

В. Ю. Кучерук, д. т. н., проф.; М. В. Глушко

КВАЛІМЕТРИЧНИЙ МЕТОД АНАЛІЗУ ЯКОСТІ ВІДГУКІВ «THE VALUE OF OPINION» ЯК ФУНДАМЕНТ СУЧАСНИХ РЕКОМЕНДАЦІЙНИХ СИСТЕМ

В статті розглянуто актуальність теми кваліметричного методу аналізу відгуків. Оскільки в 2021 році спільно із ростом e-commerce сфери, де більшість покупок відбувається після аналізу відгуків покупцем, в яких часто використовуються рекомендаційні системи. Проаналізовано основні типи психологічних портретів покупців в інтернеті.

Проаналізовано основні типи рекомендаційних систем та детально розглянуті недоліки кожної з них. Сформована мета дослідження, яка полягає в покращенні якості та систематизації її параметрів аналізу відгуків для розробки рекомендаційних систем на підставі кваліметричного методу та засобу. Проведено наукове обґрунтування на основі кваліметричного підходу аналізу відгуків до алгоритму рекомендації на основі "думки" користувача – "The Value of Opinion" та запровадження кваліметричного методу «Циклограма якості відгуків "The Value of Opinion"».

Сформована наукова задача, яка включає: систематизацію показників якості відгуків на підставі кваліметричного методу; розробку нових показників якості відгуків на підставі кваліметричного методу; можливість внесення показника "The Value of Opinion" в нову рекомендаційну систему колаборативної фільтрації, яка включає оцінку "думки" користувача.

Охарактеризовано доцільність використання кваліметричного методу оцінювання якості відгуків. Описана важливість аналізу відгуків про автомобілі. Розглянуті параметри кваліметрії та задачі, які ми вирішуємо за допомогою цих параметрів в циклограмі якості.

В цій науковій роботі розглянуто новий кваліметричний підхід до вимірювання якості відгуків як частини даних, на яких будуть базуватись алгоритми сучасних рекомендаційних систем. Запропонована система оцінки якості відгуків у форматі «Циклограма якості відгуків "The Value of Opinion"» має набір із восьми показників якості відгуків про товар. Описано важливість дослідження для бізнес-задач. Сформовані висновки та фінальні заключення по науковій роботі.

Ключові слова: відгук, рекомендаційна система, колаборативна фільтрація, якість, кваліметрія, метрика.

Вступ

Рекомендаційна система це підклас системи фільтрації інформації, яка зазвичай прагне передбачити "рейтинг" або "уподобання", які користувач надав би предмету. Вони в основному використовуються в комерційних програмах. Якість рекомендаційних систем можна визначити, оцінивши якість роботи алгоритмів, якість рекомендацій [1].

В сучасних компаніях, переважно інтернет-магазинах, велику увагу приділяють розділу відгуків про товари. Спираючись на те, що відгуки це довіра до бренду та продукту, то використання цього типу контенту є потужним інструментом маркетингу в галузі аналізу даних, які можна використовувати не тільки в рамках покращення конверсії, але й як джерело цінних даних для рекомендаційних систем.

Актуальність теми. Враховуючи, що рекомендаційна система це програма, яка на основі даних про користувача (User) і предмета (Item) дає рекомендації, така система включає в себе весь процес від отримання інформації до її надання користувачеві. Важливий кожен етап від інформації яку збирають, залежить, які алгоритми можуть застосовуватись. Ефективні алгоритми дають якісні, корисні рекомендації. Критерії оцінки результату дозволяють вибрати найбільш точні алгоритми. Рекомендаційні системи використовують переважно в e-commerce компаніях [2].

