

ПРИВЕДЕННЯ МАТЕМАТИЧНОЇ МОДЕЛІ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ВЕБ-ПОШУКУ ДО ЛІНГВІСТИЧНОЇ КАТЕГОРИЗАЦІЇ

Вінницький національний технічний університет

Анотація

Розглядається математична модель оцінювання ефективності веб-пошуку, особливості, принципи роботи, основні переваги та недоліки, пропонується подання її у вигляді лінгвістичних термів.

Ключові слова: математична модель, прогнозування, обробка даних, статистичний аналіз.

Abstract

We consider a mathematical model for evaluating the effectiveness of web search, their features, principles of operation, the main advantages and disadvantages, it is proposed to present it in the form of linguistic terms.

Ключові слова: mathematical model, forecasting, data processing, statistical analysis.

Вступ

Однією з основних проблем, пов'язаних з замовленням банківських послуг в інтернет-банкінгу, є правильний підбір саме тієї послуги яка потрібна клієнту. У даному розділі пропонується модель, що дозволяє значно знизити витрати на їх підбір без втрати ефективності за рахунок нової методики пошуку банківських послуг.

Результати дослідження

Для ефективного поширення інформації потрібно знати, на якому відкритому джерелі інформації помістити необхідний вміст[1]. Це можна досягти застосуванням моделі, яка зможе повністю зробити оцінку доцільності сайту та вирішити наступні задачі:

1. Аналіз факторів, створення показників та критеріїв ефективності веб-сайтів в Інтернеті. При створенні метрик та критеріїв ефективності вебсайту мають бути дотримані такі вимоги:

1. Критерії мають відображати характеристики інформаційного змісту веб-сайту: розміщені дані повинні бути точними, актуальними, доступними для користувача та постійно оновлюватися.

2. Критерії мають відображати характеристики сайту в загальному: актуальність, змістовність, презентабельність, зручність.

3. При обчисленні значень критеріїв мають переважати об'єктивні джерела інформації над суб'єктивними.

У процесі досліджень методом аналізу предметної галузі оцінювання ефективності розроблено інфологічну модель факторів, показників та критеріїв ефективності інтернет-сайтів, що показана в таблиці 1.

Таблиця 1 – Інфологічна модель факторів, показників та критеріїв ефективності

№ з/п	Група факторів	№ з/п	Фактор у групі	Показник	Критерій
1	Показники Юзабіліті	1	Сумарний час, проведений користувачами на сайті за день	KA	KA max
		2	Середня кількість переглянутих сторінок за один	KB	KB max
		3	Швидкість завантаження сторінки	KC	KC max
		4	Час завантаження сторінки	KD	KD min
		5	Внутрішні посилання		
		6	Зовнішні посилання	KE	KE min
2	SEO критерії	7	Кількість сторінок в Яндекс	SY	SY max
		8	Кількість сторінок в Google	SG	SG max
		9	Тематичний індекс цитувань у Яндекс	SQ	SQ max
		10	Google Page Rank	SPR	SPR max
		11	Рейтинг сайту в світі	SK	SK min

		1	Зворотні посилання	SB	SB max
		2			

Для рішення даної задачі було взято за основу показники Юзабіліті та деякі критерії продуктивності SEO-оптимізації веб-сайта. Показник Юзабіліті – це характеристика, яка показує готовність програмного забезпечення до застосування. Якщо цей показник використати для сайту, тоді це дозволить зрозуміти його функціонал та зручність у застосуванні. SEO – це набір заходів, які запроваджуються з метою досягнення високих рейтингових позицій сайту у пошуковій системі за вибраною групою ключових запитів. Значення показників, які показані у таблиці 2.1, взято з сайту google.com/analytics[2].

Вказана модель бере до уваги взаємозв'язок факторів, показників та критеріїв двох груп, це показники Юзабіліті та SEO-критерії. Вказані критерії являють собою суперечну систему, яка сигналізує про трансформацію задачі оцінювання продуктивності до багатокритерійної форми. Завдання багатокритерійного оцінювання продуктивності сайту полягає в наступному: беремо перний образ (сайт), який має найгірші показники та проводимо розрахунки, в майбутньому порівнюючи дані із іншими сайтами.

