

**ДОДАТОК А**  
**Графічний матеріал кваліфікаційної роботи**

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Інформаційних управляючих систем

**КВАЛІФІКАЦІЙНА РОБОТА**  
**ГРАФІЧНИЙ МАТЕРІАЛ**

Дослідження оптимального набору показників продуктивності веб-базованої  
інформаційної системи

(тема роботи)

Студент групи ІУСТМ-20-1  
(шифр групи)

(підпис)

Хоменко К. О.  
(прізвище, ініціали)

Науковий керівник

(підпис)

проф. Євланов М.В.  
(посада, прізвище, ініціали)

2021 р.

Таблиця А.1 – Загальна характеристика кваліфікаційної роботи

Тема кваліфікаційної роботи	Дослідження оптимального набору показників продуктивності веб-базованої інформаційної системи
Актуальність	Результатом проведення дослідження є побудова вдосконаленої моделі, що дозволить визначити оптимальний набір показників продуктивності веб-базованої інформаційної системи.
Мета досліджень	Побудова моделі, яка дасть можливість кількісно оцінити ефективність експлуатації веб-базованої інформаційної системи, ґрунтуючись на відомому наборі показників, значення яких вимірюються в ході моніторингу експлуатації системи.
Задачі досліджень	Аналіз існуючих методів оцінювання продуктивності веб-базованих інформаційних систем; опис теоретичного вирішення задачі; дослідження методів регресійного аналізу результатів моніторингу продуктивності веб-базованої інформаційної системи; практична апробація результатів магістерської кваліфікаційної роботи.
Методи досліджень	Методи регресійного аналізу, що дозволяють визначити оптимальний набір показників продуктивності експлуатованої веб-базованої інформаційної системи.

Кінець таблиці А.1

Нові наукові результати	Розроблена вдосконалена моделі, за якою проводитиметься подальший моніторинг продуктивності веб-базованих інформаційних систем.
Практична значимість роботи	Використання розробленої моделі сприятиме скороченню витрат на проведення моніторингу продуктивності ВБІС та збільшенню задовільності користувача шляхом зменшення часу відгуку.

