничого середовища і трудового процесу, які не перевищують встановлених гігієнічних нормативів, а можливі зміни функціонального стану організму відновлюються за час регламентованого відпочинку чи до початку наступної зміни та не чинять

несприятливого впливу на стан здоров'я працівників та їх потомство в найближчому і віддаленому періодах.

Шкідливі умови праці (3-й клас) – характеризуються такими рівнями шкідливих виробничих факторів, які перевищують гігієнічні нормативи і здатні несприятливо впливати на організм працівника та на його нащадків. Шкідливі умови праці за ступенем перевищення гігієнічних нормативів і вираженості можливих змін в організмі працівників поділяються на чотири ступені.

Перший ступінь – умови праці характеризуються таким рівнем шкідливих факторів виробничого середовища і трудового процесу, які, як правило, зумовлюють функціональні зміни, що виходять за межі фізіологічних коливань (останні відновлюються при тривалішій, ніж початок наступної зміни, перерві контакту зі шкідливими факторами) та збільшують ризик погіршення здоров'я.

Другий ступінь – умови праці характеризуються такими рівнями шкідливих факторів виробничого середовища і трудового процесу, які здатні спричинювати стійкі функціональні порушення, призводять у більшості випадків до зростання виробничо-зумовленої захворюваності, появи окремих ознак або легких форм професійної патології (як правило, без втрати професійної працездатності), що виникають після тривалої експозиції (10 років і більше).

Третій ступінь – умови праці характеризуються такими рівнями шкідливих факторів виробничого середовища і трудового процесу, які призводять, окрім зростання виробничо-зумовленої захворюваності, до розвитку професійних захворювань, як правило, легкого і середнього ступенів важкості (з утратою професійної працездатності в період трудової діяльності).

Четвертий ступінь – умови праці характеризуються такими рівнями шкідливих факторів виробничого середовища і трудового процесу, які здатні

призводити до значного зростання хронічної патології та рівнів захворюваності з тимчасовою втратою працездатності, а також важких форм професійних захворювань (з утратою загальної працездатності).

Небезпечні (екстремальні) умови праці (4-й клас) – умови праці характеризуються такими рівнями шкідливих факторів виробничого середовища і трудового процесу, вплив яких протягом робочої зміни (чи її частини) створює загрозу для життя, високий ризик виникнення важких форм професійних уражень.

Поширена думка, що офісні робітники працюють в приємній, безпечній обстановці. Хоча працювати в офісі не так небезпечно, як у багатьох інших місцях, тут також існує ряд джерел ризику для життя і здоров'я. Деякі з них представляють для людей, які працюють в офісі серйозну небезпеку.

Для працівника несприятливими подіями внаслідок впливу умов праці є втома, стрес, захворювання (хвороба) або травма.

Втома – це фізіологічний стан організму, що виникає внаслідок надмірно інтенсивного або тривалої діяльності, що виявляється тимчасовим зниженням функціональних можливостей людського організму. Розрізняють фізичну, розумову і емоційну втому.

Фізична втома проявляється порушенням функції м'язів: зниженням сили, точності, узгодженості і ритмічності рухів; виникає при інтенсивній або тривалій фізичній діяльності.

Розумова втома проявляється зниженням продуктивності інтелектуальної праці, ослабленням уваги (труднощі зосередження), уповільненням мислення, зниженням показників розумової активності, інтересу до роботи; виникає при інтенсивній інтелектуальній діяльності.

Емоційна втома проявляється помітним зниженням емоційних реакцій під впливом надсильних або монотонних подразників (стресів).

Надмірне робоче навантаження протягом тривалого часу або недостатній час відпочинку може привести до хронічної втоми або перевтоми. Розрізняють розумовий і психічний (душевний) перевтома. В умовах сучасного ритму праці і життя все частіше у працівників з'являється синдром хронічної втоми.

Стрес є серйозною проблемою соціально-психологічного характеру для багатьох організацій. Він може бути викликаний багатьма факторами, включаючи шум від переповнення приміщення людьми або працюючого обладнання, поганих відносин з начальством та / або колегами, збільшенням робочого навантаження і недостатньою свободою дій при виконанні роботи.

Іншим масово поширеним несприятливим наслідком праці є те або інше захворювання: нездужання, погане самопочуття, симптоми, які можуть проходити бурхливо або відносно швидко минути («гострі» - з медичної термінології) або тривати (роками) або періодично загострюватися (хронічні).

Захворювання, спровоковані умовами праці, часто називають виробничозумовленими захворюваннями.

Специфічний вплив факторів, пов'язаний з конкретними виробничими чинниками, призводить до розвитку певних, викликаних цими факторами, захворювань. Оскільки вони викликані несприятливими умовами праці конкретних робочих місць конкретних професій, то їх називають професійними захворюваннями.

Захворювання кістково-м'язової системи та пошкодження м'яких тканин типу тендиніту виникають в результаті використання офісних меблів та обладнання, які не відповідають індивідуальним анатомічним особливостям. Тендинит може розвинути в результаті повторюваних рухів окремих частин тіла. Так, наприклад, у службовців починають боліти пальці від тривалого писання або при роботі з дуже товстими папками, які доводиться постійно діставати із шаф і повертати на місце. Багато офісних робітників страждають

від різноманітних RSI, типу захворювань зап'ястя і грудної клітини, через погану підгонку меблів і устаткування і відсутності перерв у друкуванні (на клавіатурі комп'ютера) або інших періодично повторюваних діях. Недосконала конструкція меблів і устаткування призводять також до порушення постави і здавлення нервів нижніх кінцівок, тому що багато офісних робітників велику частину часу 87 проводять сидячи. Всі ці фактори сприяють виникненню захворювань попереку і нижніх кінцівок в такій же мірі, як і тривале стояння на ногах.

Безперервне використання комп'ютерів і погане загальне освітлення викликає напругу зору, в результаті чого воно поступово погіршується, з'являються головні болі, печіння і відчуття втоми в очах. Регулювання рівня освітленості приміщення і контрастності зображення на екрані комп'ютера, а також часта перефокусування очей необхідні для попередження виникнення проблем із зором. Освітлення має відповідати характеру виконуваної роботи.

5.3 Розрахунок рівня освітлення в офісі фірми

Рівень освітленості впливає на працездатність і самопочуття людини. Державні норми це враховують і визначають, яким має бути освітлення приміщень, залежно від їх призначення. Відштовхуючись від цих норм, планують освітлення об'єктів, а отже, від них безпосередньо залежить і те, які світильники і скільки лампочок необхідно купити.

Стосовно ДБН (Державно будівельні норми) є конкретні показники якості освітлення приміщень різних призначень: житлових, офісних, складських, виробничих. Спираючись на ці дані, розраховують рівень освітленості, враховуючи площу приміщення, висоту стелі, відбиваючу здатність поверхонь, тип світильника та інші параметри. Стосовно

необхідного нам показника в офісних приміщеннях, нормою освітлення буде 300 Лк.

Норми виникли після того як були враховані всі найбільш важливі особливості фізіологічних зорових процесів людини і фактори, які впливають на концентрацію, зорову напругу під час роботи, відпочинку (наприклад, контрастність, колір робочої поверхні).

Для розрахунку рівня освітленості приміщення зазвичай використовується спосіб, який передбачає використання даних про величину світлового потоку, який вимірюється люменами (Лм). Світловий потік, що випромінює лампа, який вказується на упаковці придбаної лампи.

