

DOI: <https://doi.org/10.32782/2524-0072/2024-61-118>

УДК 657.9

МОДЕЛЮВАННЯ ВЕЛИЧИНИ ДОХОДУ ВІД РЕАЛІЗАЦІЇ ПРОДУКЦІЇ КОМУНАЛЬНОГО ПІДПРИЄМСТВА

MODELING THE AMOUNT OF INCOME FROM THE SALE OF THE UTILITY COMPANY'S PRODUCTS

Прокопович Леонід Борисович

кандидат економічних наук, доцент,

Національний університет кораблебудування імені адмірала Макарова

ORCID: <https://orcid.org/0000-0003-2561-8862>

Шевчук Сергій Валентинович

кандидат економічних наук, доцент,

Національний університет кораблебудування імені адмірала Макарова

ORCID: <https://orcid.org/0000-0002-1179-9562>

Prokopovich Leonid, Shevchuk Serhii

Admiral Makarov National University of Shipbuilding

Досліджено завдання підвищення достовірності оцінювання рівня чистого доходу від реалізації продукції комунального підприємства. Проаналізована залежність величини чистого доходу від реалізації продукції від таких факторів як: собівартість реалізованої продукції, інші операційні доходи. Наявність мультиколінеарності між факторами викликало потребу побудови регресійних моделей за методом найменших квадратів за кожним фактором. Також, були побудовані моделі за методом найменших квадратів після застосування до факторів методу головних компонент. В роботі розглянуті і моделі на основі машинного навчання. Були побудовані регресійні моделі на основі k -ближніх сусідів та бінарного дерева рішень. Найліпшої серед всіх досліджених регресійних моделей оцінювання рівня чистого доходу від реалізації продукції комунального підприємства виявилася модель бінарного дерева рішень.

Ключові слова: дохід від реалізації продукції, комунальні підприємства, регресія, метод головних компонент, дерево рішення, k -ближніх сусідів.

The task of increasing the reliability of estimating the level of net income from the sale of products of a utility company is studied. The dependence of the amount of net income from sales of products on such factors as cost of sales and other operating income is analyzed. The presence of multicollinearity between the factors necessitated the construction of regression models using the least squares method for each factor. To build a stepwise model, the paper uses a transformation based on the decimal logarithm. We also built models using the least squares method after applying the principal components method to the factors. To test the internal parameters of the regression models built using the least squares method for significance, we used the value of their p -values. All models where the primary data were transformed using the principal components method failed the significance test. After checking the internal parameters in order to find the best model, the obtained regression models were compared using the following indicators: coefficient of determination, Fisher's criterion, and average approximation error. As a result of the comparison, the logarithmic model was recognized as the best single-factor regression model by the least squares method. To solve the problem of increasing the reliability of estimating the level of net income from the sale of products of a utility company, the paper also considers models based on machine learning. Regression models based on k -nearest neighbors and a binary decision tree were built. To build models using machine learning methods, the initial data were standardized and transformed using the principal component method. When comparing the resulting models, it turned out that the two models based on machine learning methods have a smaller approximation error than the models based on the least squares method. The best regression model for estimating the level of net income from the sale of a utility company's products was the binary decision tree model. The coefficient of determination of this model was 0.987, and the value of the average approximation error was 0.91%.

Keywords: revenue from sales of products, utility companies, regression, principal component method, decision tree, k -nearest neighbors.



Постановка проблеми. В умовах системних кризових явищ, які мають місце в економіці країни підприємства все більше уваги приділяють прогнозуванню показників, які характеризують фінансові результати їх діяльності. Особливу місце в завданні прогнозування фінансового результату займає величина доходу від реалізації. Питання підвищення якості прогнозування даного показника не втрачає актуальності в сучасних економічних умовах. В свою чергу, стрімкий розвиток засобів обчислювальної техніки та спеціалізованих пакетів прикладних програм дозволяє застосовувати, при вирішенні проблеми прогнозування величини доходу від реалізації, широке коло методів.

Аналіз останніх досліджень і публікацій.

