УДК 004.8: 004.654: 339.378

ANALYSIS OF TRADING NETWORK MARKETING MEASURES BY TEXT MINING METHODS

S. Hrybkov, R. Khanbabaev, O. Kharkianen
National University of Food Technologies

Key words:
Data Mining,
Text Mining,
social networks

Article history:
Received 06.10.2020
Received in revised form 20.11.2020
Accepted 07.12.2020

Corresponding author:
sergio_nuft@i.ua

ABSTRACT

The article researches and analyzes the methods of Data Mining for the formation and evaluation of marketing activities of the trade network. Special attention is paid to the analysis of Text Mining technology. This is a toolkit developed on the basis of basic Data Mining technologies in combination with statistical and linguistic analysis, methods of artificial intelligence, which allows you to search for trends, patterns and relationships in unstructured texts for management decisions. Data mining is a process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.

Building a strategy of trade network development requires the study of the information field in which the company operates, an objective assessment of the activities, comparing bonus programs of its own network and competitors.

Despite the large number of information sources for their rapid processing and classification is indispensable use of software products implementing Text Mining technology. Text mining similar to text analytics, is the process of deriving high-quality information from text. Written resources may include websites, books, emails, reviews, and articles. High-quality information is typically obtained by devising patterns and trends by means such as statistical pattern learning.

Thus, in order to evaluate the success of the marketing activities of the trade network, information support is offered using the methods of text tone and search for anomalies Text Mining technology based on data from the social network Twitter.

An example of Text Mining models implementation in Orange component software package using Twitter API is given. Working models for the practical implementation of the trade network tasks are formed on the basis of a set of Orange widgets. The result is visualized in the form of diagrams and clouds of words, which are easy to view and analyze.

DOI: 10.24263/2225-2916-2020-28-19

© С. В. Грибков, Р. Р. Ханбабаєв, О. В. Харкянен, 2020

АНАЛІЗ МАРКЕТИНГОВИХ ЗАХОДІВ ТОРГОВЕЛЬНОЇ МЕРЕЖІ МЕТОДАМИ TEXT MINING

С. В. Грибков, канд. техн. наук
Р. Р. Ханбабаєв
О. В. Харкянен, канд. техн. наук
Національний університет харчових технологій

У статті досліджено та проаналізовано методи Data Mining для формування й оцінки маркетингових заходів торговельної мережі. Особливу увагу приділено аналізу технології Text Mining. Це інструментарій, розроблений на основі базових технологій Data Mining у поєднанні зі статистичним і лінгвістичним аналізами, методами штучного інтелекту, який дає змогу здійснювати пошук тенденцій, шаблонів і взаємозв'язків у неструктурованих текстах для прийняття управлінських рішень.
Наведено приклад реалізації моделей Text Mining у компонентному програмному пакеті Orange з використанням Twitter API. Робочі моделі для практичної реалізації завдань торговельної мережі сформовані на основі набору віджетів Orange. Результат дослідження візуалізовано у вигляді діаграм і хмар зі слів, які зручно переглядати й аналізувати.
Ключові слова: *інтелектуальний аналіз даних, Text Mining, соціальні мережі.*

Постановка проблеми. Ефективність роботи торговельної мережі значною мірою залежить від успішності маркетингових кампаній. Технологія Data Mining набуває все більшої популярності в галузі маркетингу. Вона забезпечує збір, обробку й отримання нових корисних знань шляхом перетворення наявного масиву даних у зрозумілу для людини структуру для прийняття рішень.

Інформація, отримана завдяки технології Data Mining, може допомогти збільшити рентабельність інвестицій (ROI), покращити управління відносинами з клієнтами (CRM) та аналіз ринку, зменшити витрати на маркетингові кампанії та допомогти утримати клієнтів [1].