Переважно всі ІТ-компанії в сфері «e-commerce» використовують рекомендаційні системи, перша компанія яка реалізувала такі алгоритми – Amazon. Найбільші маркетплейси

в Україні, які використовують рекомендаційні системи: Rozetka, Prom, Allo, Bigl та Epicentr. Також, у всіх великих інтернет-магазинах є відгуки про товари. Найбільшу кількість відгуків має Rozetka – маркетплейс №1 в Україні (рис. 1). Проте, наскільки відомо із мережі інтернет, вказані компанії не використовують рекомендаційні системи із врахуванням відгуків про товар. Використовують відгуки як елемент довіри та впливу на прийняття рішень покупця аналітичного психотипу.

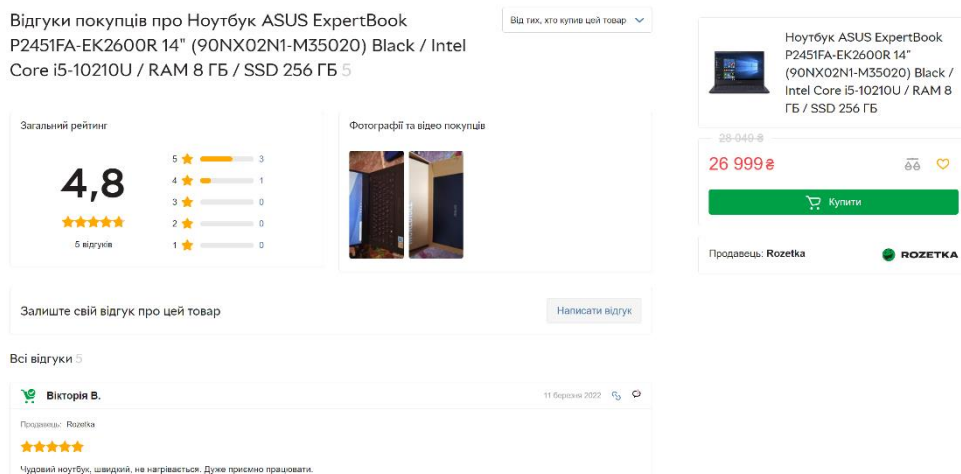


Рис. 1. Приклад відгуків про товар на Rozetka.ua – маркетплейсу із найбільшою кількістю відгуків

Кількість відгуків в мережі постійно росте, про це свідчать дані міжнародних статистичних досліджень. Для прикладу, на рис. 2 зображено кількість відгуків, створених користувачами на Goodreads (ресурс із відгуками про книги) з лютого 2012 року по липень 2019 року [3].

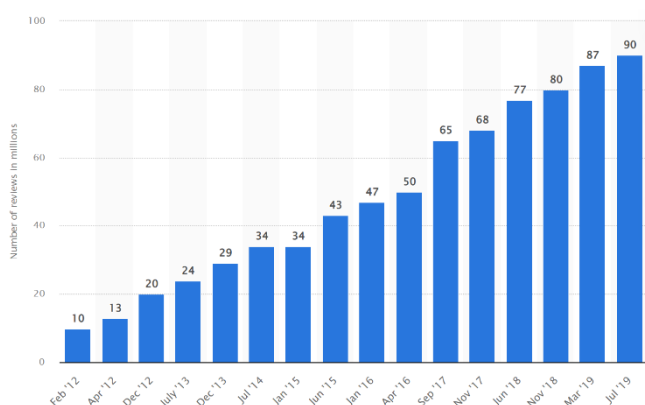


Рис. 2. Кількість створених користувачами відгуків на Goodreads 2012 – 2019 (млн)

Спираючись на аналітичні дані компанії Statista можемо побачити скільки відгуків читають користувачі, аналізуючи продукти і товари в Інтернеті [4]. Одразу бачимо, що відгуки важливі в першу чергу для молодшої аудиторії (рис. 3).