Для визначення величини світлового потоку використовується формула:

Норма освітлення * Площа * Коефіцієнт висоти стелі

При висоті стелі від 2,5 м до 2,7 м коефіцієнт дорівнює 1, при висоті 2,7-3 метрів - 1,2, якщо висота становить від 3 до 3,5 метрів - 1,5, від 3,5 м до 4,5 м - 2. Порахуємо, яка величина світлового потоку необхідна вітальні площею 30 квадратних метрів з висотою стель 2,5 м:

$$150 * 30 * 1 = 4500 \text{ (Лм)}.$$

Оскільки різні лампочки випромінюють світло різної сабо, потрібна їх кількість залежить від того, що ми виберемо – світлодіодну, енергозберігаючу, галогенну або лампу розжарювання, а також від того, якої потужності вона буде.

Таблиця 5.1 – Приклад різних видів ламп і їх потужності

Лампочка	Потужність, Вт	Світловий потік, Лм	Потрібно лампочок, шт
Розжарювання	40	470	4500/470=9,5
Галогенна	32	470	4500/470=9,5
Енергозберігаюча	15	700	4500/700=6,4
Світлодіодна	10	750	4500/750=6

Отже, для освітлення вітальні площею 30 квадратних метрів необхідно 10 галогенних ламп потужністю 30 Вт, стільки ж більш потужних лампочок розжарювання, 7 енергозберігаючих по 15 Вт кожна або 6 світлодіодних 10-89 ватних лампочок. Після підрахунків підсумок з десятковими знаками завжди округляємо в більшу сторону.

Більш складний спосіб розрахунку включає більше змінних:

$N = (E * S) / (U * p * Fi * Kz)$, де N – кількість світильників, E – норма освітленості, S – площа об'єкта, Kz – коефіцієнт запасу (залежить від виду лампочки, ступеня забруднення приміщення), U – коефіцієнт використання світлового потоку (залежить від висоти стін, площі кімнати, типу світильника, кольору покриття стелі, стін, підлоги), p – кількість патронів світильника, Fi – значення світлового потоку однієї лампи.

Наприклад, для вітальні з першого прикладу планується встановити точкові світильники. Колір покриття стелі - білий, стіни світлі, на підлозі - темно-коричневий ламінат. Для об'єкта показник E становить 150 (дивимося в таблиці), S – 30, Kz – 1 (припускаємо, що купують світлодіодні лампи потужністю 10 Вт), U – 0,46, p – 1, Fi – 750.

Підставляємо дані:

$$N = (150 * 30) / (0,46 * 1 * 750 * 1) = 4500 / 345 = 13.$$

Виходячи зі розрахованої формули було знайдено оптимальне значення кількості світильників, тобто слід розмістити 13 світильників.

5.4 Розробка заходів щодо зменшення впливу шкідливих та небезпечних факторів

З метою запобігання або зменшення впливу шкідливих і небезпечних виробничих чинників під час підготовки робочих місць, або робочої зони, необхідно слідувати загальним вимогам, що встановлені законодавством України про працю.

Відповідно до законів України «Про охорону праці» умови праці на робочому місці, безпека технологічних процесів, устаткування та інших засобів виробництва, стан засобів колективного та індивідуального захисту, що використовуються працівником, а також санітарно-побутові умови повинні відповідати вимогам законодавства.

Більшість нормативів щодо умов праці офісних працівників встановлено на рівні державних стандартів. Відповідно до ч. 1 ст. 13 Закону про охорону праці роботодавець зобов'язаний створити на робочому місці в кожному структурному підрозділі умови праці відповідно до нормативно-правових актів.

Як відомо, тривала робота за комп'ютером та з документами при недостатньому рівні освітленості може призвести до значного перенапруження зору, тому вимоги до освітлення є досить важливими. Додатково, окрім вже перелічених документів, вимоги до освітлення встановлено ДБН В.2.5-28-2006 «Природне і штучне освітлення», затвердженими наказом Мінрегіону від 15.05.2006 р. № 168.

Як вже зазначалося, відносно вікон робоче місце необхідно організовувати так, щоб природне світло було з лівого боку (п. 4.3 ДСанПіН 3.3.2.007-98). Робоче місце необхідно розміщувати таким чином, щоб уникнути попадання прямого світла в очі. Для забезпечення захисту і досягнення нормованих рівнів комп'ютерних випромінювань необхідне

застосування приєкраних фільтрів, локальних світлофільтрів (засобів індивідуального захисту очей) та інших засобів захисту, що пройшли випробування в акредитованих лабораторіях і мають щорічний гігієнічний сертифікат (п. 4.19 ДСанПіН 3.3.2.007-98).

Штучне освітлення приміщення має здійснюватись системою загального рівномірного освітлення (п. 3.2.2 ДСанПіН 3.3.2.007-98). У приміщеннях при переважній роботі з документами допускається використання системи комбінованого освітлення, тобто встановлення світильників місцевого освітлення додатково до загального.

Як джерела штучного освітлення необхідно використовувати люмінесцентні лампи. Згідно з п. 3.2.5 ДСанПіН 3.3.2.007-98 система загального освітлення має бути у вигляді суцільних або переривчатих ліній світильників, що розташовані збоку від робочих місць (зазвичай ліворуч) паралельно лінії зору працівників.

Допускається застосування ламп розжарювання у світильниках місцевого освітлення та, у разі влаштування відбитого освітлення у виробничих чи адміністративно-громадських приміщеннях, металогалогенних ламп потужністю 250 Вт.

Коефіцієнт пульсації не повинен перевищувати 5 % (п. 3.2.14 ДСанПіН 3.3.2.007-98). Рівень освітленості на робочому столі в зоні розташування документів має бути в межах 300 – 500 лк. Світильники місцевого освітлення слід встановлювати таким чином, щоб не створювати відблисків на поверхні екрана, а освітленість екрана має не перевищувати 300 лк.

Для забезпечення нормованих значень освітленості у приміщеннях відповідно до п. 3.2.15 ДСанПіН 3.3.2.007-98 необхідно мити вікна і світильники не рідше 2 разів на рік, а також своєчасно замінювати лампи, що перегоріли.

6 ОХОРОНА НАВКОЛИШНЬОГО СЕРЕДОВИЩА

Охорона навколишнього природного середовища, раціональне використання природних ресурсів, забезпечення екологічної безпеки життєдіяльності людини – невід’ємна умова сталого економічного та соціального розвитку України.

З цією метою Україна здійснює на своїй території екологічну політику, спрямовану на збереження безпечного для існування живої і неживої природи навколишнього середовища, захисту життя і здоров’я населення від негативного впливу, зумовленого забруднення навколишнього природного середовища, досягнення гармонійної взаємодії суспільства і природи, охорону, раціональне використання і відтворення природних ресурсів.

Відносини у галузі охорони навколишнього природного середовища в Україні регулюється Законом України № 1268-ХІІ від 26.06.91р. «Про охорону навколишнього природного середовища», а також розроблюваними відповідно до нього земельним, водним, лісовим законодавством, законодавством про надра, про охорону атмосферного повітря, про охорону і використання рослинного і тваринного світу та іншим спеціальним законодавством.

Цей Закон визначає правові, економічні та соціальні основи організації охорони навколишнього природного середовища в інтересах нинішнього і майбутніх поколінь.

Завданням законодавства про охорону навколишнього природного середовища є регулювання відносин у галузі охорони, використання і відтворення природних ресурсів, забезпечення екологічної безпеки, запобігання і ліквідації негативного впливу господарської та іншої діяльності

та навколишнє природне середовище, збереження природних ресурсів, генетичного фонду живої природи, ландшафтів та інших природних комплексів, унікальних територій та природних об'єктів, пов'язаних з історико-культурною спадщиною.