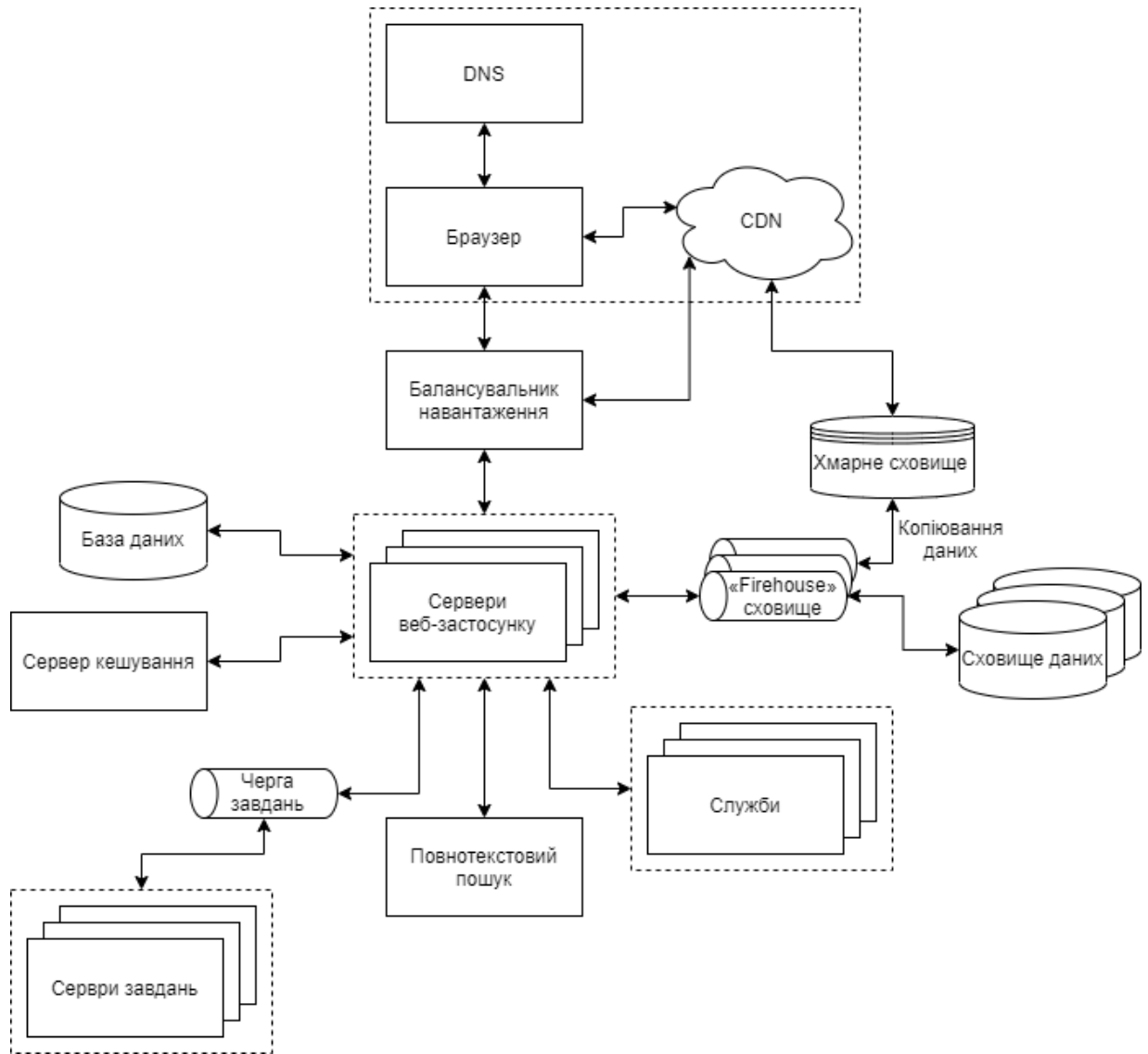


Рисунок А.1 – Схема загальної архітектури веб-застосунку

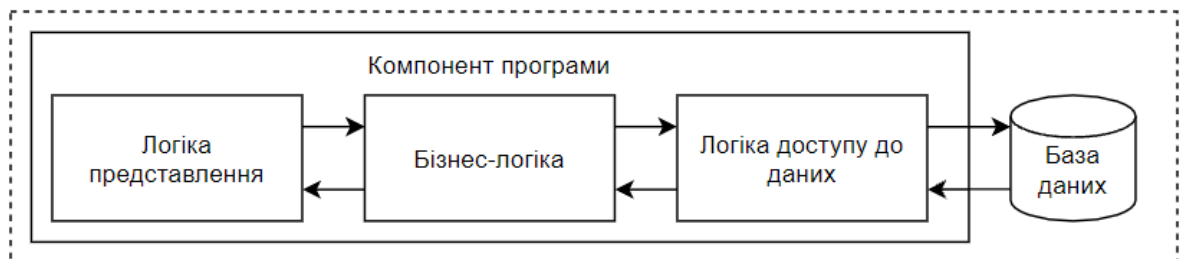


Рисунок А.2 – Схема однорівневої архітектури веб-застосунку

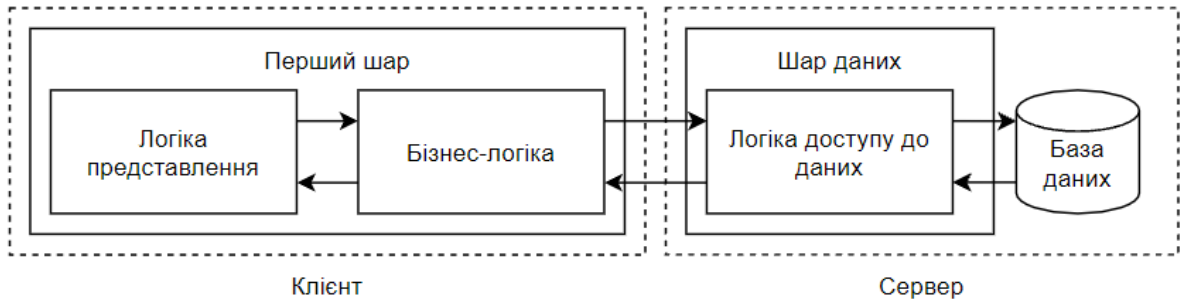


Рисунок А.3 – Схема дворівневої архітектури веб-застосунку

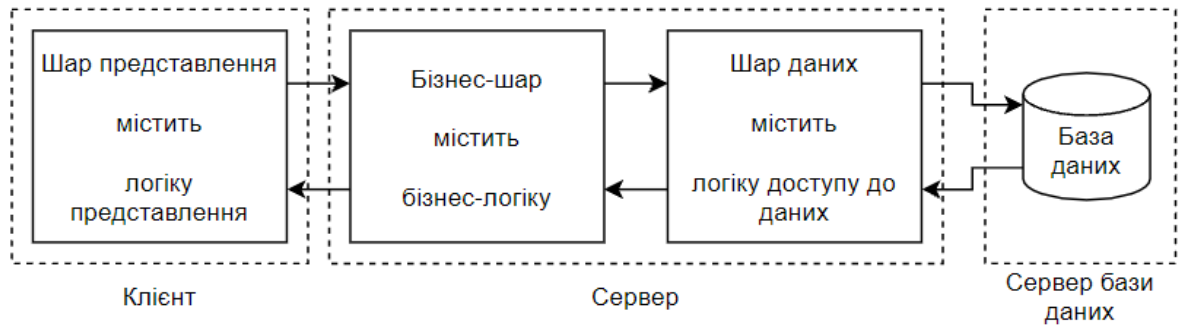


Рисунок А.4 – Схема тривірневої архітектури веб-застосунку

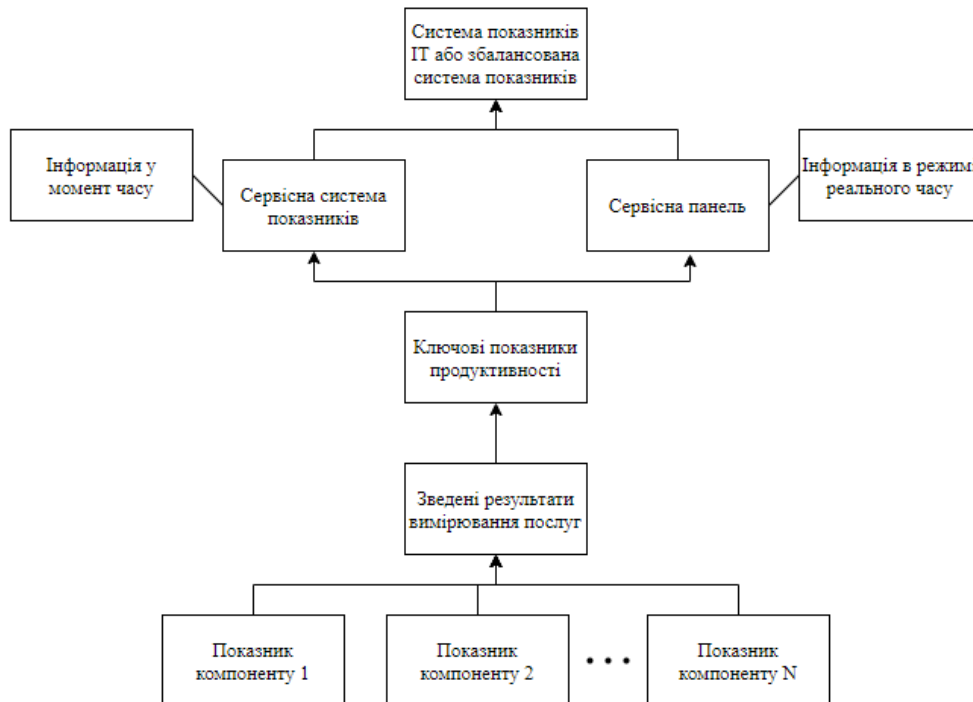


Рисунок А.5 – Схема моделі зчитування показників продуктивності сервісів

## ОПИС ТЕОРИТИЧНОГО ВИРШЕННЯ ЗАДАЧІ

Експлуатована ВБІС може бути подана у вигляді вектору, елементами якого є лінійно співзалежні елементи:

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}, \quad (\text{A.1})$$

де  $i$  – ідентифікатор елемента ВБІС,  $i = 1, \dots, n$ ;

$x_i$  – множина показників, які вимірюються у ході управлінням  $i$ -го елемента.