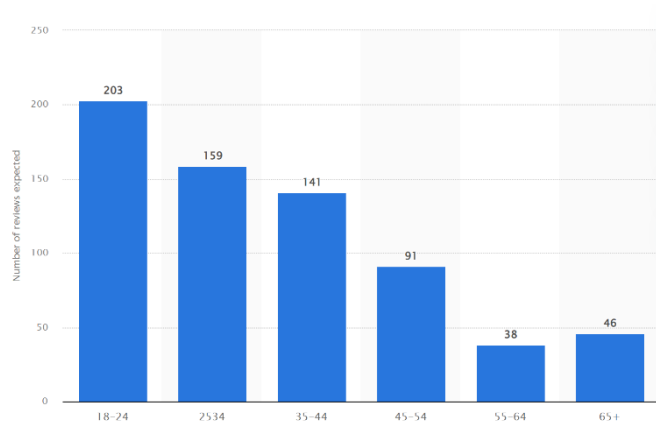


Рис. 3. Статистика по кількості переглянутих відгуків перед покупкою в інтернеті за віком

За даними Statista 40,5% покупців купують на Amazon в Сполучених Штатах станом на січень 2020 року, оскільки там є відгуки про товари та рекомендації (рис. 4) [5].

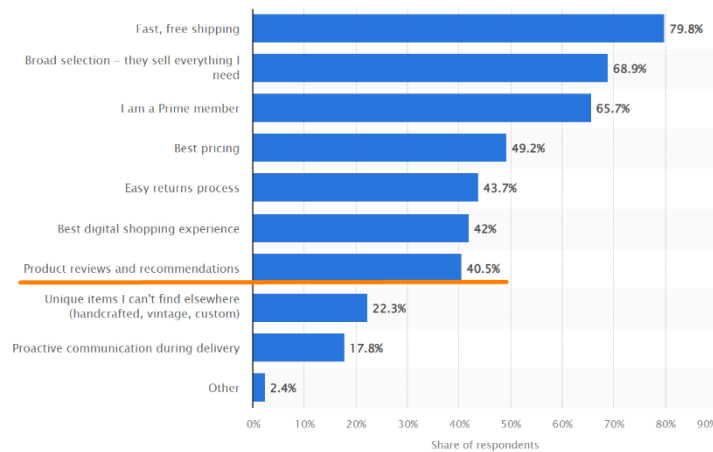


Рис. 4. Кількість покупців Amazon, які купують через наявність відгуків та рекомендації

Також, актуальність теми полягає в тому, що 11 квітня 2021 року пошукова система Google оновила ядро пошукового алгоритму під назвою «Google product reviews update 2021» [6]. Цей алгоритм фундаментально змінює принципи ранжування сайтів, оскільки в першу чергу враховує відгуки про продукти, а вже потім характеристики. Тобто, інформація від покупця тепер не менш важлива, ніж від продавця (опис товару). Відгуки про продукти в пошуку відповідають певним критеріям, як-от:

- Корисні детальні відомості, як-от переваги чи недоліки певного товару, особливості роботи продукту чи відмінність продукту від попередніх версій.
- Відгуки повинні бути від реальних людей, які дійсно використовували продукти, і показують, яким є продукт фізично або як він використовується.
- Відгуки повинні мати унікальну інформацію, крім тієї, що надає виробник, наприклад візуальні зображення, аудіо чи посилання на інший вміст, де детально описується досвід рецензента.
- Відгуки повинні містити порівняння продуктів, або пояснення, чим продукт відрізняється від продуктів конкурентів.

Інформація про відгуки також відображається під час ранжування сайтів в Google. Приклад взято з об'єкта дослідження Automoto.ua (рис. 5), проте для відображення відгуків на сайті повинні бути спеціальні фрагменти програмного коду у форматі JSON –

мікророзмітка даних, на основі якої пошукові системи розрізняють типи контенту, в цьому випадку – відгуки (Reviews).