Сучасна людина проводить багато часу в мережі Інтернет, соціальних мережах, месенджерах. При цьому майже кожна дія залишає за собою невеликі шматочки електронних даних. Кожен раз, коли використовується Інтернет або мобільний телефон, дані оброблюються, узагальнюються, зберігаються і стають основою для формування акційних пропозицій, підібраних під потреби окремого клієнта. Кожна діяльність з продажу товарів генерує певну кількість даних, які при правильному аналізі можуть забезпечити конкурентну перевагу на міжнародному ринку шляхом виявлення прихованих зв’язків між великими наборами даних. Одним із прийомів, які використовуються для аналізу великого обсягу даних, є обробка даних методами Data Mining [10].

З огляду на вищезазначене, актуальним завданням для торговельної мережі є визначення способів отримання та практичного використання маркетингової інформації для розробки нових пропозицій, оцінки проведених заходів із застосуванням сучасної технології Data Mining на основі даних соціальних мереж.

Метою статті є дослідження способів застосування методів інтелектуального аналізу даних у маркетингових завданнях торговельної мережі.

Методи і матеріали. Під час дослідження використано методи інтелектуального аналізу даних, зокрема метод Text Mining, розроблений на основі статистичного й лінгвістичного аналізів, методах штучного інтелекту і призначений для проведення аналізу, забезпечення навігації й пошуку в неструктурованих текстах.

Проектування структури даних здійснено в компонентному програмному пакеті Orange. Для аналізу використані короткі текстові повідомлення — твіти користувачів з реакцією на діяльність торговельних мереж, отримані з соціальної мережі Twitter, з використанням Twitter API.

Результати дослідження. Торговельні мережі здійснюють просування товарів за допомогою проведення різноманітних маркетингових кампаній у засобах масової інформації, на телебаченні, радіо, в Інтернеті, на щитах, застосовують зовнішню рекламу, рекламу на транспорті, надають споживачам акційні пропозиції, організують дегустації тощо. Після проведення маркетингової кампанії виникає необхідність оцінити результати, знайти слабкі та сильні сторони проведених заходів, визначити реакцію споживачів. Використання методу Text Mining технології Data Mining — це сучасний спосіб отримати, узагальнити й проаналізувати успішність маркетингової кампанії на основі реакції споживачів в мережі Інтернет, соціальних мережах, месенджерах [4—7].

У дослідженні проаналізовано маркетингову кампанію торговельної мережі «Сільпо», що проводилась на території України, а також здійснено порівняння акційних програм торговельних мереж на основі реакції споживачів у соціальній мережі Twitter. Дані отримано з відкритих джерел соціальної мережі Twitter з використанням розробленої моделі робочого процесу аналізу даних у програмному додатку Orange.

Twitter — це соціальна мережа мікроблогів, що надає можливість користувачам надсилати короткі текстові повідомлення (до 280 символів). Її головним призначенням є висловлення думок користувачів про сучасні події та явища з усіх сфер людської діяльності.

Twitter є важливим джерелом даних для бізнес-моделей різних компаній, мережа зручна у використанні як для обміну інформацією, так і для її збору. Маючи аккаунт розробника, можна отримати доступ до даних про пости, лайки, коментарі користувачів тощо. Одним із засобів збору й обробки інформації із Twitter є програмний додаток Orange.

Orange — це програмний пакет для візуального програмування на основі компонентів для візуалізації даних, машинного навчання, обміну й аналізу даних [3].

Проаналізуємо реакцію клієнтів на проведену маркетингову кампанію на прикладі торговельної мережі «Сільпо» засобами Text Mining. З метою реалізації методу Text Mining був отриманий і використаний акаунт розробника соціальної мережі Twitter.

Для проведення аналізу методом Text Mining створена модель робочого процесу у програмному додатку Orange, яка наведена на рис. 1.

Модель робочого процесу створена з використанням комбінації віджетів, що реалізують набір базових функцій і створюють структуру аналітики даних.

Віджет Twitter містить ключові слова, за якими проводиться пошук. Для аналізу проведеної маркетингової кампанії введені слова: «сільпо» та «реклама».

Віджет Preprocess Text розбиває текст на дрібні одиниці (лексеми), фільтрує їх, запускає нормалізацію (стематизацію, лематизацію), створює n-грами й теги токенів.