Дослідженню проблеми фінансових результатів як в теоретичному, так і в методологічному напрямках присвячено роботи багатьох фахівців. Серед яких можна звернути на роботи: Гудзь Н. В. [1], Кінєва Т. С., Остапенко А. Д. [2], Кобець С. П., Лузіна А. О. [3], Позднякова В. Д. [5], Ткаченко І. С., Проскурович О. В. [6]. Проте, достотне велике коло проблем фінансових результатів, особливо в частині прогнозування величини доходу від реалізації, чикають на своє розв'язання, потребують удосконалення та уточнення.

Метою дослідження є підвищення достовірності оцінювання рівня чистого доходу

від реалізації продукції комунального підприємства.

Виклад основного матеріалу дослідження. В якості початкових даних для побудови моделі оцінювання рівня чистого доходу від реалізації продукції комунального підприємства була використана інформація про фінансові результати діяльності МКП «Чернівцітеплокомуненерго» (див. табл. 2 публікації Марусяк Н. Л. [4]). Y – чистий дохід від реалізації продукції, X_1 – собівартість реалізованої продукції, X_2 – інші операційні доходи.

Для перевірки первинних даних на наявність викидів були побудовані діаграми типу «ящик з вусами», які наведені на рис. 1. Вказані діаграми підтверджують відсутність викидів у даних.

Побудовані гістограми та графіки кумулятивної функції розподілу (cumulative distribution function – CDF) показують, що початкові дані не підпорядковуються нормальному розподілу. У зв'язку з чим, для дослідження силу впливу факторів на залежну змінну був використаний коефіцієнт кореляції Спірмена. Розрахунок кореляційної матриці наведено у табл. 1.

Розрахунок кореляційної матриці свідчить, що фактор X_1 – оказує сильний вплив на залежну змінну, а фактор X_2 – середній вплив. Проте, між факторами наявна мультиколінеарність (коефіцієнт кореляції склав 0,9).

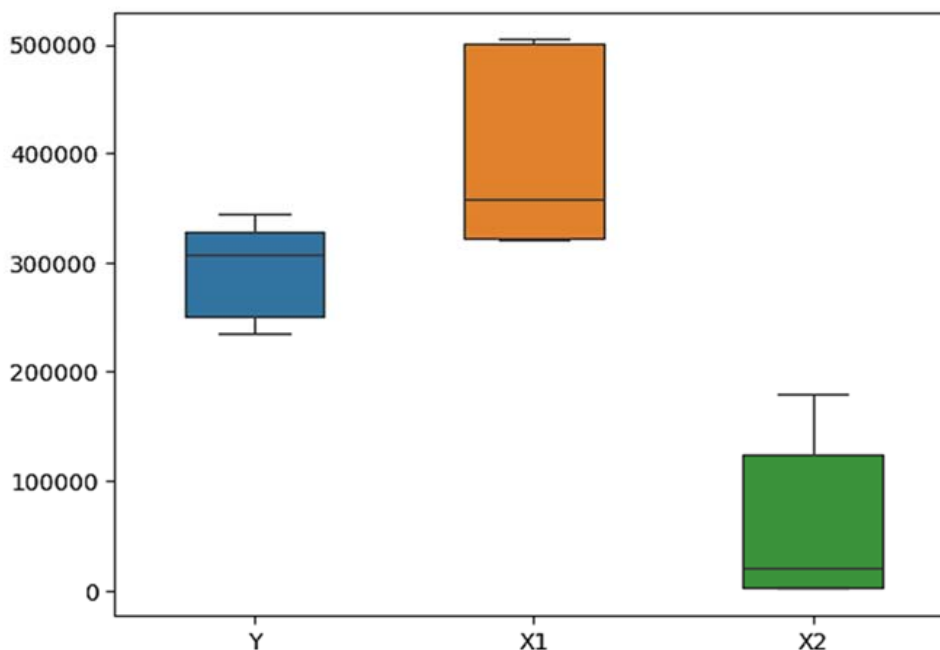


Рис. 1. Діаграми типу «ящик з вусами»

Джерело: складено авторами

Для перевірки наявності недопустимого для використання методу найменших квадратів був зроблений розрахунок зворотної матриці. Результати розрахунку наведені у табл. 2.

