

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені ТАРАСА ШЕВЧЕНКА
Факультет інформаційних технологій
Кафедра прикладних інформаційних систем**

122 «Комп'ютерні науки»
(шифр і назва спеціальності)

«Прикладне програмування»
(назва освітньої програми)

Кваліфікаційна робота бакалавра

на тему: «Веб-сервіс із аналізу ставлення користувачів Facebook до брендів»



Виконав _____

(Підпис)

Грисюк Дмитро Віталійович

(прізвище, ім'я, по батькові)

Керівник Домрачев Володимир Миколайович

(прізвище, ім'я, по батькові)

(Резолюція «До захисту»)

Попередній захист:

(Висновок: “До захисту в екзаменаційній комісії”)

Завідувач кафедри _____ Плескач В.Л. _____
(Підпис) (Прізвище, ініціали) (Дата)

Київ – 2021

ВІДОМІСТЬ ДИПЛОМНОЇ РОБОТИ

Складові частини дипломної роботи	Обсяг, арк.
Титульний аркуш	1
Завдання до дипломної роботи	1
Відомість дипломної роботи	1
Анотація	1
Анотація (іноземною мовою-англійською)	1
Зміст	1
Перелік скорочень, умовних позначень, термінів	1
Вступ	3
1	21
2	10
3	10
Висновки	1
Перелік використаних джерел	2
Додатки	4

				ДП ХХХХ 00.000.00		
	ПБ	Підп.	Дата			
Розробн.				Відомість дипломної роботи	Лист	Листів
Керівн.						
Н/контр.						
Зав.каф.	Плескач В.Л.					

АНОТАЦІЯ

Дана дипломна робота має 58 сторінок, на яких міститься 9 рисунків, 27 джерел за переліком посилань, 3 додатки. У даній роботі об'єктом дослідження є загрози і небезпеки для бізнесу (брендів) зі сторони користувачів соціальної мережі Facebook. Предметом дослідження є спостереження та аналіз інформаційних потоків соціальної мережі Facebook щодо негативного або невідповідного відношення людей до брендів. Метою даної роботи є збільшити швидкість реагування на ризики для бізнесу та розширити можливості моніторингу і аналізу ставлення користувачів до бренду завдяки створенню відповідного веб-сервісу.

У ході роботи відбувався аналіз текстової інформації за допомогою наступних методів: кластерний аналіз слідів, використовуючи метод агломеративної ієрархічної кластеризації, який шукає наближені кластери в ієрархічному дереві, сентимент аналіз - клас методів контент-аналізу, частотний аналіз - метод частоти появи певних ключових слів, латентно-семантичний аналіз - метод обробки даних. В якості інструменту збору даних з фан сторінок брендів у Facebook використовувався Facebook Graph API.

Оскільки Facebook - це найпопулярніша соціальна мережа з великим числом користувачів і відповідно великим числом брендів, а кількість тих і інших продовжує невпинно рости, то з цього слідує, що розроблений веб-сервіс є і буде потужним інструментом для здійснення моніторингу, аналізу становища свого бізнесу серед користувачів, а також ефективним рішенням негативних слідів.

Ключові слова: веб-сервіс, соціальна мережа Facebook, аналіз даних, бренд, настрої або ставлення користувачів, частотний аналіз, кластеризація.

ANNOTATION

This thesis has 58 pages, which contains 9 figures, 27 sources according to the list of references, 3 appendices. In this work, the object of the study is the threats and dangers to business (brands) from users of the social network Facebook. The subject of the research is the observation and analysis of information flows of the social network Facebook regarding negative or inappropriate attitudes of people towards brands. The purpose of this work is to increase the speed of response to business risks and expand the ability to monitor and analyze user attitudes towards the brand through the creation of an appropriate web service.

In the course of the work, textual information was analyzed using the following methods: cluster analysis of traces using the method of agglomerative hierarchical clustering, which looks for approximate clusters in a hierarchical tree, sentiment analysis - a class of content analysis methods, frequency analysis - the method of frequency of occurrence of certain keywords, latent semantic analysis is a method of data processing. The Facebook Graph API was used as a tool for collecting data from fan pages of brands on Facebook.

Facebook is the most popular social network with a large number of users and, accordingly, a large number of brands, and the number of both continues to grow steadily, it follows that the developed web service is and will be a powerful tool for monitoring and analyzing the position of your business. among users, as well as an effective solution to negative traces.