Охорона навколишнього природного середовища є системою державних і суспільних заходів, направлених на збереження, відтворення і раціональне використання природних ресурсів і поліпшення полягання природного середовища, і є частиною прикладної екології.

В міру прискорення темпів науково-технічного прогресу дія людей на природу стає все більш могутньою. І в даний час воно все сумирно із дією природних чинників, що приводить до якісної зміни співвідношення сил між суспільством і природою. На сучасному етапі людство поставлено перед чинником виникнення в природі незворотних процесів, нових шляхів переміщення і перетворення енергії і речовини. В природу втручається все більше і більше нових речовин, чужих їй, часом сильно токсичними для організмів. Частина з них не включається в природний круговорот і нагромаджується в біосфері, що приводить до небажаних екологічних наслідків.

6.1 Забруднення навколишнього середовища в процесі використання комп'ютерної техніки

Основну загрозу для навколишнього середовища в межах галузі розробки ПЗ ставлять ПК, а саме електромагнітні випромінювання від них.

Дослідження останніх років показано, що електромагнітні випромінювання вищезгаданих електричних пристроїв містять торсіонову

компоненту, котра несе інформацію про процеси, що відбуваються в тому чи іншому електронному пристрої.

Торсіонові поля мають високу проникаючу здатність і їх неможливо заекранувати.

Останніми роками при проектування, виготовленні і експлуатації технічних і систем управління в різних сферах діяльності надзвичайно широко застосовуються персональні комп'ютери і всілякі комп'ютерні програми.

Робота з комп'ютерами програмістів, операторів і інших користувачів пов'язана з додатковими шкідниками і небезпечними чинниками, що негативно впливають на організм людини. Наприклад, шкідлива дія на зір надає не притримання стандартних візуальних ергономічних параметрів екрану, розмірів мінімального елементу відображення, мерехтіння зображення, відбивна здатність і ін. Працюючий комп'ютер створює електромагнітне поле, що шкідливо діє на організм людини. Це поле може виникати радіоперешкоди, тобто завантажити роботі радіо- і телеапаратури, що призводить до зниження надійності технічної системи або системи управління, до збільшення ризику виникнення аварійної ситуації і виробничому середовищі. Для забезпечення безпеки роботи з комп'ютером розроблені і повинні повсюди застосовуватися стандарти на монітори, вимоги до приміщення для експлуатації комп'ютерів і до організації і устаткування робочих місць.

6.2 Використання програмного забезпечення для вирішення екологічних проблем

Широко використовуються фахівцями-екологами програми розрахунку забруднення атмосфери та інші спеціалізовані екологічні програми на рівні

окремого підприємства досить повно вирішують завдання, що стоять в цій області.

Розроблена фірмою «Інтеграл» комп'ютерна система розрахункового моніторингу стану якості повітряного басейну міста (регіону) «Еколог-Місто» призначена для автоматизації діяльності територіальних природоохоронних органів, екологічних служб адміністрацій міст (регіонів), проведення зведених розрахунків забруднення атмосфери міст.

Результати зведених розрахунків забруднення атмосфери можуть бути використані з метою нормування викидів забруднюючих речовин і для вирішення інших завдань.

Система «Еколог-місто» складається з двох підсистем:

- підсистеми ведення банку даних;
- підсистеми отримання попередньої оцінки якості атмосферного повітря.

Взаємодія між підсистемами здійснюється як на рівні обміну інформацією, так і на рівні передачі управління.

Основним завданням підсистеми ведення банку даних є підготовка інформації про викиди шкідливих речовин в атмосферу та інших даних, необхідних для коректної роботи підсистеми отримання оцінок забруднення атмосфери.

Основними особливостями файлів даних, записаних в розробленому форматі є:

- 1) універсальність з точки зору можливості використання різних програмних засобів при створенні і перетворенні файлів. Текстові файли даних можуть бути створені і заповнені або за допомогою програм системи «Екологмісто» або за допомогою будь-якого текстового редактора;
- 2) сегментування інформації в файлах.

Кожен файл даних складається з секцій, кожна з яких містить певний вид інформації. Інформація в секціях різниться:

- по виду описуваних нею величин і властивостей.
- за рівнем описуваних нею об'єктів в ієрархії об'єктів, що розглядаються під час вирішення завдань оцінки забруднення атмосфери міста (регіону).

Використовується інформація наступних видів:

- про параметри джерел забруднення атмосфери;
- про інвентаризацію викидів забруднюючих речовин;
- про умови розсіювання домішок в атмосфері;
- про використовувані системах координат;
- про місцезнаходження на місцевості зон зі специфічними вимогами до якості атмосферного повітря;
- про узагальнених характеристиках міста, його районів, підприємств, їх майданчиків і цехів.

За рівнем описуваних об'єктів інформація сегментована на розділи, що містять опис:

- міста (регіону);
- району міста (регіону);
- підприємства;
- майданчики і цеху підприємства;
- джерела забруднення атмосфери, що належить конкретній ділянці (можливо, також і цеху) підприємства.

Підсистема ведення банку даних складається з чотирьох самостійних великих блоків, взаємодія між якими здійснюється на рівні обміну інформацією:

- блоку підготовки вихідної інформації;
- блоку ведення міського банку даних;

- блоків перетворення даних про джерела викидів з форматів, використовуваних в програмі УПРЗА "Еколог" та інших програмах в єдиний формат даних системи "Еколог-Місто";

Блок підготовки вихідної інформації надає користувачеві наступні можливості:

- 1) занести дані про джерела викидів підприємства або прийняти дані про джерела викидів підприємства в текстовому форматі з файлу на дискеті або на жорсткому диску;
- 2) провести перевірку коректності занесених даних;
- 3) записати занесені дані в текстовий формат для надання їх в цьому форматі на дискеті в територіальні органи охорони навколишнього середовища;
- 4) блок ведення міського банку даних дозволяє:
- 5) прийняти інформацію про джерела викидів підприємств з файлу в єдиному форматі, INT, або підготувати її безпосередньо;
- 6) переглянути інформацію у вигляді таблиць і при необхідності її доповнити;
- 7) переглянути взаємне розташування джерел на тлі топооснови місцевості як для окремого підприємства, так і для сукупності підприємств.

ВИСНОВКИ

В кваліфікаційній роботі вирішене наукове завдання з підвищення достовірності оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, за допомогою нелінійної регресійної моделі.

1) Здійснено аналіз існуючих регресійних рівняння для оцінювання розміру web застосунків. З'ясовано, що на сьогоднішній день існують лінійні регресійні рівняння для оцінювання розміру web застосунків. Проте існують чотири основні припущення, які виправдовують використання моделей лінійної регресії, одним з яких є нормальність розподілу помилок. Але це припущення справедливе лише в окремих випадках. Це призводить до необхідності використання нелінійних регресійних моделей для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

2) Здійснено збір набору даних, за яким будується нелінійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, та перевірку цих даних на наявність багатовимірних викідів. З'ясовано, застосування критерія на основі квадрата Махаланобіса до підібраних даних без їх нормалізації показало, що дані містять багатомірний викид.

3) Отримані результати свідчать про те, що нелійна регресійна модель для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp, має високу якість оцінювання. Ця нелінійна модель в порівнянні з лінійною регресійною моделлю має більші значення множинного коефіцієнту детермінації R^2 і відсотку прогнозованих результатів $PRED(0,25)$, менше значення середньої величини відносної

похибки MMRE та менші ширини довірчого інтервалу і інтервалу передбачення для більшої кількості значень незалежних змінних.

