

ДОСЛІДЖЕННЯ ТА СТВОРЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ АНАЛІЗУ ВІДГУКІВ КОРИСТУВАЧІВ ІНТЕРНЕТ СЕРВІСІВ

Запорізька державна інженерна академія, кафедра ПЗАС

Ми живемо в час стрімкого розвитку Інтернету та інформаційних технологій. Все частіше з'являються сервіси, сайти та системи які надають змогу людині виконати певні дії вдома, не виходячи на вулицю. Прикладами таких дій можуть бути Інтернет покупки, сплата комунальних послуг, покупка квитків. Важливою частиною цих сервісів є зворотній зв'язок з користувачами, завдяки якому сервіс розвивається, покращується. Вся робота з користувачами зводиться до відгуків, на базі яких вносяться якісь зміни зі сторони розробників, та будується думка про сервіс чи товар зі сторони інших користувачів. Ми можемо зустріти відгуки будь де в Інтернеті, будь то фільм, товар в магазині, чи продавець на торгівельній інтернет площадці. Вони допомагають іншим користувачам визначитись чи хочуть вони переглянути цей фільм, чи купити цей товар або краще вибрати інший, чи замовити щось у цього продавця або вибрати іншого. Також відгуки позитивно впливають на сам сервіс чи сайт, так як покращують його ранжованість у пошукових сервісів, так як розширюють вміст даної сторінки тими чи іншими критеріями та словами. Системи аналізів відгуків зазвичай створюються для внутрішнього використання у сервісах та сайтах, для кращої оцінки стану справ. А також вони використовуються в різних розширеннях від торгових Інтернет площадок для браузерів, для створення рейтингу продавців, по якому користувач розуміє чи варто купувати щось у даного продавця, чи краще купити такий же товар у іншого.

Статистика свідчить:

- За даними Reevo, 50 і більше відгуків про товар дозволяють підняти конверсію на 4,6%.
- 63% споживачів віддають перевагу Інтернет-магазинам, в яких товари супроводжуються відгуками (дані iPerceptions).
- Відвідувачі, які читають відгуки в Інтернет-магазині і використовують можливість задати питання про товар, оформляють замовлення вдвічі частіше «не допитливих», при цьому грошей вони витрачають на 11% більше (дані Bazaarvoice, Conversation Index).
- Рівень довіри до споживацьких відгуків майже в 12 разів вище рівня довіри до опису товару (дані eMarketer).
- Статистика, зібрана і проаналізована Reevo, показує, що відгуки товарів дозволяють підвищити продажі в середньому на 18%.

Пошукові системи з великою симпатією ставляться до сайтів, які часто оновлюються, особливо, якщо ці оновлення – це публікація унікального контенту. Відгуки користувачів – прекрасний приклад унікального наповнення інтернет-магазину. Багато власників і менеджерів Інтернет-магазинів не бажають працювати на власним описом товарів (або платити за них) і використовують те, що представляє постачальник/виробник. В нішах, повних сторінок з однаковими текстами, піднятися в пошуковій видачі допомагають користувачські відгуки про товари. Контент, що генерується в вигляді відгуків про товари, позитивно впливає на позиції по низькочастотним запитам.

Метою роботи є створення системи аналізу відгуків для дослідження їх впливу на прийняття рішення про покупку чи використання людиною даного сервісу, товару, сайту, та методів аналізу відгуків. Дослідження впливу позитивних та негативних відгуків на самі інтернет сервіси та сайти і площадки. Аналіз алгоритмів інтелектуальної автоматизованої обробки текстової інформації. Поділ інформації на класи та за емоційним відтінком.

Система поділяється на дві функціональні частини:

1. Парсер.
2. Аналізатор відгуків.

Для того щоб результати роботи програми були як можна більш правильними потрібна велика кількість даних, котра має бути отримана з різних Інтернет сервісів.

Парсер реалізує процес отримання інформації про товар і власне самих відгуків про них та збереження цієї інформації в базі даних для зручного їх використання в майбутньому, що дозволить уникнути часто повторюваного процесу парсингу даних, який є довготривалим та залежним від ресурсів системи.

Після дослідження предметної області було визначено, що актуальною для аналізу та класифікації відгуку буде задача визначення тональності тексту відгуків.

Аналізатор відгуків використовує методи машинного навчання для вирішення проблем, пов'язаних з обробкою природної мови (NLP). Однією з таких проблем є автоматичне визначення емоційного забарвлення (позитивний, негативний, нейтральний) текстових даних, тобто аналізу тональності (sentiment analysis).

Окрім цього аналізатор виконує функцію представлення. Тобто він відображає результати аналізу, виконаного на попередньому кроці, у зрозумілому для людей вигляді, тим самим полегшуючи процес вибору товару чи сервісу.

Таким чином створений комплекс програмних засобів дає змогу виконати швидкий аналіз великої кількості даних, їх класифікації та представлення. Отримані результати можуть бути використані як звичайним користувачем так і власниками / менеджерами Інтернет сервісів, сайтів, площадок Інтернет продажів. Система отримання відгуків (парсингу) може бути розширена для роботи з багатьма веб-ресурсами і використовуватись як окремий програмний продукт.

Висновки:

- Проведено дослідження проблеми збору відгуків користувачів Інтернет ресурсів, їх аналізу та відображення інформації.
- Запропоновано метод вирішення цієї проблеми шляхом створення програмної системи збору та інтелектуальної обробки відгуків.
- Розроблено архітектуру системи, яка складається з парсера та аналізатора відгуків.
- За допомогою розробленої системи створена база даних відгуків та досліджено вплив відгуків на користувача, а також на власників і менеджерів Інтернет-сервісів, сайтів, Інтернет-площадок для продажів.
- Виявлено прямий зв'язок наявності та правильного представлення відгуків користувачів про товар чи послугу зі збільшенням відвідувань сайтів та користування послугами.
- Доведено актуальність розвитку та реалізації теми роботи в подальшому.

Література

1. I. Chetviorkin, P. Braslavskiy, N. Loukachevich, "SentimentAnalysisTrackatROMIP 2011," In Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2012", Bekasovo, 2012.
2. T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," In European Conference on Machine Learning (ECML), Springer Berlin/Heidelberg, 1998, pp. 137-142.
3. J. Furnkranz, T. Mitchell, and E. Riloff, «A Case Study in Using Linguistic Phrases for Text Categorization on the WWW,» In AAAI/ICML Workshop on Learning for Text Categorization, 1998, pp. 5-12.