Keywords: web service, social network Facebook, data analysis, brand, mood or attitudes of users, frequency analysis, clustering.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ	6
ВСТУП	7
РОЗДІЛ 1. ЯК FACEBOOK ВПЛИВАЄ ТА ФОРМУЄ ГРОМАДСЬКУ ДУМКУ І ЧОМУ БРЕНДИ ВИМУШЕНІ МОНІТОРИТИ ТА АНАЛІЗУВАТИ СТАВЛЕННЯ ДО СЕБЕ У FACEBOOK?	10
1.1. Концепція вірусності та популярності інформації у Facebook.	10
1.2. Особливості Facebook та його алгоритм ранжирування публікацій в новинній стрічці.	14
1.3. Вплив і маніпулювання суспільною свідомістю.	17
1.4. Необхідність і способи моніторингу ставлення користувачів Facebook до брендів.	21
Висновки до розділу 1	27
РОЗДІЛ 2. МЕТОДИ АНАЛІЗУ ІНФОРМАЦІЙНИХ ПОТОКІВ	29
2.1. Аналіз інформаційних потоків і контекстів	31
2.2. Виявлення інформаційного потоку	32
2.3. Аналіз контексту	33
2.4. Результати експерименту.	34
Висновки до розділу 2	36
РОЗДІЛ 3. РЕАЛІЗАЦІЯ МЕТОДІВ АНАЛІЗУ ТА МОНІТОРИНГУ СТАВЛЕННЯ КОРИСТУВАЧІВ FACEBOOK ДО БРЕНДІВ.	37
3.1. Безпосередньо сама реалізація	39
3.2. Результати	45
Висновок до розділу 3	47
ВИСНОВКИ	48
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	49
ДОДАТКИ	51

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

SMM - Social Media Marketing - один з інструментів інтернет-маркетингу.

Like, Love, Care - реакції на публікації: подобається, кохання, тримаймося.

ЗМІ - засоби масової інформації.

PR - public relations, піар - зв'язки з громадськістю.

Covid-19 - CoronaVirus Disease 2019 - коронавірусна інфекція 2019 року - потенційно важка гостра респіраторна інфекція.

FFP - Facebook fan page – сторінка фанів бренду у Facebook.

B2B - business-to-business - бізнес до бізнесу - термін, що визначає вид інформаційного та економічної взаємодії.

API - програмний інтерфейс програми, інтерфейс прикладного програмування.

Graph API - це основний спосіб читання і запису даних із Facebook.

NLP - обробка текстів написаних природною мовою.

NLTK - пакет бібліотек і програм для символної та статистичної обробки природної мови.

ВСТУП

Сьогодні більша частина людства не може уявити свого життя без соціальних мереж, адже вони відкривають нових друзів, дозволяють читати цікаві новини, підключатися до обговорень або навпаки ділитися чимось. А з початку пандемії коронавірусу вони стали чи не основною площадкою для спілкування із друзями та рідними. Соціальні мережі стрімко впроваджуються в усі соціальні процеси та істотно змінюють характер комунікативної взаємодії. Аудиторія соціальних мереж величезна, і продовжує рости. Наприклад, у найпопулярнішої соціальної мережі Facebook є 1.84 мільярд щоденних активних користувачів з усього світу станом на 31 грудня 2020 р., а це на 10.8% більше, ніж в попередньому році. Не в останню чергу через пандемію зросла й кількість часу, який люди проводять в інтернеті. Відповідно до дослідження, в середньому типовий користувач інтернету проводить в онлайні майже 7 годин на день, з яких близько 2,5 годин йде на використання соціальних мереж. Якщо скласти усе разом, то сумарно всі онлайн-користувачі витрачають більш ніж 10 мільярдів годин на соціальні мережі за день. Тому упускати такий широкий майданчик зі свого плану не раціонально. Соціальні мережі все частіше використовуються для розвитку та просування брендів. В такому соціальному інтернет просторі формуються позитивні образи і нове сприйняття відомих і невідомих брендів, створюються нові ціннісні орієнтири цільових споживачів. Саме тому уже близько 82% компаній мають свої сторінки як мінімум однієї соціальної мережі. А за даними Social Media Marketing Industry Report 2020 94% маркетологів використовують соціальну мережу Facebook для просування брендів, що являється найкращим показником серед інших соціальних мереж. Однак, окрім позитивних образів бренда часто можуть створюватися і негативні образи, а це в свою чергу суттєво може впливати на кількість продаж, на рейтинги і навіть призвести до занепаду бренда. Зазвичай, бізнес створює спеціальні канали для роботи з клієнтами: гарячу лінію, форми на сайті, чат-

ботів, щоб реагувати на проблеми клієнтів. Але останнім часом склалася тенденція, що незадоволеному покупцеві чи користувачеві зручніше піти в соціальну мережу і там поділитися своєю проблемою із усіма. Можливо, його публікація залишиться непоміченою. Або навпаки - набере багато репостів, коментарів і лайків, викличе народний гнів і бунт. В такому випадку мовчання може дорого обійтися для компанії. Щоб швидко виходити на зв'язок з клієнтами, а також нівелювати негатив, потрібно регулярно моніторити та аналізувати згадки бренду в соціальних мережах. Вручну це робити нереально, саме тому я хочу запропонувати механізм моніторингу та аналізу ставлення користувачів до бренду, який створений на основі обробки інформації відкритих інформаційних потоків найпопулярнішої соціальної мережі Facebook. Цей механізм дозволить бізнесу розуміти реальне ставлення людей до бренду, допоможе виміряти охоплення аудиторії, а головне допоможе швидко і вчасно реагувати на ситуації, які можуть нашкодити компанії.