Таблиця 1

Кореляційна матриця

	Y	X1	X2
Y	1	0,9	0,7
X1	0,9	1	0,9
X2	0,7	0,9	1

Джерело: складено авторами

Таблиця 2

Зворотна матриця, R^{-1}

	X1	X2
X1	5,263158	-4,73684
X2	-4,73684	5,263158

Джерело: складено авторами

Значення на головній діагоналі зворотної матриці перевищують граничне значення (4,0) для використання обох факторів при побудові регресійної моделі за методом найменших квадратів. У зв'язку з чим вирішено,

спробувати побудувати однофакторні моделі за допомогою методу найменших квадратів за кожним фактором окремо за наступними рівняннями (див. рівняння 1–8).

$$\hat{y} = a_0 + a_1 x + \varepsilon, \quad (1)$$

$$\hat{y} = a_1 x + \varepsilon, \quad (2)$$

$$\hat{y} = a_0 + a_1 \ln(x) + \varepsilon, \quad (3)$$

$$\hat{y} = a_1 \ln(x) + \varepsilon, \quad (4)$$

$$\hat{y} = a_0 + a_1 x + a_2 x^2 + \varepsilon, \quad (5)$$

$$\hat{y} = a_0 + a_1 x + a_2 x^2 + \varepsilon, \quad (6)$$

$$\hat{y} = a_0 x^{a_1} + \varepsilon, \quad (7)$$

$$\hat{y} = x^{a_1} + \varepsilon, \quad (8)$$

Для побудови моделей за рівняннями 7 та 8 було здійснено перетворення початкових даних за допомогою десятичного логарифму. Отримані величини внутрішніх параметрів моделей за методом найменших квадратів наведені у табл. 3.

З метою перевірки отриманих величин внутрішніх параметрів моделі на суттєвість був зроблений розрахунок їх p -значень. Даний розрахунок наведено й табл. 4.

Інформація, яка наведена у табл. 4 показує, що у 50% моделей за фактором X1 виявилися несуттєві внутрішні параметри (моделі

Таблиця 3

Параметри однофакторних моделей за методом найменших квадратів

Номер рівняння	Внутрішні параметри		
	a_0	a_1	a_2
для фактору X1			
1	106482,629	0,466064606	-
2	-	0,720242048	-
3	-2195988,566	193262,6324	-
4	-	22839,49113	-
5	-1237835,458	7,214978409	-8,12043E-06
6	-	1,012816529	-6,71142E-07
7	53,56862999	0,66743281	-
8	-	0,976381745	-
для фактору X2			
1	262711,381	0,467095566	-
2	-	2,271429344	-
3	150245,205	14437,36144	-
4	-	29086,75025	-
5	265805,4524	0,166857312	1,74151E-06
6	-	7,054793363	-2,9859E-05
7	180154,9441	0,048059007	-
8	-	1,2280011	-

Джерело: складено авторами

Таблиця 4

p-значення внутрішніх параметрів моделей за методом найменших квадратів

Номер рівняння	p-значення для параметру		
	a ₀	a ₁	a ₂
<i>для фактору X1</i>			
*1	0,139888874	0,037198615	-
2	-	3,25925E-05	-
3	0,04198963	0,030477523	-
4	-	0,000111638	-
*5	0,125721536	0,097613279	0,109270544
*6	-	0,00476308	0,112184932
*7	0,196803501	0,037508457	-
8	-	1,03405E-09	-
<i>для фактору X2</i>			
*1	0,001084159	0,1155322	-
*2	-	0,084754822	-
*3	0,274786371	0,286659099	-
4	-	0,00017605	-
*5	0,011233686	0,909118705	0,834754957
*6	-	0,311197196	0,463838898
*7	8,26583E-05	0,317533306	-
8	-	0,000320608	-

* - моделі, де величина внутрішніх параметрів виявилася несуттєвою

Джерело: складено авторами

за рівняннями: 1, 5–7). Для моделей найменших квадратів побудованих за фактором «інші операційні доходи», перевірку на суттєвість внутрішніх параметрів пройшли тільки дві моделі.

Для моделей (2–4, 8 за фактором X1 та 4, 8 за фактором X2), внутрішні параметри яких виявилися суттєвими, була побудована порівняльна таблиця (див. табл. 5).