Загальна продуктивність ВБІС складається із суми продуктивностей функцій:

$$x_{perf} = (a_1 * x_{tr,a_1} + b_1) + \dots + (a_j * x_{tr,a_j} + b_j) = \sum_{i=1}^j a_i * x_{tr,a_i} + b_i, \quad (\text{A.2})$$

де  $x_{tr}$  – продуктивність функції, що бере участь у виконанні бізнес-сценарію;

$b$  – продуктивність інтерфейсу, через який опосередковано вплив функцій одна на одну;

$a$  – номер функції;

$z$  – кількість функцій.

У свою чергу, продуктивність функції визначається продуктивностями елементів, що беруть участь у транзакціях:

$$\begin{aligned} x_{tr,a} &= (a_1 * x_{t,a_1b_1} + b_1) + \dots + (a_k * x_{t,a_kb_k} + b_k) + \dots + (a_m * \\ &* x_{t,a_mb_m} + b_m) + \dots + (a_k * x_{t,a_kb_k} + b_k) + \dots + (a_1 * x_{t,a_1b_1} + \\ &+ b_1) = (a_m * x_{t,a_mb_m} + b_m) + 2 * \sum_{k=1}^m a * x_{t,ab} + b, \end{aligned} \quad (\text{A.3})$$

де  $x_{t,ab}$  – продуктивність  $b$ -го елемента ВБІС, що бере участь у  $a$ -ій транзакції;

$y$  – кількість елементів;

$a$  – номер транзакції;

$b$  – продуктивність інтерфейсу, через який опосередковано вплив елементів один на одного.

У експерименті, у якому проводиться  $n$  вимірювань значень показників  $x_i$ , де  $i = 1, 2, \dots, k$  – кількість показників, значення показнику  $x_i$  у  $j$ -му вимірюванні ( $j = 1, 2, \dots, n$ ) позначається як  $x_{ij}$ . Тоді набір точок  $x^j$  називатиметься планом експерименту, а планом експерименту буде матриця наступного вигляду:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{k1} \\ x_{21} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}. \quad (\text{A.4})$$

## ДОСЛІДЖЕННЯ МЕТОДІВ РЕГРЕСІЙНОГО АНАЛІЗУ

F-тест, або критерій Фішера – статистичний критерій, тестова статистика якого при виконанні нульової гіпотези має розподіл Фішера (F-розподіл). Критерій дозволяє оцінити доцільність введення додаткової незалежної змінної в лінійну модель множинної регресії, рівняння якої має наступний вигляд:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon, \quad (\text{A.5})$$

де  $y$  – залежна змінна;

$x_k, k = 1, 2, \dots, n$  – незалежні змінні;

$\beta_k, k = 1, 2, \dots, n$  – параметри моделі;

$n$  – об'єм вибірки;

$\varepsilon$  – випадкова складова.

Ідея цього критерію заснована на такому понятті, як сума квадратів регресії  $SSR$ :

$$SSR = \sum_{k=1}^n (\widehat{y}_k - \bar{y})^2, \quad (\text{A.6})$$

де  $n$  – об'єм вибірки;

$\widehat{y}_k$  – оцінка, отримана на основі регресійної моделі,  $k = 1, 2, \dots, n$ ;

$\bar{y}$  – середнє за всіма спостереженнями  $y$ .

Пояснювальна здатність моделі внаслідок введення нової змінної  $x_{extra}$  збільшується за наступною формулою:

$$SSR_{extra} = SSR_{full} - SSR_{initial}, \quad (\text{A.7})$$

де  $SSR_{extra}$  – внесок змінної  $x_{extra}$  у пояснення загальної мінливості  $y$ .