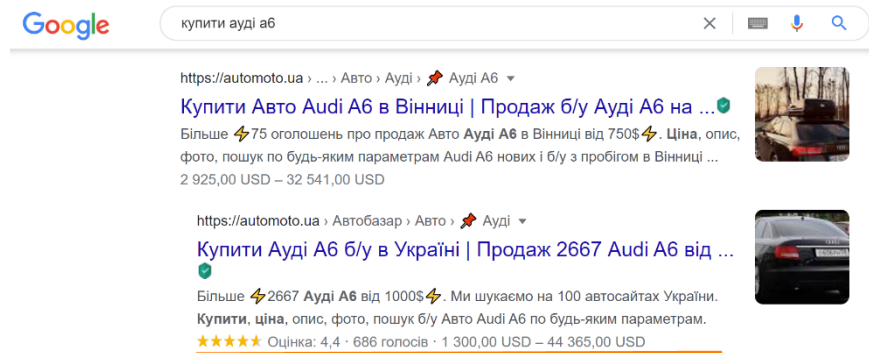


Рис. 5. Відображення інформації про відгуки користувачів про автомобілі

Мета та наукові задачі. Метою дослідження є систематизація параметрів якості відгуків на підставі кваліметричних методів та засобів отримання показників для того, щоб покращити рекомендаційні системи. Наукове обґрунтування полягає у відсутності єдиних прийнятих показників якості відгуків, які можуть бути застосовані для покращення якості рекомендаційних систем.

Для реалізації окресленої мети, необхідно виконати такі завдання:

- провести аналіз рекомендаційних систем із використанням відгуків та наукових публікацій щодо поставленої мети;
- сформулювати принципи визначення якості відгуків, які можуть бути використані для покращення рекомендацій;
- дослідити основні підходи до структурування характеристик оцінювання якості відгуків;
- запропонувати кваліметричний метод оцінки якості відгуків для оцінки “думки” користувача – «Циклограма якості відгуків “The Value of Opinion”».

Кваліметрія (англ. qualimetry, нім. Qualimetrie f) (лат. Quales – якість + грец. Μέτρον – міряю) — наука про оцінку якості об’єктів, вивчає та реалізує методи і засоби кількісної оцінки якості продукції. Сьогодні кваліметрію розглядають як частину теорії прийняття рішень. Якість відгуків ми визначаємо саме за рахунок кваліметричних методів вимірювання [7].

Об’єктом дослідження є сайт-агрегатор пошуку автомобілів в – Automoto.ua. Це пошукова система, спеціалізацією якої є пошук оголошень про продаж автомобілів, мотоциклів, спецтехніки та інших транспортних засобів в Україні. Automoto.ua дає можливість здійснювати пошук пропозицій про продаж авто по всій Україні, надаючи максимально повні та актуальні результати. Сьогодні сайт обробляє інформацію з понад 100 автосайтів України. Щодня в базі розміщують більше 250 000 тисяч оголошень, 9 – 16 тисяч з яких є «свіжими» надходженнями за поточний день.

Предметом дослідження є кваліметричний метод аналізу відгуків.

Результати дослідження

На основі попереднього дослідження алгоритму рекомендаційної системи було використано метод колаборативної фільтрації на основі користувачів (user to user) [8]. Але цей метод має дві основні проблеми:

1. Винятковість даних: У випадку наявності великої кількості елементів, з якими

взаємодіяв користувач, їх кількість зменшується до невеликого відсотка, що робить коефіцієнт кореляції менш надійним.

2. Користувачі швидко змінюються (знаходять потрібний автомобіль та більше не шукають), і всю модель системи потрібно було перерахувати, що є затратно по часу і за обчисленням дорогим [9].