Величини коефіцієнту детермінації свідчать, що одна модель має прийнятне значення ($0,5 < R^2 < 0,8$), ще одна модель має добру точність апроксимації ($0,8 < R^2 < 0,95$), а чотири моделі мають високу точність апроксимації ($0,95 < R^2 < 1,00$). Порівняння розрахункових та табличних величин критерія Фішера свідчать, що всі моделі виявилися статистично значимими. Дослідження середньої помилки апроксимації виявила, що у трьох моделей вона перевищує критичне значення (10%), тому дані моделі були усунені від подальшого порівняння. Всі моделі, які за вказаними у табл. 5 показниками досягли прийнятних значень, виявилися моделі за фактором X1. У логарифмічній моделі за рівнянням 3 найліпша величина середньої

помилки апроксимації, але моделі за рівняннями 2 і 4 ліпше за величинами коефіцієнту детермінації та критерієм Фішера. Тому, хоча найліпшою моделлю серед моделей за фактором X1 була обрана логарифмічна модель (див. рівняння 9), для оцінки рівня чистого доходу від реалізації продукції комунального підприємства можна використовувати й інші знайдені регресійні моделі за методом найменших квадратів (див. рівняння 10, 11).

$$\hat{y} = -2195988,566 + 193262,6324 \ln(x) + \varepsilon, \quad (9)$$

$$\hat{y} = 0,720242048x + \varepsilon, \quad (10)$$

$$\hat{y} = x^{0,976381745} + \varepsilon, \quad (11)$$

На наступному етапі дослідження, з метою усунення впливу мультиколінеарності було вирішено перетворити значення факторів за допомогою методу головних компонент. Перед використанням методу головних компонент первинні дані були стандартизовані. В наслідок застосування методу головних компонент встановлено, що значення дисперсії, які пояснюються головними компонентами складають 97,5% (для першої головної компоненти PC1) та 2,5% (для PC2). Тому, вирішено для побудови регресійної моделі за методом

Таблиця 5

Порівняння однофакторних моделей за методом найменших квадратів

Фактор, Номер рівняння	Коефіцієнт детермінації, R ²	Критерій Фішера, F		Середня помилка апроксимації, А у %
		розрахунковий	табличний	
X1, 2	0,99069177	425,727	7,709	7,93
X1, 3	0,833293888	14,996	10,128	4,87
*X1, 4	0,982795684	228,500	7,709	13,09
X1, 8	0,999947489	76170,200	7,709	7,90
*X2, 4	0,978411305	181,282	7,709	12,58
*X2, 8	0,970903503	133,474	7,709	27,38

* - моделі які не досягли прийнятних значень

Джерело: складено авторами

найменших квадратів використовувати дані першої головної компоненти за рівняннями 1, 2, 5, 6, 12, 13.

$$\hat{y} = a_0 + a_1x + a_2x^2 + a_3x^3 + \varepsilon, \quad (12)$$

$$\hat{y} = a_0 + a_1x + a_2x^2 + a_3x^3 + \varepsilon. \quad (13)$$

Отримані в наслідок побудови моделей за РС1 їх внутрішні параметри та *p*-значення наведені у табл. 6. Як можна побачити, всі отримані регресійні моделі не пройшли перевірку на суттєвість внутрішніх параметрів та були усунені від подальшого дослідження.

Також, дані першого головного компонента були використанні при побудові моделей за

методами машинного навчання: *k*-ближніх сусідів та методом бінарного дерева рішень. В якості гіперпараметрів моделей були обрані: кількість сусідів (для *k*-ближніх сусідів), максимальна глибина дерева (для дерева рішень). В наслідок побудови вказаних регресійних моделей оцінювання рівня чистого доходу від реалізації продукції комунального підприємства була побудована порівняльна таблиця (див. табл. 7).