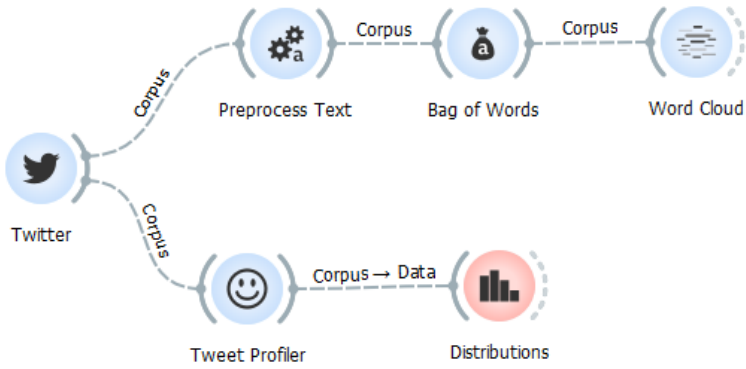

Рис. 1. Модель робочого процесу в Orange

Віджет Bag of Words потрібен для підрахунку частоти появи ключових слів. Він визначає кількість слів для кожного екземпляра даних (документа). Кількість може бути абсолютною, двійковою (міститься або не міститься слово в документі) або підлінійною (відношення числа входжень обраного слова до загальної кількості слів документа).

Віджет Tweet Profiler використовується для аналізу емоційного забарвлення твітів. Tweet Profiler отримує інформацію із сервера про емоції кожного твіту (або документа). Віджет надсилає дані на сервер, де модель Twitter обчислює ймовірні емоції та/або бали.

Для візуалізації результатів обробки необхідні два віджети: *Word Cloud* та *Distributions*.

Word Cloud відображає лексеми у вигляді хмари, їхній розмір означає частоту використання слова у тексті. Віджет Distributions відображає розподіл значень дискретних або безперервних атрибутів. Якщо дані містять змінну класу, розподіли можуть бути обумовлені класом. Для дискретних атрибутів графік, який відображається віджетом, показує скільки разів кожне значення атрибуту з'являється в даних. Якщо дані містять змінну класу, також будуть відображені розподіли класів для кожного зі значень атрибутів [8].

Візуалізація результатів здійснюється на основі отриманої хмари зі слів, представленої на рис. 2.

Рис. 2. Результат аналізу у вигляді хмари зі слів

У хмарі зі слів розмір літер слова позначає частоту його використання, тобто чим літери більше, тим частіше слово з'являється у постах з реакцією на маркетингову кампанію торговельної мережі «Сільпо».

Результат аналізу емоційного забарвлення твітів на рис. 3 представлено у вигляді діаграми, де синій колір означає злість, червоний — роздратування, зелений — страх, оранжевий — радість, жовтий — смуток, фіолетовий — здивування.