Для вирішення цих проблем ми використали колаборативну фільтрацію item-to-item та отримали приріст клікабельності на 22% в середньому. Але проблема системи полягає в тому, що вона аналізує тільки поведінку «потенційних покупців», але не враховує дані від «реальних власників» (не продавців) [10]. Тому, пропонується створити третій новий алгоритм колаборативної фільтрації, який буде враховувати «думку» користувачів про автомобілі. Мається на увазі зібрати відгуки про автомобілі із десятків автомобільних сайтів України та світу, та конвертувати текст відгуку в числове значення «The Value of Opinion». Алгоритм пропонує наступні покращення: створення найбільшого в світі каталогу відгуків про автомобілі (до 3 мільйонів відгуків); аналіз всього відгуку, а не тільки блоків «плюси» та «мінуси»; набір із 8 кваліметричних метрик, які комплексно оцінюють якість відгуку та формують комплексне значення «The Value of Opinion»; внесення показника «The Value of Opinion» в основну формулу ранжування рекомендаційних систем із найбільшою кількістю оголошень в Україні.

На основі наших наукових робіт та досвіду, проаналізувавши наукові роботи українських та закордонних вчених, ми виявили, що у алгоритмів рекомендацій, які включають аналіз відгуків, присутній ряд недоліків:

- Гонитва за покращенням точності, яка вимірюється 1 – 2 показниками, але не комплексно;
- Немає кореляції впливу досліджень на реальний бізнес;
- Немає кореляції впливу досліджень на донесення інформації клієнтам в пошуку товарів;
- Відсутність комплексної оцінки якості на основі кваліметричних досліджень;
- Відсутність систематизації та принципів визначення якості відгуків, які можуть бути використані для покращення рекомендацій.

Тому, в рамках кваліметричних методів вимірювань запроваджуємо «Циклограму якості відгуків» (рис. 6), де пропонується комплексний показник «The Value of Opinion» [9].

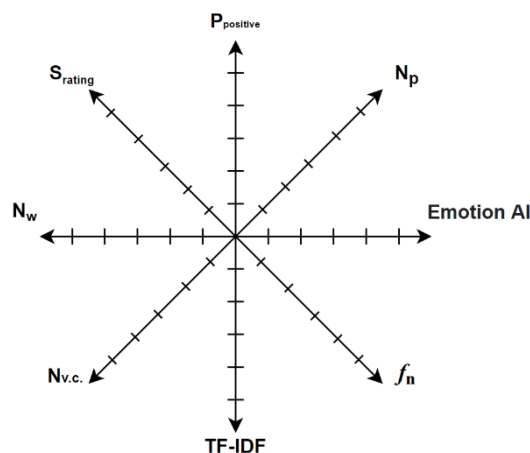


Рис. 6. Циклограма якості відгуку «The Value of Opinion» для побудови рекомендаційних систем

Розглянемо детальніше кожен метрику, яка характеризує якість відгуків:

N_w – кількість слів у написаному відгуку від власника. Чим більша кількість слів – тим

більш якісний відгук з точки зору повноти опису та характеристик про товар.

S_{rating} – середня оцінка продукту користувачем виставлена під час написання відгуку. Зазвичай оцінка ставиться від 1 до 5.

$P_{positive}$ – відсоток позитивних думок (слів, тверджень, словосполучень) про автомобіль. Показник необхідний для базового розуміння, чи сподобалась інформація власнику автомобіля.

$$P_{positive} = \frac{N_{\text{позитивних думок}}}{N_{\text{позитивних думок}} + N_{\text{негативних думок}}}; \quad (1)$$

Для вимірювання $P_{positive}$ необхідно створити семантичне ядро із порівно поділених позитивних та негативних слів, які будуть зчитуватись при аналізі кожного відгуку. Приклад наведено в таблиці 1.

Таблиця 1

	<i>Позитивна думка</i>	<i>Негативна думка</i>
1.	Витрати пального низькі	Витрати пального високі
2.	Двигун на 100 тис. км. ремонт не потребує	Двигун на 100 тис. км. потребує ремонту
3.	Коробка передач працює добре	Коробка передач працює погано
<i>N</i>
50.	Форсунки замінював 2 рази	Форсунки не замінював

N_p – загальна кількість відгуків про обрану марку/модель/рік авто, які вдалось знайти для рекомендацій. Цей показник необхідний для розуміння впливу кількості відгуків на загальну формулу ранжування. За малої кількості відгуків про конкретні моделі автомобілів вплив на рекомендації буде менший, оскільки зменшується репрезентативність даних.