4) На підставі запропонованої нелінійної регресійної моделі розроблено програму для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp. Була розроблена наступна програмна документація: технічне завдання, текст програми, опис програми, інструкція користувача, та методика випробувань, що наведені відповідно у додатках А, Б, В, Г, Д. Були проведені випробування програми згідно з програмою та методикою випробувань, що наведена у додатку Д. За результатами випробувань зроблено висновок про працездатність розробленої програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Tan H.B.K. Estimating LOC for information systems from their conceptual data models / H. B. K. Tan, Y. Zhao, H. Zhang // Proceedings of the 28th International Conference on Software Engineering (ICSE '06), Shanghai, China, May 20-28, 2006. – P. 321-330.
2. Prykhodko S. Estimating the Software Size of Open-Source PHP-Based Systems Using Non-Linear Regression Analysis / Sergiy Prykhodko, Natalia Prykhodko, Lidiia Makarova // Proceedings of International Conference “Advanced Computer Information Technologies” (ACIT-2018). CEUR Workshop Proceedings. – 2019. – Vol.2300. – Ceske Budejovice, CZECH REPUBLIC. CEUR-WS.org – P. 199-202. ISSN 1613-0073.
3. Приходько С.Б. Нелінійна регресійна модель для оцінювання розміру Web-застосунків, що створюються з використанням фреймворку Laravel / С. Б. Приходько, Н. В. Приходько, М. В. Ворона, І. О. Беловол // Інформаційні технології та комп'ютерна інженерія. – 2021. – Т.50. – № 1. – С. 115-121.
4. Приходько С.Б. Трьохфакторна нелінійна регресійна модель для оцінювання розміру Php-застосунків з відкритим кодом. [Текст] / С. Б. Приходько, Н. В. Приходько, Т. А. Фаріонова, М. В Ворона. // Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки. – Київ, 2020. – № 4 (56), – С. 124-131.
5. Ворона М.В. Трьохфакторні нелінійні регресійні моделі для оцінювання розміру web-застосунків, що створюються з використанням php-фреймворків. [Текст] / М. В. Ворона, А. С. Приходько, І. С. Шутко // Інформатика, інформаційні системи та технології: тези доповідей вісімнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 23 квітня 2021 р. – Одеса, 2021. – С. 111-113.

6. Приходько С.Б. Нелінійні регресійні моделі для оцінювання кількості строк коду web застосунків, що створюються за допомогою РНР-фреймворків / С. Б. Приходько, А. С. Приходько, І. С. Шутко // Інформаційні технології: моделі, алгоритми, системи: Матеріали II Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2021 року). – Миколаїв: НУК імені адмірала Макарова, 2021. – С. 11-12.

7. Приходько С.Б. Нелінійні регресійні моделі для оцінювання розміру web-застосунків, що створюються за допомогою РНР-фреймворків / С. Б. Приходько, А. С. Приходько, І. С. Шутко // Проблеми моделювання та розроблення інформаційних систем : матеріали V науково-практичної інтернет-конференції (Дрогобич, 23 квітня 2021 року). – Дрогобич : ДДПУ ім. І. Франка, 2021. – С. 11-12.

8. Boehm B. W. Software engineering economics [Text] / B. W. Boehm. – Englewood Cliffs, NJ : Prentice Hall, 1981. – 768 p.

9. Trendowicz A. Software project effort estimation. foundations and best practice guidelines for success / A. Trendowicz, R. Jeffery. – Springer International Publishing, 2014. – 491 p. DOI: <https://doi.org/10.1007/978-3-319-03629-8>.

10. Tan H.B.K. Conceptual data model-based software size estimation for information systems / H. B. K. Tan, Y. Zhao, H. Zhang // Transactions on Software Engineering and Methodology. – 2009. – Vol. 19. – Issue 2. – October 2009. – Article No. 4.

11. Prykhodko N.V. Constructing the non-linear regression models on the basis of multivariate normalizing transformations / N. V. Prykhodko, S. B. Prykhodko // Electronic modeling. – 2018. – T.40. – № 6. – С. 99-108.

12. Montgomery D.C. Introduction to Linear Regression Analysis /Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining. – Wiley, 2021. – 702 p.

13. Fahrmeir L. Regression: Models, Methods and Applications / Ludwig Fahrmeir, Thomas Kneib, Stefan Lang, Brian Marx. – Springer, 2013. – 712 p.

14. Derek S. Young. Handbook of Regression Methods / S. Derek Young. – Chapman and Hall, 2017. – 654 p.
15. Härdle W.K. Regression Models / W. K. Härdle, L. Simar // In: Applied Multivariate Statistical Analysis. – Springer, Berlin, Heidelberg, 2015. DOI: https://doi.org/10.1007/978-3-662-45171-7_8
16. Johnson R.A. Applied Multivariate Statistical Analysis / R. A. Johnson, D. W. Wichern. – Pearson Prentice Hall, 2007. – 800 p.
17. Chatterjee S. Handbook of Regression Analysis / S. Chatterjee, J. S. Simonoff. – New York: John Wiley & Sons, 2013. – 252 p.
18. Prykhodko S.B. Developing the software defect prediction models using regression analysis based on normalizing transformations [Text] / S. B. Prykhodko // Modern Problems in Testing of the Applied Software : the Research and Practice Seminar (PTASS-2016), Poltava, Ukraine, May 25-26, 2016 : abstracts. – P. 6-7.
19. Prykhodko S.B. Building the equations, confidence and prediction intervals of non-linear regressions on the basis of multivariate normalizing transformations / S. B. Prykhodko, N. V. Prykhodko, K. S. Pugachenko // Прикладні науково-технічні дослідження: матеріали II міжнар. наук.-практ. конф., 3-5 квіт. 2018 р. – Академія технічних наук України. – Івано-Франківськ: Симфонія форте, 2018. – С. 16.
20. Prykhodko S. Building the Non-Linear Regression Equations on the Basis of Multivariate Normalizing Transformations / Sergiy Prykhodko, Natalia Prykhodko, Lidiia Makarova // Proceedings of the 2018 IEEE First International Conference on SYSTEM ANALYSIS & INTELLIGENT COMPUTING (SAIC), 08-12 October, 2018, Kyiv, Ukraine. – P. 48-52. DOI: <https://doi.org/10.1109/SAIC.2018.8516742>
21. Prykhodko S. Estimating the Software Size of Open-Source PHP-Based Systems Using Non-Linear Regression Analysis / Sergiy Prykhodko, Natalia Prykhodko, Lidiia Makarova // Proceedings of International Conference

“Advanced Computer Information Technologies” (ACIT-2018). CEUR Workshop Proceedings. – 2019. – Vol.2300. – Ceske Budejovice, CZECH REPUBLIC. CEUR-WS.org – P. 199-202.

22. Приходько С.Б. Побудова рівняння нелінійної регресії для оцінювання розміру програмного забезпечення РНР-систем з відкритим кодом / С. Б. Приходько, О. В. Спінов // Прикладні науково-технічні дослідження: матеріали II міжнар. наук.-практ. конф., 3-5 квіт. 2018 р. – Академія технічних наук України. – Івано-Франківськ: Симфонія форте, 2018. – С. 36.

23. Спінов О.В. Оцінювання розміру програмного забезпечення РНР-систем з відкритим кодом за рівнянням нелінійної регресії / О. В. Спінов, С. Б. Приходько // Інформатика, інформаційні системи та технології: тези доповідей п'ятнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 27 квітня 2018 р. – Одеса, 2018. – С. 25-26.

24. Prykhodko S.B. Constructing the non-linear regression equation to estimate the software size of open source PHP-based information systems / S. B. Prykhodko, N. V. Prykhodko, T. G. Smykodub, A. V. Spinov // Проблемы информационных технологий. – 2018. – № 1 (023). – С. 118-125.