Даний критерій дозволяє перевірити гіпотезу  $H_0$ : внесок  $SSR$ , що вноситься  $x_{extra}$ , є недостатньо великим, через що цю змінну не слід включати в модель, та її альтернативу  $H_1$ : внесок  $SSR$ , що вноситься  $x_{extra}$ , є значущим, через що цю змінну слід включати в модель.

Для того, щоб перевірити ці гіпотези, слід перейти від показника  $SSR_{extra}$  до статистики наступного виду:

$$\gamma = \frac{SSR_{extra}}{MSE_{full}}, \quad (A.8)$$

де  $MSE_{full}$  – сума квадратів помилок  $SSE$ , що припадає на одну ступінь свободи  $df_{SSE}$ , тобто розрахункове значення F-критерію.

Значення  $MSE_{full}$  знаходиться за наступною формулою:

$$MSE_{full} = F_{real} = \frac{SSE}{df_{SSE}} = \frac{\sum_{k=1}^n (y_k - \hat{y}_k)^2}{n-m-2}, \quad (A.9)$$

де  $y_k$  – істинне значення результуючої змінної;

$\hat{y}_k$  – оцінка, отримана на основі регресійної моделі;

$n$  – об'єм вибірки;

$m$  – кількість змінних вихідної моделі (без  $x_{extra}$ ).

Загальне рівняння лінійної моделі множинної регресії у результаті застосування методу гребеневої регресії приймає вигляд (A.5).

У стандартній лінійній регресії вектор-стовпець  $y$  у розміром  $n \times 1$  має бути спроектований на простір стовпців матриці даних  $X$   $n \times m$ , стовпці якої сильно корельовані. Звичайна оцінка методом найменших квадратів коефіцієнтів  $\beta$  має наступний вигляд:

$$\beta = (X^T X)^{-1} X^T y, \quad (A.10)$$

де  $X^T$  – транспонована матриця  $X$ .

В умовах мультиколінеарності матриця  $X^T X$  стає погано обумовленою. Для вирішення цієї проблеми на величину коефіцієнтів накладається обмеження, тоді рівняння приймає вигляд:

$$\beta = (X^T X + \lambda I_n)^{-1} X^T y, \quad (\text{A.11})$$

де  $\lambda$  – невід’ємний параметр;

$I_n$  – одинична матриця розмірністю  $n \times n$ .

Опис методу регресійного аналізу, що дозволить повною мірою вирішити задачі, наведені у розділі 2, позначимо  $M_{RA}^*$ , при цьому описи розглянутих методів регресійного аналізу позначимо  $M_{R1}$  та  $M_{R2}$  для методу покрокової та гребеневої регресії відповідно ( $M_{Ri}, i = 1, 2$ ). Тоді задача вибору методу регресійного аналізу буде полягати у виділенні такого опису методу регресійного аналізу, відмінності якого від опису  $M_{RA}^*$  будуть мінімальними. Дана задача може бути описана наступною моделлю:

$$F = M_{RA}^* - M_{Ri} \rightarrow \min. \quad (\text{A.12})$$

Оскільки  $M_{RA}^*$  та  $M_{Ri}$  повинні існувати як текстові описи, необхідне виконання наступних умов:

$$\begin{cases} M_{RA}^* \neq \emptyset \\ M_{Ri} \neq \emptyset, i = 1, 2. \end{cases} \quad (\text{A.13})$$

Узагальнений опис  $M_R$  буде складатися із множини локальних критеріїв, яку можна описати наступним кортежем:

$$M_R = \langle C_1, C_2, C_3, C_4, C_5, C_6 \rangle, \quad (\text{A.14})$$

де  $C_1$  – критерій «відповідність опису вихідних даних»;

$C_2$  – критерій «придатність працювати із лінійно залежними значеннями»;

$C_3$  – критерій «ступінь відповідності даних припущенням методу»;

$C_4$  – критерій «складність (при виконанні методу витрачається небагато часу)»;

$C_5$  – критерій «невибагливість до знань та вмінь аналітика, що використовує метод»;

$C_6$  – критерій «ступінь відповідності результату застосування методу результату вирішення поставлених задач».