Emotion AI – показник, пов'язаний з виявленням емоцій за допомогою штучного інтелекту, також відомий як афективні обчислення. Обчислення відноситься до аналізу тональності тексту на основі алгоритмів машинного навчання. І є класом методів контент-аналізу в комп'ютерній лінгвістиці, яке призначено для автоматизованого виявлення в текстах емоційно забарвленої лексики та емоційної оцінки об'єктів авторами.

Основною метою аналізу тональності є знаходження думок у тексті та виявлення їх властивостей. Від поставленого завдання залежить, які властивості будуть досліджуватись. Метою аналізу є автомобіль, відгук до якого написав власник, якому належить «думка». Думки поділяються на два типи: безпосередня думка та порівняння.

Безпосередня думка містить висловлювання автора про один об'єкт – автомобіль. Формальне визначення безпосередньої думки виглядає як сукупність із п'яти елементів (*e, f, op, h, t*), де:

- (*entity, feature*) – Об'єкт тональності *e* (сутність, щодо якої висловлюється автор) або його властивості *f* (атрибути, частини об'єкта);
- *orientation* або *polarity* – тональна оцінка (емоційна позиція автора щодо згаданої теми);
- *holder* – суб'єкт тональності (автор, тобто комусь належить ця думка);
- момент часу *time* коли було залишено думку.

Приклади тональних оцінок:

- позитивна;
- негативна;
- нейтральна.

Під «нейтральним» мається на увазі, що текст не містить емоційного забарвлення. Також можуть існувати інші тональні оцінки.

За допомогою алгоритмів машинного навчання SVM (*support vector machine*) та LDA (*Latent Dirichlet allocation*) визначаємо чи відгук дійсно відноситься до авто (рис. 7), а далі визначаємо наскільки думка негативна, чи позитивна за шкалою від 0 до 1.

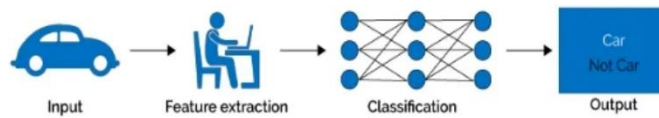


Рис. 7. Визначення ознак відгуку про автомобіль за допомогою машинного навчання

f_n – показник закону Ципфа («ранг-частотність»). Це емпірична закономірність розподілу частотності слів природної мови: якщо всі слова мови (або просто досить великого тексту) впорядкувати за зменшенням частотності їх використання, то частотність n -го слова в такому списку виявиться приблизно обернено пропорційною його порядковому номеру n , тобто рангу цього слова (рис. 8).

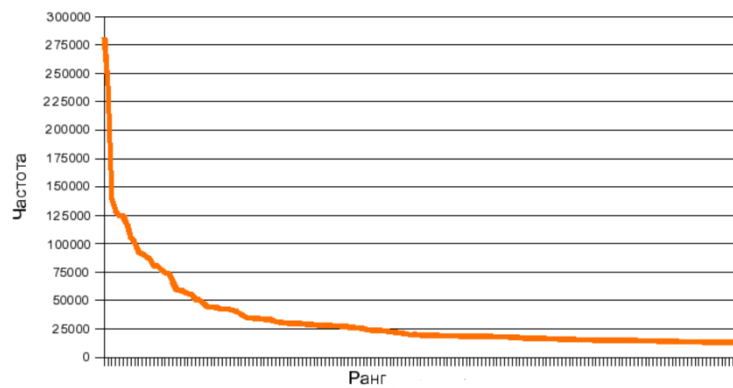


Рис. 8. Графік для частотності слів із статей з рангами від 3 до 170 за законом Ципфа

Наприклад, друге за використанням слово зустрічається приблизно вдвічі рідше, ніж перше, третє втричі рідше, ніж перше, і так далі. Формально: N – кількість елементів; k – їхній ранг; s – значення показника, що характеризує розподіл.