II. Створення моделі вироблення[3] висновку про продуктивність упровадження нового сайту поміж множини інших ресурсів в інтернеті є зведенням переліку суперечливих часткових критеріїв таблиці 2.1 до узагальноної оцінки. Для реалізації потрібно вибрати нелінійну схему компромісів Вороніна А.М., виходячи із доведених переваг вказаної схеми агрегації. Сутністю зміни значень, які роблять опис часткових критеріїв (табл. 2.1) - дискретна форма. Тому зазначена вище інформація буде мати наступний вид:

$$Y(y_0) = \sum_{l=1}^k \gamma_{0l} (1 - y_{0l})^{-1} \rightarrow \min \quad (1)$$

де $l = 1 \dots k$ – число включених у згортку часткових критеріїв; γ_{0l} – ваговий коефіцієнт; $1 - y_0$ – нормований частковий критерій оптимальності. Нормування критеріїв, які є в складі згортки (1), створюється залежно від найвищого та найнижчого значень показників для унікальності певного сайту. Вагові коефіцієнти нормуються за формулою:

$$\gamma_{0l} = \frac{\gamma_l}{\sum_{l=1}^k \gamma_l} \quad (2)$$

де γ_l – поточне (ненормоване) значення вагового коефіцієнта.

Надалі здійснюється послідовне (у межах кожної з двох груп факторів) зведення часткових критеріїв до узагальнених за групою (за (1)) та до інтегрованого критерію ефективності. Цим реалізується технологія вкладених згорток і забезпечується чутливість кінцевого рішення до зміни кожного з 12 часткових критеріїв таблиці 1. Приймаючи надалі для спрощення запису в (1) вагові коефіцієнти за одиниці з позначенням нормувань критеріїв нулем, отримуємо узагальнені критерії ефективності за кожною групою:

$$\begin{aligned} F_1 &= \gamma_{0K_A} (1 - K_{A0})^{-1} + \gamma_{0K_B} (1 - K_{B0})^{-1} + \gamma_{0K_C} (1 - K_{C0})^{-1} + \\ &+ \gamma_{0K_D} (1 - K_{D0})^{-1} + \gamma_{0K_E} (1 - K_{E0})^{-1} + \gamma_{0K_F} (1 - K_{F0})^{-1} \rightarrow \min, \\ F_2 &= \gamma_{0S_Y} (1 - S_{Y0})^{-1} + \gamma_{0S_G} (1 - S_{G0})^{-1} + \gamma_{0S_Q} (1 - S_{Q0})^{-1} + \\ &+ \gamma_{0S_{PR}} (1 - S_{PR0})^{-1} + \gamma_{0S_K} (1 - S_{K0})^{-1} + \gamma_{0S_B} (1 - S_{B0})^{-1} \rightarrow \min. \end{aligned} \quad (3)$$

Для визначення інтегрованої оцінки ефективності за дискретною згорткою (1) із узагальнених критеріїв (3) здійснюється їх нормування відносно найгіршої оцінки (максимального значення показника, що характеризує частковий критерій). З урахуванням зазначеного інтегрована оцінка ефективності інтернет-сайтів формується таким чином:

$$I = \gamma_{10} (1 - F_{10})^{-1} + \gamma_{20} (1 - F_{20})^{-1} \rightarrow \min \quad (4)$$

III. Інтерпретація отриманого рішення полягає в приведенні значення інтегрованої оцінки (5) до єдиної шкали зміни, наприклад, від 0 (найгірша) до 1 (найкраща). Це досягається шляхом нормування (4) до абстрактної найгіршої оцінки найгіршого сайту, що дає змогу: 1) здійснювати агрегацію часткових критеріїв до інтегрованого вигляду за згорткою; 2) встановлювати значення критеріїв оцінки в межах від 0 до 1 з найкращим значенням 0, найгіршим 1. Останнє дає змогу оцінювати одини окремих сайт з подальшим порівнянням декількох:

$$I_0 = 1 - \frac{I}{\max I}, \max I = \sum_{i=1}^2 (1 - [\max F_i - \Delta])^{-1}, \quad (5)$$

де $\max F_i$ – найгірше з можливих значень часткового показника; D – коефіцієнт запасу, що забезпечує уникнення некоректних операцій при нормуванні, приймається як $D=0,1-0,3$.