Для підвищення точності порівняння змісту описів методів подамо інтегральний критерій  $C_{conc}(Ri)$  як нормовану суму критеріїв збігу окремих розділів описів методів:

$$C_{conc}(Ri) = \frac{\sum_{j=1}^6 C_{conc}(C_j)}{6}, \quad (A.15)$$

де  $j$  – кількість критеріїв описів,  $j = 1, 2, \dots, 6$ ;

$C_{conc}(C_j)$  – критерій збігу  $j$ -го локального критерію описів  $M_{RA}^*$  та  $M_{Ri}$ .

Отже, базуючись на запропонованій моделі (A.12) – (A.13) можна виділити спосіб вибору методу регресійного аналізу, що складатиметься з таких етапів.

Етап 1. Визначення множини локальних критеріїв оцінювання, зазначених у (A.14).

Етап 2. За обраними на Етапі 1 критеріями визначити множину потреб аналітиків, що проводять аналіз моніторингу продуктивності ВБІС, у методі регресійного аналізу.

Етап 3. В результаті обробки визначених на Етапі 2 потреб з врахуванням відповідності обраним на Етапі 1 критеріям оцінювання елементів опису  $M_{RA}^*$  необхідно сформулювати опис такого методу регресійного аналізу  $M_{RA}^*$ , який надасть можливість в повній мірі вирішити поставлені

задачі.

Етап 4. Сформувати описи досліджуваних методів регресійного аналізу  $M_{Ri}$  шляхом пошуку інформації, що знаходиться у відкритому доступі, з врахуванням відповідності елементів кожного опису  $M_{Ri}$  обраним на Етапі 1 критеріям оцінювання.

Етап 5. Обрати групу з  $p$  експертів із числа аналітиків, що проводять аналіз моніторингу продуктивності ВБІС на підприємстві.

Етап 6. Запропонувати експертам групи, обраної на Етапі 5, для порівняння паспорти  $M_{RA}^*$  та  $M_{Ri}$ .

Етап 7. Визначити значення локальних критеріїв  $C_{conc}(C_j)$  шляхом анкетування експертів групи, обраної на Етапі 5.

Етап 8. Розрахувати усереднене значення локального критерію за середнім арифметичним.

Етап 9. Розрахувати значення інтегрального критерію збігу  $C_{conc}(C_j)$  за формулою (А.15), базуючись на отриманих на Етапі 7 результатах.

Етап 10. Знайти рішення задачі вибору методу регресійного аналізу за допомогою моделі (А.12) – (А.13) та завершити виконання методу.

Під час виконання Етапів 6, 7 та 8 було проведено анкетування обраних експертів, визначені значення локальних критеріїв  $C_{conc}(C_j)$ . Результати наведено у таблицях А.2-А.3.

Таблиця А.2 – Результати визначення та обробки значень локальних критеріїв  $C_{conc}(C_j)$  для покрокового методу регресійного аналізу

Локальний критерій	Значення першого експерту, %	Значення другого експерту, %	Значення третього експерту, %	Значення четвертого експерту, %	Усереднене значення локального критерію, %
$C_{conc}(C_1)$	95	100	96	99	97,5
$C_{conc}(C_2)$	5	3	10	5	5,75
$C_{conc}(C_3)$	100	95	94	96	93,75
$C_{conc}(C_4)$	94	95	100	95	96
$C_{conc}(C_5)$	100	96	95	95	96,5
$C_{conc}(C_6)$	75	80	73	70	74,5

Таблиця А.3 – Результати визначення та обробки значень локальних критеріїв  $C_{conc}(C_j)$  для гребеневого методу регресійного аналізу

Локальний критерій	Значення першого експерту, %	Значення другого експерту, %	Значення третього експерту, %	Значення четвертого експерту, %	Усереднене значення локального критерію, %
$C_{conc}(C_1)$	100	95	96	99	97,5
$C_{conc}(C_2)$	95	90	100	93	94,5
$C_{conc}(C_3)$	100	96	93	96	96,25
$C_{conc}(C_4)$	93	92	95	96	94
$C_{conc}(C_5)$	90	94	96	95	93,75
$C_{conc}(C_6)$	100	95	95	96	96,5

Під час виконання Етапу 8 були розраховані значення інтегрального критерію  $C_{conc}(C_j)$  для кожного з досліджених методів регресійного аналізу за формулою (А.15). Результат наведено у таблиці А.4.