25. Mardia K.V. Measures of multivariate skewness and kurtosis with applications / K. V. Mardia // Biometrika. – 57. – 1970. – P. 519–530. DOI: <https://doi.org/10.1093/biomet/57.3.519>

26. Chatterjee S., Price B. Regression Analysis by Example. – New York: John Wiley & Son, 1977. – 228 p.

27. Prykhodko S. Detecting Outliers in Multivariate Non-Gaussian Data on the basis of Normalizing Transformations / S. Prykhodko, N. Prykhodko, L. Makarova, K. Pugachenko // Proceedings of the 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON) «Celebrating 25 Years of IEEE Ukraine Section», Kyiv, Ukraine, May 29 – June 2, 2017. – P. 846-849.

28. Prykhodko S. Application of the Squared Mahalanobis Distance for Detecting Outliers in Multivariate Non-Gaussian Data / S. Prykhodko, N. Prykhodko, L. Makarova, A. Pukhalevych // Proceedings of 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, February 20–24, 2018. – P. 962-965.

29. Методичні вказівки та завдання до виконання лабораторних робіт з дисципліни "Обробка експериментальних даних на комп'ютері" / С. Б. Приходько, Л. М. Макарова, К. С. Пугаченко. – Миколаїв: НУК, 2018. – 76 с.

30. Foss T. A simulation study of the model evaluation criterion MMRE / T. Foss, E. Stensrud, B. Kitchenham, I. Myrtveit // IEEE Transactions on software engineering, 11 (29), 2003. – P. 985–995.

31. Prykhodko S. Mathematical Modeling of Non-Gaussian Dependent Random Variables by Nonlinear Regression Models Based on the Multivariate Normalizing Transformations / S. Prykhodko, N. Prykhodko // In: Shkarlet S., Morozov A., Palagin A. (eds) Mathematical Modeling and Simulation of Systems (MODS'2020). MODS 2020. Advances in Intelligent Systems and Computing, Vol 1265, 2021. – P. 166-174. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-58124-4_16

32. Приходько С.Б. Метод покращення нелінійних регресійних моделей на основі багатовимірних нормалізуючих перетворень / С. Б. Приходько, Н. В. Приходько // Прикладні науково-технічні дослідження: матеріали III міжнар. наук.-практ. конф. (Івано-Франківськ, 3–5 квітня 2019 р.). – Івано-Франківськ: Сімфонія Форте, 2019. – С. 20.

ДОДАТОК А – ТЕХНІЧНЕ ЗАВДАННЯ

Вступ

Назва розробки – “Програма для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp ”, далі програма.

Підстави для розробки

Завдання на кваліфікаційну роботу.

Призначення розробки:

Програма призначена для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Вимоги до програми:

Вимоги функціональних характеристик.

Вимоги до складу функцій, що виконуються.

Програма має виконувати такі функції:

1) Вводити з інтерфейсу програми довірчу ймовірність, три метрики діаграми класів, за якими потрібно оцінити розмір ПЗ web застосунків, що створюються з використанням фреймворку CakePhp в тисячах строк коду: перша X_1 , друга X_2 та третя X_3 метрики визначають відповідно загальну кількість класів, середню кількість методів на кожен клас та глибину дерева успадкування DIT.

2) Виконувати обчислення розміру web застосунка, що створюється з використанням фреймворку CakePhp.

3) Виконувати обчислення довірчого інтервалу нелінійної регресії.

4) Виконувати обчислення інтервалу передбачення нелінійної регресії.

Вимоги до організації вхідних та вихідних даних

Вхідні дані:

- три метрики діаграми класів, за якими потрібно оцінити розмір web застосунку, що створюється з використанням фреймворку CakePhp: загальну кількість класів, загальну кількість зв'язків між класами та глибину дерева успадкування DIT.

- значення довірчої ймовірності, трьох метрик діаграми класів, за якими потрібно оцінити розмір web застосунку, що створюється з використанням фреймворку CakePhp: загальну кількість класів, середню кількість методів на кожен клас та глибину дерева успадкування DIT.

Вихідні дані:

- значення нелінійної регресії та границь довірчого інтервалу нелінійної регресії розміру web застосунку, що створюється з використанням фреймворку CakePhp.

- значення коефіцієнту детермінації.

Вимоги до надійності не висуваються.

Вимоги до умов експлуатації не висуваються.

Вимоги до складу й параметрів технічних засобів.

Програма має експлуатуватись на обладнанні з параметрами не нижчими за наступні: 64-бітний процесор на архітектурі x86 або ARM, оперативна пам'ять 4Гб (2 Гб мінімум), простір на жорсткому диску 100 Мб мінімум.

Вимоги до інформаційної та програмної сумісності.

Для запуску та роботи з програмним забезпеченням необхідна операційна система Windows 8/10/11, Ubuntu 18/19/20 або Mac OS 10.12/10.13/10.14.

Вимоги до маркування та пакування.

Визначаються вимогами до пакування CD дисків.

Вимоги до транспортування та зберігання.

Програма може зберігатись та транспортуватися на CD дисках.

Вимоги до програмної документації.

Програмна документація має включати: технічне завдання на розробку програми, текст програми, опис програми, інструкція користувача, програма та методика випробувань.

Техніко-економічні показники не розраховуються.

Стадії та етапи розробки.

Стадії та етапи розробки програми представлені у таблиці.

Таблиця – Стадії та етапи розробки

Стадія розробки	Етапи робіт	Термін виконання
1. Технічне завдання	Розробка та затвердження технічного завдання	21.09.2021
2. Ескізний проект	Розробка ескізного проекту	24.09.2021
	Затвердження ескізного проекту	26.09.2021
3. Технічний проект	Розробка технічного проекту	08.10.2021
	Затвердження технічного проекту	10.10.2021
4. Робочий проект	Розробка програми	05.11.2021
	Розробка програмної документації	12.11.2021
	Випробування програми	21.11.2021

Порядок контролю і приймання

Порядок контролю та приймання програми здійснюється представниками замовника у присутності представника розробника.

Для приймання програми надається програмна документація, яка була розроблена відповідно до технічного завдання. Представники замовника установлюють відповідність програми технічному завданню.

За результатами приймання програми представниками замовника

складається акт приймання програми.

ДОДАТОК Б – ТЕКСТ ПРОГРАМИ

```

import sys
import math
import numpy as np
import scipy as sp
from scipy import stats

from PySide2 import QtCore, QtWidgets, QtGui

intercept = -1.8048396509693518
coeffs = np.array([0.98186532, 1.18673812, 0.31978913])
N = 32
MSE = 0.0009806190154350383
ZX = np.array([
    [0.07183667, 0.03496394, 0.20641009],
    [0.03496394, 1.40223534, -0.12683976],
    [0.20641009, -0.12683976, 5.45787676]])
Z_MEAN = [1.355858, 0.605359, 0.229471]

class Window(QtWidgets.QMainWindow):
    def __init__(self):
        super().__init__()

        input_group_box = QtWidgets.QGroupBox(self)
        result_group_box = QtWidgets.QGroupBox(self)

        self.methods_edit = QtWidgets.QLineEdit("0", self)
        self.class_edit = QtWidgets.QLineEdit("0", self)
        self.dit_edit = QtWidgets.QLineEdit("0", self)
        self.confidence_edit = QtWidgets.QLineEdit("0,95", self)
        self.project_size_edit = QtWidgets.QLineEdit("0", self)

        self.methods_edit.setValidator(QtGui.QDoubleValidator(self))
        self.class_edit.setValidator(QtGui.QIntValidator(self))
        self.project_size_edit.setValidator(QtGui.QDoubleValidator(self))
        self.dit_edit.setValidator(QtGui.QDoubleValidator(self))
        self.confidence_edit.setValidator(QtGui.QDoubleValidator(0.0, 1.0, 2,
parent=self))

        self.methods_edit.textChanged.connect(self.calculate)
        self.class_edit.textChanged.connect(self.calculate)
        self.project_size_edit.textChanged.connect(self.calculate)
        self.dit_edit.textChanged.connect(self.calculate)
        self.confidence_edit.textChanged.connect(self.calculate)

        central_widget = QtWidgets.QWidget()
        self.setCentralWidget(central_widget)

        layout = QtWidgets.QVBoxLayout(central_widget)
        layout.addWidget(input_group_box)
        layout.addWidget(result_group_box)

        layout = QtWidgets.QFormLayout(input_group_box)