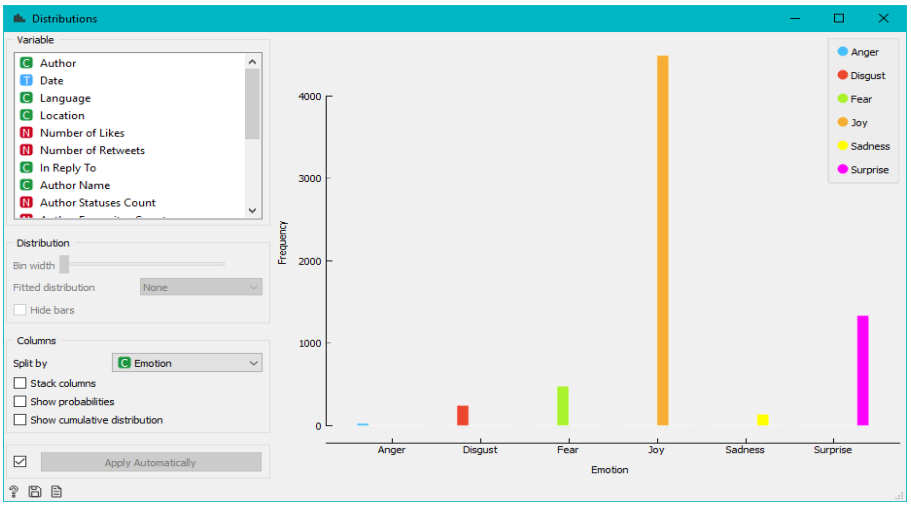

Рис. 3. Результат розподілу твітів користувачів за емоціями

Після візуалізації результатів обробки можна зробити висновок, що проведена маркетингова кампанія припала до душі користувачам соціальної мережі Twitter, що видно з графіка емоцій. У графіку домінує емоція «радість», що свідчить про позитивне ставлення користувачів до реклами торговельної мережі «Сільпо». Негативні ж емоції має невелика частина твітів.

На хмарі зі слів виділяються слова «лучшая», «ахахах», «отлично», «молодежь» тощо, що підтверджує висновок про позитивне враження від реклами (див. графік на рис. 3). Також виділено слово «драконов», що, ймовірно, стосується стилістичного оформлення одного з магазинів у Хмельницькому у китайському стилі. Можна припустити, що з цим також пов'язана значна частина твітів з емоційним забарвленням «здивування».

На основі отриманих результатів можна зробити висновок, що проведена маркетингова кампанія була вдалою і запам'яталась споживачам, їм до вподоби настрій і тип реклами, що була задіяна. Тож при проведенні наступної кампанії слід врахувати цей досвід і продовжити проводити рекламні заходи в тому ж напрямку.

Здійснимо порівняльний аналіз акційних пропозицій торговельної мережі «Сільпо» з конкурентом — торговельною мережею супермаркетів «Фуршет». Порівняльний аналіз проведений на прикладі акційних програм «Власний рахунок» і «Fishka».

На рис. 4 зображено пошук інформації про акційні програми в соціальній мережі Twitter за ключовими словами «власний рахунок» і «fishka». Для порівняння пости були згруповані за тегами «@silpoua» та «@Furshet UA».

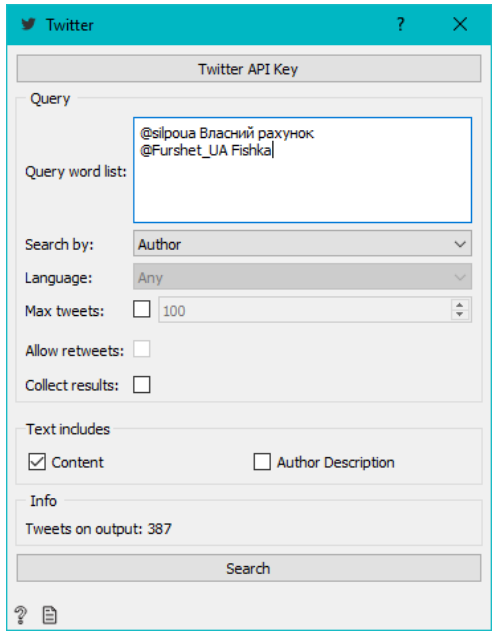

Рис. 4. Пошук інформації в соціальній мережі Twitter

У моделі робочого процесу (рис. 5) використано віджет Box Plot, що дає змогу перевірити нові дані, щоб швидко виявити будь-які аномалії.

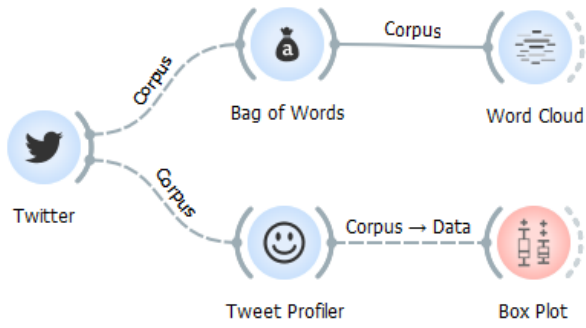

Рис. 5. Модель робочого процесу для порівняння акційних програм

На рис. 6 представлено результат аналізу у вигляді розподілу знайдених постів за емоційним забарвленням.