Таблиця А.4 – Результати розрахунків значень інтегрального критерію

№ методу	Назва методу	Значення $C_{conc}(C_j)$ , %
1	Покрокова регресія	77,33
2	Гребенева регресія	95,42

## ПРАКТИЧНА АПРОБАЦІЯ РЕЗУЛЬТАТІВ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ

1	Дані з моніторингу													
	2	3	4	5	6	7	8	9	10	11	12	13	14	
Година	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	У	
1	1 8455,064	140,97	1588,36	19,90817	44,35711	8,3875	23,98166	98,67104	1827,471	389	518,67	0,013151	0,262886	
2	1 5347,259	1503,373	0,03	7,419245	67,94697	47,53	90,75185	103,0829	724,4776	9062,87	642,03	0,077143	0,567261	
3	1 4438,211	1202,124	392	8,526448	93,56469	38,79	46,51523	68,39912	1967,338	2388	1589,92	0,106705	0,968989	
4	1 8343,132	1964,316	1406,58	5,763166	58,48795	8138,91	5,605058	143,7761	1121,153	9057,65	1408,39	0,771589	1,142715	
5	1 5296,89	1237,973	38,88	16,70424	125,9984	8134,71	13,89273	6,685375	4069,171	0,03	14,62	0,939116	1,149048	
6	1 9950,422	712,7673	1401,42	11,99009	93,70457	100	72,58879	4,28862	5546,184	8141,49	21,61	0,105263	0,217975	
7	1 1651,79	475,0435	553,31	13,4794	128,3542	8,4337	42,45189	122,6046	1787,41	47,16	553,4	0,398045	2,576149	
8	1 3001,993	265,1628	2388,15	9,484875	86,90782	47,43	51,76661	109,5069	4876,794	21,61	2388,07	0,505407	1,532368	
9	1 7541,883	2217,944	0,03	2,152735	149,8179	38,7	29,97249	61,83777	4252,149	2388,03	9053,65	0,255994	2,566316	
10	1 3383,283	2087,192	9058,86	7,852471	73,29234	522,22	38,14575	109,0047	5693,092	391	1,3	0,797087	2,880084	
11	1 4590,764	1337,452	1,3	6,752851	95,41225	21,61	33,91613	66,00476	4348,992	8,414	47,71	0,491923	0,155267	
12	1 6508,605	627,8576	2388,13	11,78917	26,3503	2388,08	1,288576	132,6222	5397,609	521,97	522,2	0,264127	1,829231	
13	1 2540,168	527,371	2388	4,578114	15,49371	9049	68,21862	59,76985	1022,928	8,4277	2388,13	0,119506	1,679905	
14	1 7663,022	2130,061	642,03	19,73044	116,5855	8,4139	36,67318	139,8708	1157,615	100	8131,17	0,040005	0,279467	
15	1 3478,767	676,4846	553,01	2,623421	103,0982	2388,06	30,80447	35,08266	1278,54	1396,57	8,433	0,713315	1,269543	
16	1 5824,282	1788,045	8132,69	0,243167	100,7208	9055,37	92,55016	147,3795	3362,27	100	0,03	0,616758	1,199607	
17	1 6069,699	609,7054	1593	17,98855	27,14215	2388,13	21,36911	55,97132	2422,806	23,3718	392	0,971083	1,878216	
18	1 9731,492	26,96478	9050,6	1,498853	65,8038	642,22	52,09938	64,65279	2769,947	47,14	2388	0,729927	0,54926	
19	1 2033,461	273,5361	2388,12	10,24281	92,17393	393	37,0917	101,9363	3014,409	8,3569	100	0,907518	0,33831	
20	1 8850,385	162,3259	8,4439	1,968011	112,1759	100	48,7403	148,2516	1212,143	1391,67	38,99	0,907943	0,636701	
21	1 6708,568	2486,893	8133,81	0,646609	10,20472	38,8	21,01357	7,770841	3690,026	39,08	23,296	0,721618	1,841974	
22	2 6127,262	1370,028	521,43	12,8598	26,58297	2388,08	22,3578	145,7292	294,5038	47,27	518,67	0,918431	0,696241	