Потім закон Ципфа передбачає, що із сукупності з N елементів нормована частота елемента рангу k , $f(k; s, N)$, є:

$$f(k; s, N) = \frac{1/k^s}{\sum_{n=1}^N (1/n^s)} ; \tag{2}$$

TF-IDF (від англ. *TF – term frequency*, *IDF – inverse document frequency*) – статистичний показник, що використовується для оцінки важливості слів у контексті документа. Вага (значимість) слова пропорційна кількості вживань цього слова у документі, і обернено пропорційна частоті вживання слова у інших документах колекції.

Показник *TF-IDF* використовується в задачах аналізу текстів та інформаційного пошуку. Його можна застосовувати як один з критеріїв релевантності документа до пошукового запиту, а також при розрахунку міри спорідненості документів при кластеризації.

Кожний відгук про автомобіль має набір N ключових слів, які описують авто. Вимірювання показника важливості ключових слів у контексті відгуку пропорційно до вимірювання якості написаного відгуку.

TF (term frequency – частота слова) – відношення числа входжень обраного слова t_i до загальної кількості слів відгуку. Таким чином, оцінюється важливість слова в межах

обраного відгуку.

$$TF = \frac{n_i}{\sum_k n_k}; \quad (3)$$

де n_i є число входжень слова в документ, а в знаменнику – загальна кількість слів в документі.

IDF (*inverse document frequency – обернена частота документа*) – інверсія частоти, з якою слово зустрічається в каталозі відгуків. Використання IDF зменшує вагу широковживаних слів.

$$IDF = \frac{|D|}{|(d_i \supset t_i)|}; \quad (4)$$

де

- $|D|$ – кількість відгуків каталогу;
- $|(d_i \supset t_i)|$ – кількість відгуків, в яких зустрічається слово t_i (коли $n_i \neq 0$).

Вибір основи логарифму у формулі не має значення, адже зміна основи призведе до зміни ваги кожного слова на постійний множник, тобто вагове співвідношення залишиться незмінним.

Іншими словами, показник TF-IDF це добуток двох множників: TF та IDF.

$$TF - IDF = TF \cdot IDF; \quad (5)$$

Більшу вагу TF-IDF отримують слова з високою частотою появи в межах відгуку та низькою частотою вживання в інших відгуках каталогів.

$N_{v.c.}$ – (*number of visual content*) це бінарне значення на рівні відгука, та кількісне на рівні каталогу відгуків, яке відповідає за наявність візуального контенту (фото, відео), який надає додаткову цінність та якість відгуку для користувача, та є фактором впливу Google при аналізі відгуків.

Висновки

Встановлено, що не існує єдиної системи оцінки якості відгуків, які би використовувались для покращення рекомендаційних систем. Також, наявність відгуків на сайті використовується переважно як інструмент маркетингу та довіри, про що свідчить аналіз відкритих даних про компанії та рекомендаційні системи, які використовують.

Принципи визначення якості відгуків, а саме набір метрик, які є кваліметричною основою якості: N_w , S_{rating} , $P_{positive}$, N_p , $Emotion AI$, f_n , $TF-ID$, $N_{v.c.}$ сформовані. Вказані 8 метрик, формують «Циклограму якості відгуків “The Value of Opinion”». Крім запропонованого кваліметричного підходу вимірювання, циклограма включає в себе новий показник – «*Emotion AI*», який базується на алгоритмах машинного навчання SVM та LDA для аналізу тональності текстів, і дасть змогу проаналізувати емоції власника автомобіля та визначити його «думку», яка подається в форматі числового показника.

Запропонована «Циклограма якості відгуків “The Value of Opinion”» може використовуватись для аналізу будь-яких відгуків про товари в інтернеті в рамках покращення якості рекомендаційних систем, та користувацького досвіду при покупках в цілому, а отже є універсальною для використання.