Дана робота має величезну актуальність, так як на теперішній час для аналізу ставлення користувачів до бренду в якості предмета дослідження досить рідко використовують саме соціальні мережі.

Метою даної **роботи** є збільшити швидкість реагування на ризики для бізнесу та розширити можливості моніторингу і аналізу ставлення користувачів до бренду методом створення відповідного веб-сервісу. Щоб досягти поставлену мету я сформулював і виконав наступне:

- Визначення списку подій чи явищ, що відбулися на території України і світу в конкретні проміжки часу і несли за собою впливові наслідки;
- Формування списку брендів, в тому числі й особистих брендів, які користуються високою популярністю в Україні та за її межами, а також є активними учасниками соціальної мережі Facebook.
- Структурування механізму пошуку необхідної інформації для пошуку
- Аналіз знайденої інформації

Об'єктом дослідження є загрози і небезпеки для бізнесу (брендів) зі сторони користувачів соціальної мережі Facebook.

Предметом дослідження є спостереження та аналіз інформаційних потоків соціальної мережі Facebook щодо негативного або невідповідного відношення людей до брендів.

В цій роботі використовуються наступні методи для аналізу інформації:

- латентно-семантичний аналіз для пошуку схожих змістів;
- частотний аналіз для знаходження частоти появи ключових слів та словосполучень;
- аналіз тональності для визначення емоцій в постах;

РОЗДІЛ 1. ЯК FACEBOOK ВПЛИВАЄ ТА ФОРМУЄ ГРОМАДСЬКУ ДУМКУ І ЧОМУ БРЕНДИ ВИМУШЕНІ МОНІТОРИТИ ТА АНАЛІЗУВАТИ СТАВЛЕННЯ ДО СЕБЕ У FACEBOOK?

Як ми уже довідалися зі вступу, Facebook охоплює неймовірно гігантську аудиторію, а його користувачі проводять чималу частину свого активного дня за переглядом різноманітних фото, відео, читанням новин та постів свої друзів, а також за поширенням та обговоренням тих чи інших матеріалів, які залишили їх небайдужими. З особистого досвіду хочу додати, що Facebook настільки міцно увійшов в буденне життя більшості людей, що зараз неможливо уявити початку дня без оновлення стрічки цієї соціальної мережі. А як можна не відкрити додаток Facebook, коли їдеш в транспорті на роботу чи в університет або ж поки очікуєш у черзі за смачною булочкою? Звісно, що така присутність соціальної мережі в нашому житті не може залишатися безслідно і має вплив на наші думки та рішення.

1.1. Концепція вірусності та популярності інформації у Facebook.

Щоб розібратися детальніше про цей вплив, я проаналізував і дослідив концепцію вірусності і узагальнені фактори, що впливають на популярність інформації, яка розміщена в Facebook. Були виділені дві основні групи факторів вірусності: контент (його тематика і форма), аудиторія. Технологічні особливості розробки і функціонування всесвітньої мережі і комунікаційні технології привели в цілому до значних розбіжностей в процесах і методах залучення і формування аудиторій офлайн і онлайн. Миттєве оновлення інформації в соціальній мережі, збільшення кількості джерел інформації та інші особливості функціонування нових медіа дозволяють говорити про інформаційні надлишкові пропозиції, що в свою чергу, призводить до конкуренції між бізнес брендами чи між особистими брендами в онлайн, в результаті цього відбувається не тільки поліпшення якості та подання

контенту, але і скорочення аудиторії конкретних бізнес-сайтів і персональних сайтів, що в розрізі торгівлі дає можливості для молодих і малих брендів конкурувати із великими. Тому один з інструменти для збільшення відвідуваності чи продажів з Facebook є SMM-просування - вірусність контенту. В додаток, вірусний контент може істотно впливати на громадську думку і маніпулювати суспільною свідомістю в умовах гібридної війни. Використання факторів популярності контенту в соціальних мережах особливо емоцій, які він викликає, стиль подачі, час публікації, предметної області, і беручи до уваги принципи ранжирування звітів по соціальній мережі, бренди можуть забезпечити ефективне джерело аудиторії, популярність в новому медіа-просторі може стати засобом протидії агресії в соціальній мережі.