Вирази (3)–(5) становлять математичну модель, яка ґрунтується на послідовному угрупованні частинних критеріїв до інтегрованої оцінки ефективності інтернет-сайтів. Отриману числову оцінку можливо привести до лінгвістичної категорії відповідно до фундаментальної шкали оцінювання, наприклад, у вигляді таблиці 2.

Таблиця 2 – Фундаментальна шкала оцінювання

Інтегрована оцінка ефективності I_0	Лінгвістична категорія ефективності
1,0–0,7	Дуже висока
0,7–0,6	Висока
0,6–0,5	Вище середнього
0,5–0,4	Середня
0,4–0,3	Нижче середнього
0,3–0,2	Низька
0,2 і менше	Дуже низька

Приклад застосування оцінювання ефективності[4]. В інтернет-середовищі розміщено множину різноманітної інформації, оформленої у вигляді окремих сторінок, що логічно пов'язані між собою. Ця інформація становить собою інтернет-сайти. Для доведення дієвості запропонованого підходу проводилися розрахунки за початковими даними таблиці 3 для інтернет-сайтів трьох Веб-мані, Приватбанк та Простокредит.

Таблиця 3 – Початкові дані для розрахунків

№ з/п гру пи	№ з/п фактора	Показник	Діапазон зміни		Поточне значення	Нормуюче значення
			min	max		
1	1	K_A	259 хв.	444 хв.	262 хв.	259 хв. $-\Delta$
	2	K_B	3,9	5,7	3,9	3,9 $-\Delta$
	3	K_C	178,54 кб/с	333,31 кб/с	333,31 кб/с	178,54 кб/с $-\Delta$
	4	K_D	2,41 с	4 с	3,94 с	4 с $+\Delta$
	5	K_E	18	291	291	18 $-\Delta$
	6	K_F	12	41	14	41 $+\Delta$
2	7	S_Y	89 000	407 000	135 000	89 000 $-\Delta$
	8	S_G	265 000	437 000	265 000	265 000 $-\Delta$
	9	S_Q	380	2300	1200	380 $-\Delta$
	10	SPR	6	7	6	6 $-\Delta$
	11	S_K	79 102	349 358	349 358	349 358 $+\Delta$
	12	S_B	565	1 733	565	565 $-\Delta$

Порядок розрахунків щодо формування узагальнених значень показників показано лише для першої категорії критеріїв ефективності показників Юзабіліті з початковими даними для НАУ. Відповідно до (3) маємо модель розрахунку узагальненої оцінки для першої групи факторів, де спочатку проводиться нормування часткових критеріїв:

$$F_{oi}^{\min} = \frac{F_i^{\min}}{\max F_i^{\min} + \Delta}, \quad F_{oi}^{\max} = \frac{\min F_i^{\max} - \Delta}{F_i^{\max}},$$

$$K_{A0} = \frac{259 - \Delta}{262} = 0,9874, \quad K_{B0} = \frac{3,9 - \Delta}{3,9} = 0,923, \quad K_{C0} = \frac{178,54 - \Delta}{333,31} = 0,5348,$$

$$K_{D0} = \frac{3,94}{4 + \Delta} = 0,9163, \quad K_{E0} = \frac{18 - \Delta}{291} = 0,0608, \quad K_{F0} = \frac{14}{41 + \Delta} = 0,339, \quad \Delta = 0,3.$$