Інформація табл. 7 показує, що за величиною коефіцієнту детермінації R², одна модель має прийнятне значення (метод *k*-ближніх сусі-

Таблиця 6

Внутрішні параметри моделей за методом найменших квадратів та їх *p*-значення за першим головним компонентом (РС1)

Номер рівняння	a ₀	a ₁	a ₂	a ₃
внутрішні параметри				
1	293675,8001	61191,24848	-	-
2	-	61191,24688	-	-
5	320888,9508	80509,27971	-74127,12812	-
6	-	-130203,837	734421,0101	-
12	343604,17	19508,53917	-186415,4812	194081,9746
13	-	60313,24671	872709,0804	-486820,8114
<i>p</i> -значення				
*1	0,000178114	0,061790926	-	-
*2	-	0,813613317	-	-
*5	0,026473483	0,206218427	0,648424714	-
*6	-	0,256280134	0,012578832	-
*12	0,131738481	0,88463861	0,577796741	0,631166317
*13	-	0,884026974	0,099492734	0,640866562

* - моделі які не досягли прийнятних значень

Джерело: складено авторами

Таблиця 7

Порівняння моделей побудованих за допомогою метода k-ближніх сусідів та бінарного дерева рішень

Значення гіперпараметру	Коефіцієнт детермінації, R ²	Середня помилка апроксимації, A у %	PRED (0,25)
<i>метод k-ближніх сусідів</i>			
2	0,829	4,68	1,00
3	0,654	7,51	1,00
<i>метод дерева рішень</i>			
2	0,987	0,91	1,0

Джерело: складено авторами

дів при $n_neighbors = 3$), одна модель добра (метод k -ближніх сусідів при $n_neighbors = 2$) та одна модель, за даним показником, має високу точність апроксимації (метод дерева рішень при $max_depth = 2$). За всіма моделями величина рівня прогнозування PRED склала 1.0, що є добрим. Також, величини середньої помилки апроксимації моделей не перевищують граничну межу (10%). Найліпші величини коефіцієнту детермінації та середньої помилки апроксимації у моделі за методом дерева рішень. Дана моделі представлена у вигляді схеми на рис. 2.

У підсумку, можна казати, що для оцінювання рівня чистого доходу від реалізації продукції комунального підприємства рекомендується застосовувати побудоване регресійне дерево. Запропонована модель має високе значення коефіцієнту детермінації та

найменшу середню помилку апроксимації із досліджених моделей.

Висновки. В роботі, досліджені регресійні моделі для оцінювання рівня чистого доходу від реалізації продукції комунального підприємства. Для побудови моделей, в якості факторів використовувалися: собівартість реалізованої продукції, інші операційні доходи. Були побудовані моделі за методами найменших квадратів, k -ближніх сусідів та бінарного дерева рішень. Перевірка отриманих моделей та їх порівняння дозволило виявити найліпшу модель – модель на основі дерева рішень. Перспективи подальших досліджень полягають у побудові довірчих інтервалів за отриманими моделями та вивчення інших підходів до побудови більш якісних моделей оцінювання рівня чистого доходу від реалізації продукції комунального підприємства.

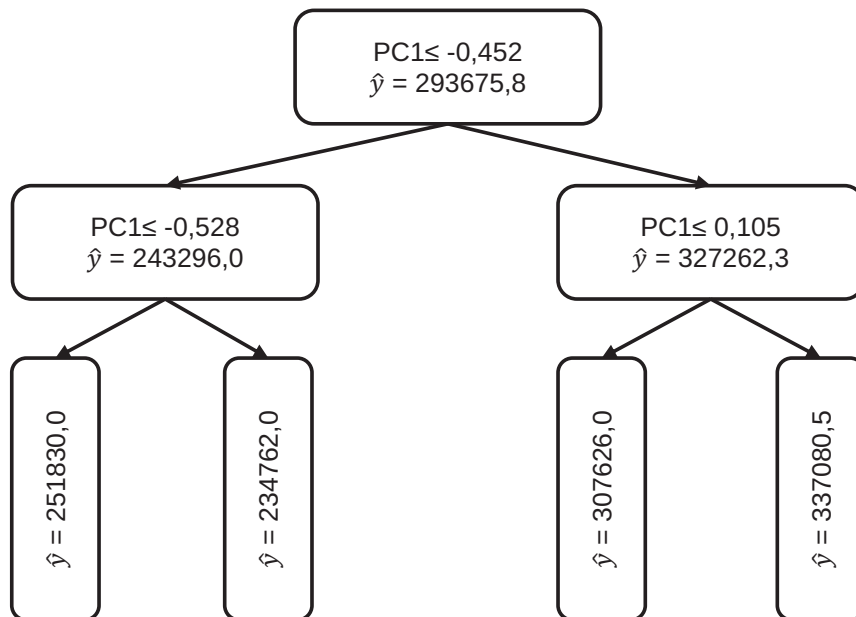


Рис. 2. Регресійне дерево для оцінювання рівня чистого доходу від реалізації продукції комунального підприємства