СПИСОК ЛІТЕРАТУРИ

1. Виявлення 4XX та 5XX помилок на сайті як впливний фактор на результат точності рекомендаційних систем [Електронний ресурс] / Volodymyr Kucheruk, Mykhailo Hlushko // Conference <https://conferences.vntu.edu.ua>. – 2019. – Режим доступу : <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2019>.
2. Покращення алгоритму "item to item" методу колаборативної фільтрації для розробки рекомендаційних систем на основі косинусної міри шляхом оцінки релевантності [Електронний ресурс] / Volodymyr Kucheruk, Mykhailo Hlushko // ScienceRise. – 2018. – № 1. – С. 20 – 24. – Режим доступу : http://nbuv.gov.ua/UJRN/text_2018_1_6.
3. Number of user-generated reviews on Goodreads from February 2012 to July 2019 [Електронний ресурс] // Simicart.com. – 2019. – Режим доступу до ресурсу: <https://www.statista.com/statistics/252998/number-of-user-generated-reviews-on-goodreadscom/>.
4. How many reviews do you expect when looking at a product online? [Електронний ресурс] // Statista.com. – 2021. – Режим доступу : <https://www.statista.com/statistics/1019495/online-shoppers-expectations-product-reviews-in-the-us/>.
5. Reasons for internet users in the United States to shop on Amazon as of January 2020 [Електронний ресурс] // Statista.com. – 2021. – Режим доступу : <https://www.statista.com/statistics/670499/us-amazon-usage-reason/>.
6. Google March 2022 product reviews update is finished rolling out [Електронний ресурс] // Statista.com. – 2021. – Режим доступу : <https://searchengineland.com/google-march-2022-product-reviews-update-is-finished-rolling-out-383499>.
7. Квалиметрия: методы количественного оценивания качества различных объектов (курс лекций и практических занятий): учеб. пособие / [Астратова Г. В., Латыпова Л. В., Шапошников В. А. и др.]. – ГОУ ВПО ХМАО – Югры «Сургут. гос. пед. ун-т». – Сургут : РИО СурГПУ, 2014. – 160 с. – ISBN 978-5-93190-321-7.
8. Sci-conf.com [Електронний ресурс] / Volodymyr Kucheruk, Mykhailo Hlushko // Режим доступу : <https://sci-conf.com.ua/ii-mezhdunarodnaya-nauchno-prakticheskaya-konferentsiya-actual-trends-of-modern-scientific-research-16-18-avgusta-2020-goda-myunhen-germaniya-arhiv/>.
9. Improving Accuracy of Recommender systems based on Collaborative Filtering Algorithm Item-To-Item [Електронний ресурс] / Volodymyr Kucheruk, Mykhailo Hlushko // International Science Group : Technical research and development. Технічні дослідження і розробки (Технічні науки) – Primedia eLaunch LLC, США, 2021. – Режим доступу : <https://isg-konf.com/wp-content/uploads/2021/04/Monograph-USA-Technical-2021-I-isg-konf.pdf>.
10. Методи оцінки якості рекомендаційних систем [Електронний ресурс] / В. Ю. Кучерук, М. В. Глушко // «Конференції ВНТУ електронні наукові видання, Науково-технічна конференція факультету комп'ютерних систем і автоматики (2021)» 03 березня 2021 р. – Режим доступу : <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2021/paper/view/12509>.

Рекомендована кафедрою Метрології та промислової автоматики.

Стаття надійшла до редакції 18.09.2021.

Стаття пройшла рецензування 26.09.2021.

Кучерук Володимир Юрійович – д. т. н., професор, завідувач кафедри метрології та промислової автоматики, факультет комп'ютерних систем і автоматики.

Глушко Михайло Васильович – студент-аспірант кафедри метрології та промислової автоматики, факультет комп'ютерних систем і автоматики, e-mail: Mikhail.Hlushko@gmail.com.

Вінницький національний технічний університет.