Рисунок А.6 – Частина значень показників продуктивності, зібраних у ході проведення моніторингу експлуатованої ВБІС

Ridge Regression Summary for Dependent Variable: Y (Дані з моніторингу in ANALYSIS)						
N=504	b*	Std.Err. of b*	b	Std.Err. of b	t(491)	p-value
Intercept			1,414595	0,186925	7,567713	0,000000
X1	-0,042918	0,043585	-0,000013	0,000013	-0,984714	0,325250
X2	0,107068	0,043640	0,000129	0,000053	2,453455	0,014496
X3	0,055750	0,044079	0,000019	0,000015	1,264760	0,206557
X4	-0,021288	0,043729	-0,003250	0,006676	-0,486820	0,626603
X5	-0,026617	0,044267	-0,000532	0,000884	-0,601285	0,547928
X6	0,007077	0,043856	0,000008	0,000048	0,161379	0,871861
X7	0,026875	0,044056	0,000855	0,001402	0,610024	0,542128
X8	-0,022562	0,043782	-0,000447	0,000867	-0,515327	0,606557
X9	-0,025168	0,044118	-0,000014	0,000024	-0,570459	0,568628
X10	-0,039673	0,043734	-0,000014	0,000015	-0,907144	0,364776
X11	0,065651	0,044116	0,000023	0,000015	1,488140	0,137356
X12	0,030104	0,044018	0,091559	0,133877	0,683904	0,494359

Рисунок А.7 – Підсумок гребеневої регресії

Як можна побачити на рисунку А.7, для досліджуваного набору даних значущим виявився лише коефіцієнт при змінній  $X_2$ , а також вільний член рівняння регресії; p-value для даних коефіцієнтів є меншим за 0,05.

Таким чином, рівняння регресії матиме наступний вигляд:  $Y = 1,1416 + 0,000129 * X_2$ .

## ВИСНОВКИ

У ході виконання магістерської кваліфікаційної роботи було проведено дослідження існуючих видів моделей веб-базованих інформаційних систем та визначено відмінності між ними. Водночас усі сучасні веб-застосунки мають спільну особливість: моніторинг продуктивності для них проводиться як для сукупності чорних ящиків, які взаємодіють один з одним через певні інтерфейси. Було визначено, що такий підхід унеможлиблює оцінювання продуктивності ВБІС в цілому та може стати причиною збільшення витрат на проведення процесу моніторингу.

Було визначено основний спосіб здійснення моніторингу продуктивності у вигляді проведення пасивного експерименту. У ході проведення такого експерименту фіксуються значення показників найнижчого рівня абстракції. Запропоновано моделі визначення продуктивності як окремих транзакцій, так і бізнес-сценаріїв.

У якості методів проведення аналізу моніторингу продуктивності ВБІС було запропоновано покроковий та гребеневий методи регресійного аналізу та складено їхній опис. Вивчення покрокової регресії виявило, що даний метод потребує додаткової обробки результатів моніторингу через неможливість обробляти лінійно залежні дані, які можуть з'явитися у ході проведення моніторингу. При проведенні порівняльного аналізу із використанням інтегрального критерію збігу та експертних оцінок було зроблено висновок, що для проведення аналізу результатів моніторингу доцільно використовувати метод гребеневої регресії.

Із використанням даних моніторингу продуктивності експлуатованої веб-системи КСАСК було проведено аналіз даних із використанням гребеневої регресії, який дозволив побудувати регресійне рівняння для вказаного набору даних та виявити значущі показники продуктивності ВБІС.