```

```

layout.addRow("Class count", self.class_edit)
layout.addRow("Avg methods count", self.methods_edit)
layout.addRow("Depth of inheritance tree", self.dit_edit)
layout.addRow("Confidence level", self.confidence_edit)
layout.addRow("Actual size", self.project_size_edit)

self.kloc = QtWidgets.QLineEdit(self)
self.kloc.setReadOnly(True)
self.kloc.setMinimumWidth(150)

self.confidence_interval = QtWidgets.QLineEdit(self)
self.confidence_interval.setReadOnly(True)
self.confidence_interval.setMinimumWidth(150)

self.predicition_interval = QtWidgets.QLineEdit(self)
self.predicition_interval.setReadOnly(True)
self.predicition_interval.setMinimumWidth(150)

self.mre = QtWidgets.QLineEdit(self)
self.mre.setReadOnly(True)
self.mre.setMinimumWidth(150)

layout = QtWidgets.QFormLayout(result_group_box)
layout.addRow("KLOC", self.kloc)
layout.addRow("Confidence interval", self.confidence_interval)
layout.addRow("Prediction interval", self.predicition_interval)
layout.addRow("MRE", self.mre)

def calculate(self):
    vector = []
    for edit in [self.class_edit, self.methods_edit, self.dit_edit]:
        if edit.text() and edit.hasAcceptableInput():
            vector.append(float(edit.text().replace(",", ".")))
        else:
            return

    value = (10 ** intercept) * (vector[0] ** coeffs[0]) * (vector[1] **
coeffs[1]) * (vector[2] ** coeffs[2])

    zx = np.log10(vector) - Z_MEAN
    se_z = np.matmul(np.matmul(zx.T, ZX), zx)

    conf_level = float(self.confidence_edit.text().replace(",", "."))

    pred_val = (sp.stats.t.ppf(1 - (1 - conf_level) / 2, N - 4) *
math.sqrt(MSE * (1 + 1.0 / N + se_z)))
    conf_val = (sp.stats.t.ppf(1 - (1 - conf_level) / 2, N - 4) *
math.sqrt(MSE * (1.0 / N + se_z)))

    pred_interval = (10 ** (math.log10(value) - pred_val), 10 **
(math.log10(value) + pred_val))
    conf_interval = (10 ** (math.log10(value) - conf_val), 10 **
(math.log10(value) + conf_val))

    self.kloc.setText('%.4f' % value)
    self.predicition_interval.setText('%.4f - %.4f' % pred_interval)
    self.confidence_interval.setText('%.4f - %.4f' % conf_interval)

```



```
        actual_size = float(self.project_size_edit.text().replace(",", "."))
        mre = abs(actual_size - value) / value * 100

        self.mre.setText('%.2f%%' % mre)

if __name__ == "__main__":
    app = QtWidgets.QApplication([])

    widget = Window()
    widget.resize(400, 300)
    widget.show()

    sys.exit(app.exec_())
```

ДОДАТОК В – ОПИС ПРОГРАМИ

Загальні відомості

Програма для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Функціональне призначення

Програма CakePhpSizeEst призначена оцінювання для розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Технічні засоби, що використовуються

Програма має експлуатуватись на обладнанні з параметрами не нижчими за наступні: 64-бітний процесор на архітектурі x86 або ARM, оперативна пам'ять 4Гб (2 Гб мінімум), простір на жорсткому диску 100 Мб мінімум.

Виклик та завантаження

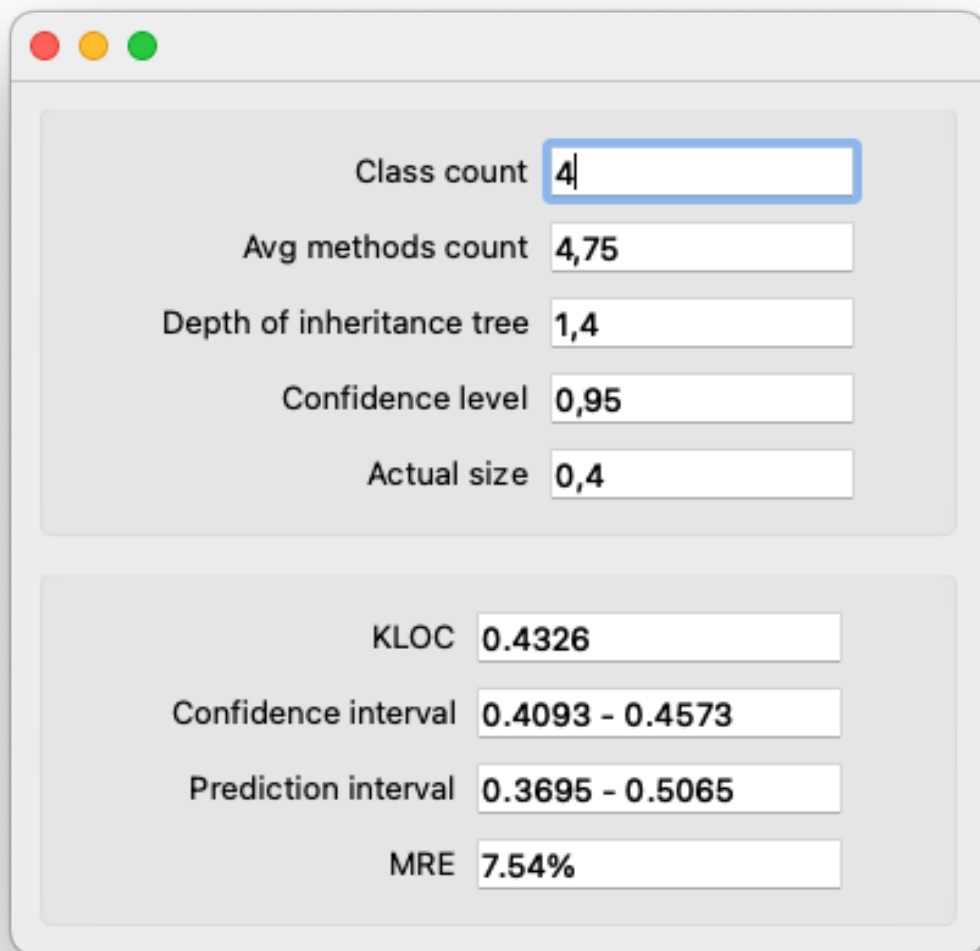
Для запуску та роботи з програмним забезпеченням необхідна операційна система Windows 8/10/11, Ubuntu 18/19/20 або Mac OS 10.12/10.13/10.14.

Завантаження програми здійснюється шляхом відкриття файлу CakePhpSizeEst.py.

ДОДАТОК Г – ІНСТРУКЦІЯ КОРИСТУВАЧА

Загальний опис роботи з ПЗ

Після завантаження файлу CakePhpSizeEst.py на екрані з'являється форма для відображення початкового стану програми для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp (рисунок Г.1).