З розподілу можна зробити висновок, що обидві програми мають найрізноманітніші відгуки, але варто відмітити, що «Фуршет» має 55,31% постів з емоційним забарвлення «радість», тоді як «Сільпо» — лише 38,51%. Також «Сільпо» має значні 21,74% емоцій «страху» та 32,30% — «здивування», що може свідчити про невдоволення клієнтів.

Для аналізу побудовано хмари із слів по кожній акційній програмі для більш детального її вивчення (рис. 7, 8).

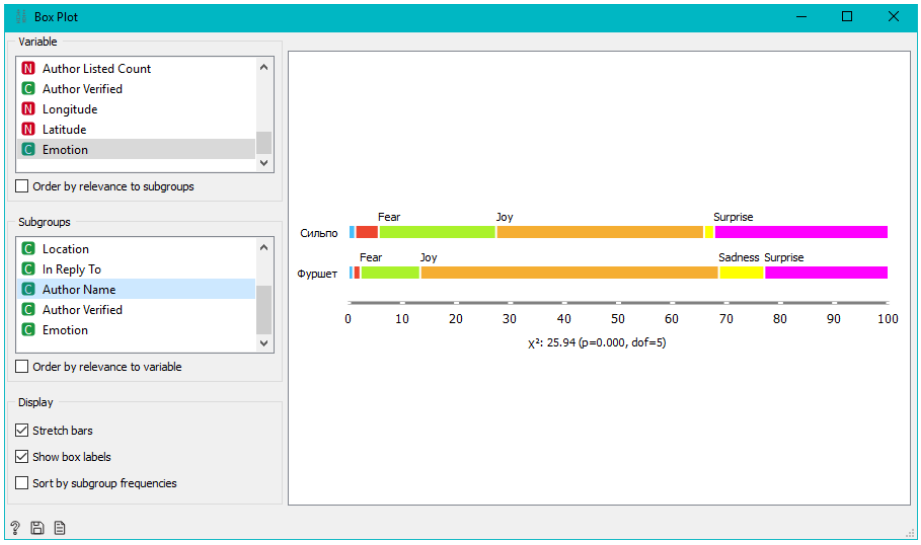

Рис. 6. Розподіл постів за емоційним забарвленням

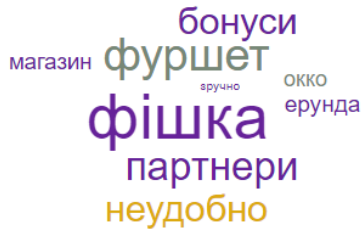

Рис. 7. Хмара із слів за запитом «Фішка»

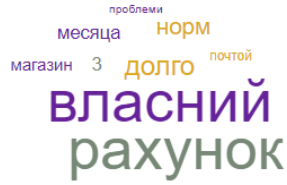

Рис. 8. Хмара із слів за запитом «Власний рахунок»

Хмари із слів дають змогу зрозуміти, що покупців торговельної мережі «Фуршет» (але далеко не всіх) приваблює можливість користування акційною програмою і за межами «Фуршету», адже «Фішка» — це єдина картка для великої партнерської мережі.

У мережі «Сільпо» можна виділити таку проблему: бонуси приходять один раз на три місяці, та ще й в окремих випадках поштою, що незручно для багатьох покупців.

На основі проведених досліджень можна сформувати рекомендації для наступної маркетингової кампанії: переглянути термін видачі бонусів; розробити маркетингові заходи щодо популяризації мобільного додатку «Сільпо», що забезпечить можливість покупцям отримувати бонуси на смартфон.

Отже, використання технології Text Mining дає змогу (без залучення додаткових коштів на проведення досліджень) здійснити аналіз проведених торговельною мережею маркетингових заходів, виявити переваги й недоліки акційної програми мережі порівняно з конкурентами, а також сформувати рекомендації щодо майбутніх маркетингових заходів.