Джерело: складено авторами

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Гудзь Н. В. Облікова модель формування фінансових результатів діяльності підприємств в умовах євроінтеграційних процесів в Україні. *Економіка та суспільство*. № 13. 2017. С. 1339–1346. URL: https://economyandsociety.in.ua/journals/13_ukr/223.pdf
2. Кінева Т. С., Остапенко А. Д. Звіт про фінансові результати в системі стратегічного управління підприємством. *Науковий огляд*. 2015. № 9. С. 1–7. URL: <https://core.ac.uk/download/pdf/217451031.pdf>
3. Кобець С. П., Лузіна А. О. Застосування адаптивних моделей для прогнозування чистого доходу від реалізації продукції. 2019. № 4. URL: http://www.economy.nayka.com.ua/pdf/4_2019/42.pdf
4. Марусяк Н. Л. Фінансові ризики та їх вплив на фінансовий стан підприємства. *Ефективна економіка*. 2024. № 1. URL: <https://www.nayka.com.ua/index.php/ee/article/view/2885/2921>
5. Позднякова В. Д. Економетрична модель оцінювання фінансових результатів діяльності банків України. *Економіка і суспільство*. 2017. № 11. С. 582–587. URL: https://economyandsociety.in.ua/journals/11_ukr/94.pdf
6. Ткаченко І. С., Проскурівич О. В. Економіко-математичне моделювання фінансового результату підприємства. *Економіка: реалії часу*. 2017. № 3. С. 84–94. URL: <https://economics.net.ua/files/archive/2017/No3/84.pdf>

REFERENCES:

1. Hudz N. V. (2017) Oblikova model formuvannya finansovykh rezultativ diialnosti pidpriemstv v umovakh yevrointehratsiinykh protsesiv v Ukraini [Accounting model of formation of financial results of activity of enterprises in conditions of integration processes in Ukraine]. *Economy and Society*, no. 13, pp. 1339–1346. Available at: https://economyandsociety.in.ua/journals/13_ukr/223.pdf
2. Kineva T. S., Ostapenko A. D. (2015) Zvit pro finansovi rezultaty v systemi stratehichnoho upravlinnia pidpriemstvomi [Report on the financial results in a system of strategic management]. *Scientific review*, no. 9, pp. 1–7. Available at: <https://core.ac.uk/download/pdf/217451031.pdf>
3. Kobets S., Luzina A. (2019) Zastosuvannya adaptyvnykh modelei dlia prohnozuvannya chystoho dokhodu vid realizatsii produktsii [Application of adaptive models for forecasting a net sales]. *Efektivna ekonomika*, no. 4. Available at: http://www.economy.nayka.com.ua/pdf/4_2019/42.pdf
4. Marusiak N. (2024) Finansovi ryzyky ta yikh vplyv na finansovyi stan pidpriemstva [Methodical approaches to the assessment of the volume of services provided by enterprises in the transport industry]. *Efektivna ekonomika*, no. 1. Available at: <https://www.nayka.com.ua/index.php/ee/article/view/2885/2921> (in Ukrainian)
5. Pozdnyakova V. D. (2017) Ekonometrychna model otsiniuvannya finansovykh rezultativ diialnosti bankiv Ukrainy [Econometric model of estimation of financial results of the activities of banks of Ukraine]. *Economy and Society*, no. 11, pp. 582–587. Available at: https://economyandsociety.in.ua/journals/11_ukr/94.pdf
6. Tkachenko I. S., Proskurovych O. V. (2017) Ekonomiko-matematychno modeliuвання finansovoho rezultatu pidpriemstva [Economic and mathematical modeling of enterprise's financial results]. *Economics: time realities*, no. 3. Available at: <https://economics.net.ua/files/archive/2017/No3/84.pdf>