Нормування частинних показників для узагальненої оцінки здійснюється таким чином:

$$K_{A0}^{\min} = \frac{259 - \Delta}{259} = 0,9988, \quad K_{B0}^{\min} = \frac{3,9 - \Delta}{3,9} = 0,923, \quad K_{C0}^{\min} = \frac{178,54 - \Delta}{178,54} = 0,9983,$$

$$K_{D0}^{\max} = \frac{4}{4 + \Delta} = 0,93, \quad K_{E0}^{\min} = \frac{18 - \Delta}{18} = 0,9833, \quad K_{F0}^{\max} = \frac{41}{41 + \Delta} = 0,9927, \quad \Delta = 0,3.$$

Відповідно до визначених початкових даних, з використанням виразів (3)–(5) отримано значення інтегрованої оцінки ефективності інтернет-сайтів, яке становить $I_0 = 0,7212$. Розрахунок інтегрованої оцінки ефективності для решти двох інформаційних ресурсів проводився аналогічним чином, його результати висвітлено в таблиці 4.

Таблиця 4 – Інтегровані оцінки ефективності

	КПІ	НАУ	КНУ
I_0	0,7463	0,7212	0,6824

Лінгвістична категорія ефективності згідно з фундаментальною шкалою оцінювання для обраного сайту має «високий» рівень. Отже, слід позитивно оцінити інтернет-сайт «Приватбанк», порівняно з іншим запропонованим інформаційним ресурсом. Із розрахункових даних таблиці 4 випливає, що, порівняно з міжінформаційними ресурсами, кращі показники має «Веб мані» зі значенням інтегрованої оцінки $I_0 = 0,7463$. Проведені розрахунки доводять дієвість запропонованої моделі багатокритерійного оцінювання.

Висновки

Було сформовано математичну оптимізаційну модель оцінювання ефективності веб-пошуку. Вона відрізняється від відомих аналогів використанням запропонованої системи факторів, показників і критеріїв, розбитих на дві групи за належністю до певної категорії ефективності. З метою врахування всіх критеріїв у результаті оцінювання застосовано технологію вкладених згорток. Формування моделі ґрунтується на використанні методів зв'язку задач, зокрема методу зведення її до форми шляхом згортки часткових критеріїв. Приклад використання розробленої моделі підтвердив її придатність для вирішення практичних завдань оцінювання ефективності веб-пошуку. Дану модель можна застосовувати для унітарної та порівняльної оцінки.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Солошенко М. В. Методика оценки экономической ценности Internet сайта как средства коммуникации / М.В. Солошенко // Маркетинг и маркетинговые исследования. – 2000. – № 6. – С. 2–5.
2. Iskovych–Lototsky R. D., Zelinska O. V., Ivanchuk Y. V., Veselovska N. R. Development of the evaluation model of technological parameters of shaping workpieces from powder materials // Eastern–European Journal of Enterprise Technologies. Engineering technological systems. – 2017. – Vol. 1, № 1(85). – P. 9–17. doi: 10.15587/1729-4061.2017.59418.
3. Іскович–Лотоцький Р. Д. Моделювання робочих процесів в піролізній установці для утилізації відходів / Р. Д. Іскович–Лотоцький, Я. В. Іванчук, Я. П. Веселовський // Східно–європейський журнал передових технологій. – Харків, 2016. – Том 1, № 8(79). – С.11–20. doi: 10.15587/1729-4061.2016.59419.
4. Іскович–Лотоцький Р. Д. Математичне моделювання робочих процесів інерційного вібропрес–молота з електрогідравлічною системою керування гідроімпульсного привода для формоутворення заготовок з порошкових матеріалів // Р.Д. Іскович–Лотоцький, В.П. Міськов, Я.В. Іванчук // Вісник Хмельницького національного університету. Серія: Технічні науки. – 2016, – №3(237). – С. 176 – 180.

Вознюк Роман Володимирович — студент групи КН-19м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, e-mail: vozroman@ukr.net.

Науковий керівник: **Сілагін Олексій Віталійович** — к-т техн. наук, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця

Vozniuk Roman V. — Department of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, email : vozroman@ukr.net.

Supervisor: **Silagin Olexsiy V.** — Ph. D., Associate Professor, Associate Professor of Computer Scienses, Vinnytsia National Technical University, Vinnytsia.