The screenshot displays a graphical user interface for the CakePhpSizeEst.py application. It features two main sections with input fields and calculated metrics. The top section contains five input fields: 'Class count' (value: 4), 'Avg methods count' (value: 4,75), 'Depth of inheritance tree' (value: 1,4), 'Confidence level' (value: 0,95), and 'Actual size' (value: 0,4). The bottom section displays four calculated metrics: 'KLOC' (0.4326), 'Confidence interval' (0.4093 - 0.4573), 'Prediction interval' (0.3695 - 0.5065), and 'MRE' (7.54%).

Parameter	Value
Class count	4
Avg methods count	4,75
Depth of inheritance tree	1,4
Confidence level	0,95
Actual size	0,4
KLOC	0.4326
Confidence interval	0.4093 - 0.4573
Prediction interval	0.3695 - 0.5065
MRE	7.54%

Рисунок Г.1 – Форма для відображення початкового стану програми

У графічному вікні (рис. Г.1) у відповідних вікнах редагування потрібно відредагувати: загальну кількість класів, загальну кількість методів на кожен клас, глибину дерева успадкування, довірчу ймовірність.

Крім того у графічному вікні (рис. Г.1) за бажанням можна ввести фактичний розмір ПЗ для подальшого порівняння з відповідною оцінкою кількістю строк коду.

Після цього з'являться результати оцінювання розміру ПЗ.

ДОДАТОК Д - ПРОГРАМА ТА МЕТОДИКА ВИПРОБОВУВАНЬ

Об’єкт випробувань: Програма для оцінювання розміру web застосунків, що створюються з використанням фреймворку CakePhp.

Ціль випробувань: перевірка відповідності програми CakePhpSizeEst до вимог технічного завдання і повноти реалізації.

Склад програмної документації:

- технічне завдання;
- інструкція користувача;
- опис програми;
- текст програми;
- програма та методика випробувань.

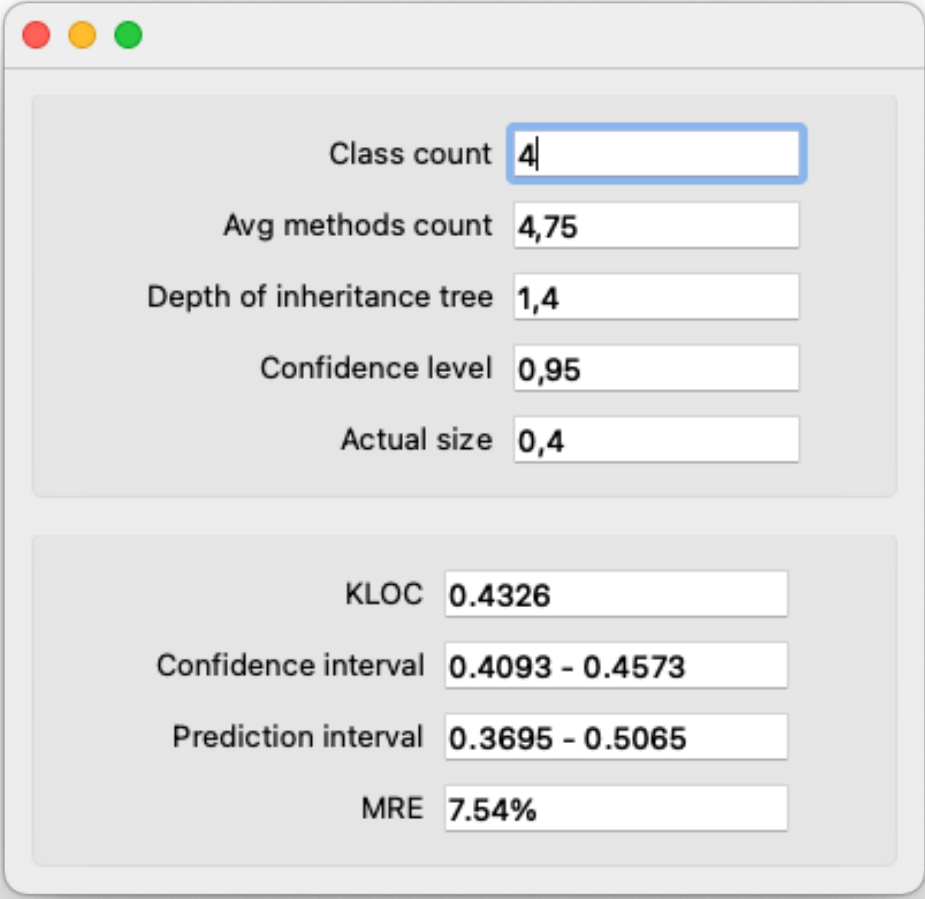
Технічні вимоги.

Програма має експлуатуватись на обладнанні з параметрами не нижчими за наступні: 64-бітний процесор на архітектурі x86 або ARM, оперативна пам’ять 4Гб (2 Гб мінімум), простір на жорсткому диску 100 Мб мінімум.

Методи випробувань.

Випробування програми виконується представником замовника в присутності представника розробника. Відповідно до опису програми виконується запуск програми CakePhpSizeEst. Далі виконується перевірка роботи програми. Виконується перевірка всіх екранних форм програми, а також всіх допустимих для виконання операцій всередині екранних форм згідно до посібника користувача. Якщо програма демонструє адекватне функціонування згідно до вимог ТЗ та представленої програмної документації, то робота з розробки програми вважається виконаною, і виконується впровадження програми до експлуатації. В протилежному випадку програма має бути відправлено на доробку розробникові.

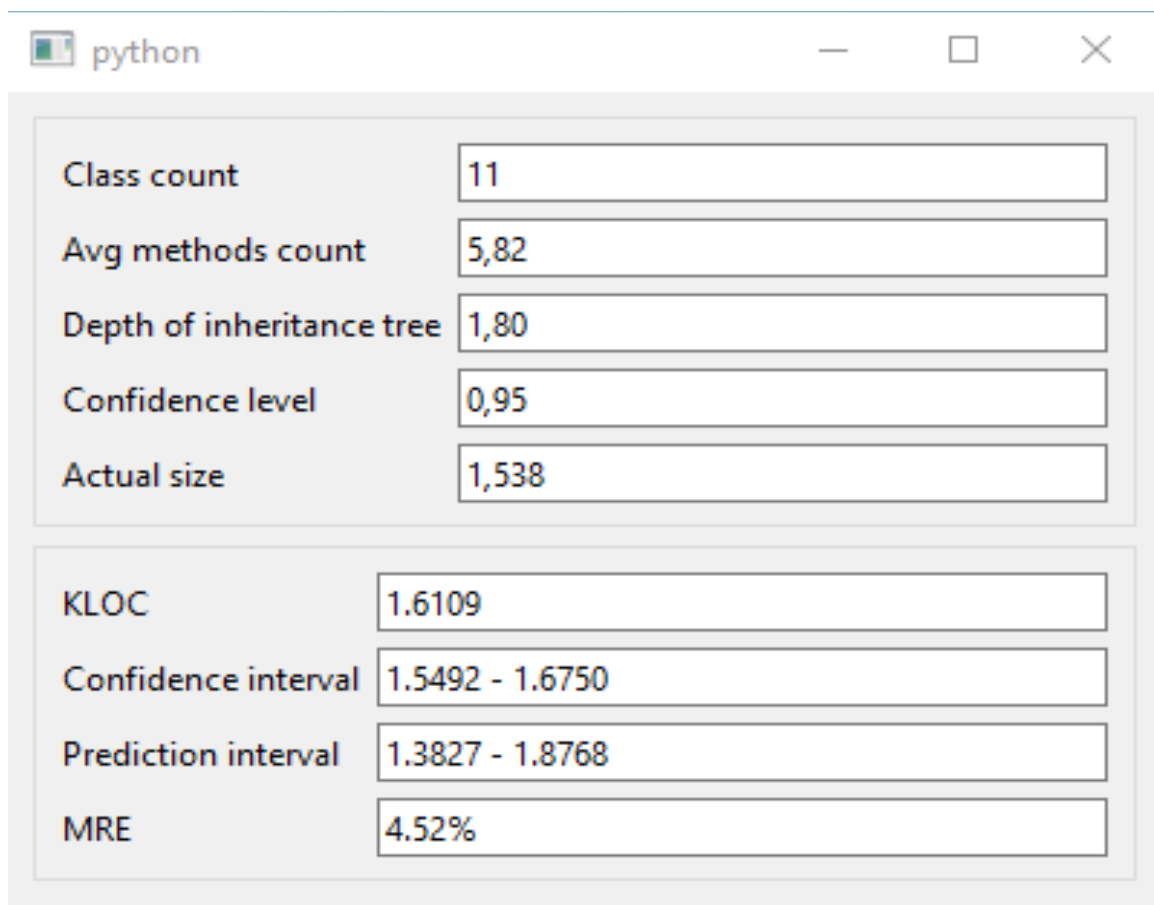
Використовуються тести за методологією „чорної скриньки” (рис.Д.1-Д.3).