Висновки. За результатами проведених досліджень:

- встановлено ефективність використання технології Text Mining для інформаційної підтримки формування й аналізу маркетингових заходів торговельної мережі;
- адаптовано методи технології Text Mining для практичної реалізації завдань торговельної мережі;
- здійснено оцінку проведеної маркетингової кампанії та порівняльний аналіз акційних програм конкуруючих торговельних мереж методами аналізу тональності тексту та виявлення аномалій технології Text Mining.

ЛІТЕРАТУРА

1. Mushtaq A. Data Mining For Marketing / Aiman Mushtaq, Hina Kanth // International Journal on Recent and Innovation Trends in Computing and Communication. — 2015. — Vol. 3, no. 3. — P. 985—991. DOI: 10.17762/ ijritcc.v3i3. 39 53.

2. Demšar J. Orange: data mining toolbox in Python / Janez Demšar, Tomaž Curk, Aleš Erjavec // Journal of Machine Learning Research. — 2013. — 14(1). — P. 2349—2353.

3. Ромакін В. В. Комп'ютерний аналіз даних: навчальний посібник / Ромакін В. В. — Миколаїв: Вид-во МДГУ ім. Петра Могили, 2006. — 144 с.

4. Нога Р. Аналітичний огляд методів та засобів опрацювання текстової інформації / Р. Нога, Н. Б. Шаховська // Вісн. Нац. Ун-ту «Львів. політехніка». — 2011. — № 715. — С. 323—332.

5. Sumathy L. Article: Text Mining: Concepts, Applications, Tools and Issues An Overview. / K. L. Sumathy, M. Chidambaram // International Journal of Computer Applications. — 2013. — 80(4). — P. 29—32.

6. Галина А. В. Обзор технологии Text Mining / А. В. Галина, Е. А. Есина // Аллея Науки. — 2018. — Т. 2, № 1. — С. 393—396.

7. Луговская Е. М. Методы и этапы автоматической обработки текста / Е. М. Луговская // Вестник Современных Исследований. — 2019. — С. 33—38.

8. Zhao Y. Analysing Twitter Data with Text Mining and Social NetworkAnalysis / Y. Zhao // In Proceedings of the 11th Australasian Data Mining and Analytics Conference (AusDM 2013). — 2013. — 23 p.

9. Оксанич І. Г. Інтелектуальний аналіз масиву текстових документів на основі технології Text Mining / І. Г. Оксанич, Д. М. Піскунов, Д. П. Черниш // Системи обробки інформації. — 2013. — Вип. 2. — С. 139–143.

10. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко. — СПб.: БХВ-Петербург, 2007. — 384 с.

11. Интеллектуальные системы: учебное пособие / Н. Соловьев, А. Семенов, А. Цыганков, Е. Чернопрудова // Оренбургский гос. ун-т. — Оренгбург: ОГУ, 2013. — 236 с.

АНАЛИЗ МАРКЕТИНГОВЫХ МЕРОПРИЯТИЙ ТОРГОВОЙ СЕТИ МЕТОДАМИ TEXT MINING

С. В. Грибков, Р. Р. Ханбабаев, Е. В. Харкянен
Национальный университет пищевых технологий

В статье исследованы и проанализированы методы Data Mining для формирования и оценки маркетинговых мероприятий торговой сети. Особое внимание уделено анализу технологии Text Mining. Это инструментарий, разработанный на основе базовых технологий Data Mining в сочетании со статистическим и линг-

вистическим анализами, методами искусственного интеллекта, который позволяет осуществлять поиск тенденций, шаблонов и взаимосвязей в неструктурированных текстах для принятия управленческих решений.
Приведен пример реализации моделей Text Mining в компонентном программном пакете Orange с использованием Twitter API. Рабочие модели для практической реализации задач торговой сети сформированы на основе набора виджетов Orange. Результат исследования визуализирован в виде диаграмм и облаков со слов, которые удобно просматривать и анализировать.
Ключевые слова: *интеллектуальный анализ данных, Text Mining, социальные сети.*