The image shows a graphical user interface window with a light gray background and a standard macOS-style title bar with red, yellow, and green window control buttons. The window is divided into two main sections by a horizontal line. The top section contains five input fields, each with a label and a value. The bottom section contains four input fields, each with a label and a value. The 'Class count' field is highlighted with a blue border.

Label	Value
Class count	4
Avg methods count	4,75
Depth of inheritance tree	1,4
Confidence level	0,95
Actual size	0,4
KLOC	0.4326
Confidence interval	0.4093 - 0.4573
Prediction interval	0.3695 - 0.5065
MRE	7.54%

Рисунок Д.1 – Графічне вікно програми

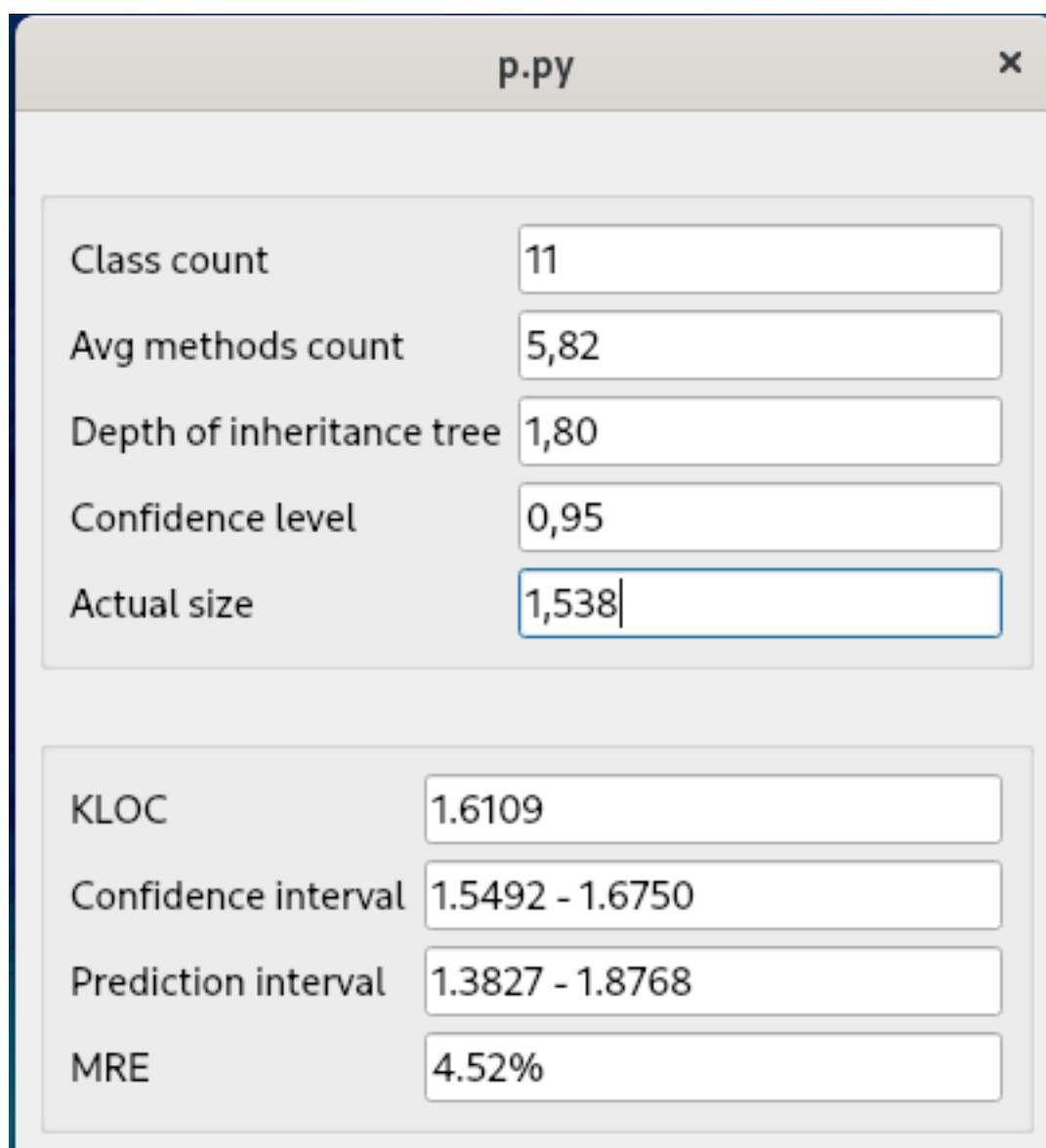


The image shows a screenshot of a Windows application window titled "python". The window contains two sections of data, each with labels on the left and text input fields on the right. The first section includes "Class count" (11), "Avg methods count" (5,82), "Depth of inheritance tree" (1,80), "Confidence level" (0,95), and "Actual size" (1,538). The second section includes "KLOC" (1.6109), "Confidence interval" (1.5492 - 1.6750), "Prediction interval" (1.3827 - 1.8768), and "MRE" (4.52%).

Class count	11
Avg methods count	5,82
Depth of inheritance tree	1,80
Confidence level	0,95
Actual size	1,538

KLOC	1.6109
Confidence interval	1.5492 - 1.6750
Prediction interval	1.3827 - 1.8768
MRE	4.52%

Рисунок Д.2 – Графічне вікно програми на платформі Windows 10



The image shows a terminal window titled 'p.py' with a close button 'x' in the top right corner. The window contains a form with two sections. The first section has five rows, each with a label and a text input field. The second section has four rows, each with a label and a text input field. The values entered in the fields are: Class count (11), Avg methods count (5,82), Depth of inheritance tree (1,80), Confidence level (0,95), Actual size (1,538), KLOC (1.6109), Confidence interval (1.5492 - 1.6750), Prediction interval (1.3827 - 1.8768), and MRE (4.52%).

Class count	11
Avg methods count	5,82
Depth of inheritance tree	1,80
Confidence level	0,95
Actual size	1,538

KLOC	1.6109
Confidence interval	1.5492 - 1.6750
Prediction interval	1.3827 - 1.8768
MRE	4.52%

Рисунок Д.3 – Графічне вікно програми на платформі Linux