Багато місцевих та зарубіжні дослідники вивчали медіа в соціальних мережах. Проаналізувавши статті з «Times», американські вчені Дж. Бергер і К. Мілкман виділяють ряд особливостей змісту, які впливають на його вірусність, зокрема емоції читачів, які викликає контент. Оскільки дослідження проводилися на американській аудиторії, тому я вирішив дослідити ці особливості саме для української аудиторії. Зокрема, дослідження потребувало використання методів узагальнення та термінологічного аналізу для уточнення визначення термінів «вірусність» і «зміст вірусності». Аналітичний і синтетичний метод дозволив виділити фактори вірусності контенту і скласти уявлення про всю систему заходів, що застосовуються для просування інформації в соціальних мережах. Метод спостереження і вимірювання використовувалися для збору фактичного вихідного матеріалу для теоретичного вивчення факторів вірусності контенту під українські умови; Метод порівняння дозволив визначити особливості популяризації інформації при особливих обставинах, таких як під час гібридної війни; Метод контент-аналізу для поглибленого вивчення текстових повідомлень соціальних мереж з метою вибору індикаторів, які дозволили інтерпретувати результати

дослідження в конкретні рекомендації по просуванню медіа-контенту в сучасних умовах.

В результаті такого аналізу я сформував ряд визначень. Вірусний контент - це публікації, що володіють здатністю до вірусоподібного самовідтворенню, тобто до природного відтворення матеріалів користувачами Facebook. Аудиторія використовує кнопки «Share» (наприклад, «розкажи своїм друзям», «поділитися») для поширення цікавих публікацій. Дослідники інтернет-комунікації пояснюють феномен вірусності за законом Меткалфа, згідно з яким корисність соціальної мережі пропорційна квадрату кількості його користувачів. Згідно з цим законом, причина вірусності контенту може полягати в збільшенні числа зацікавлених в ньому користувачів, і в результаті цінність контенту збільшується. Однак слід зазначити, що феномен вірусності більш серйозний, складний і універсальний. В цілому можна говорити про три основних складових факторів вірусності:

- зміст - його особливості, форма, предметна область і т. д. Наприклад, стаття, що складається з 10 000 знаків буде менш популярною серед користувачів, ніж яскрава інфографіка на ту ж саму тему;

- аудиторія (користувачі соціальних мереж діляться тільки тією інформацією, яка буде цікавою або корисною для друзів, які бажають отримати « Bonus » взамін на «Like», подяку, захоплення і т. д.);

- особливості соціальної мережі, що впливають на форму і тип контент, а також активність і характеристики аудиторії (кращі годинник розміщення, особливості ранжирування матеріалів в стрічці новин та ін.).

З огляду на перший фактор вірусності - сам контент – один. Слід згадати результати досліджень американських вчених Дж. Бергера. і К. Мілкман, які проаналізували близько 7000 статей, опублікованих в «Times» в пошуках матеріалів, якими найчастіше ділилися користувачі. Приймаючи до уваги такі фактори, як час публікації статті, популярність автора, репутація автора, довжину і складність матеріалу, Дж. Бергер і К. Мілкман розкрили дві характеристики, що визначають вірусність контенту. Вони дізналися, що

популярність матеріалу залежить в першу чергу від того, наскільки позитивна його основна ідея - і наскільки сильні емоції користувача він може викликати. Загалом є кілька особливостей, що впливають на вірусність матеріалів в Facebook:

- публікації, які викликають емоції. Чим більш емоційний матеріал, тим частіше ним діляться в соціальних мережах. Позитивні емоції сприяють популярності більше, ніж негативні;

- життєво важлива інформація. Користувачі соціальних мереж емоційно реагують на інформацію, що стосується їх основних потреб (за даними Maslow's pyramid), події, які суттєво впливають або можуть вплинути на їхнє життя і т. д.;

- корисність інформації. Користувачі діляться матеріалами, які можуть бути корисними, мають цінність і є практичними. Зазвичай ця інформація публікується і поширюється у вигляді списків і нотатків, інструкцій чи коротких відео;

- найсвіжіша інформація. Контент, який спочатку повідомляє ексклюзивну інформацію, зазвичай набуває значної популярності серед користувачів Facebook.

Однак, з огляду на фактор особливостей змісту разом з ще одним фактором - особливості аудиторії, а саме психологічне сприйняття, які спонукають людей ділитися деякою інформацією, найважливішим фактором інформаційної вірусності є її здатність емоційно активувати аудиторію. Тільки ті матеріали, які викликають стан збудження і готовність до дії або спонукання користувачів до емоційної активності мають шанс, щоб стати вірусним обміном в Інтернеті. Це пояснюється психологічним станом людини, викликане різними емоціями: позитивними емоціями, такими як натхнення, велика радість, сміх спонукають ділитися враженнями. Також можуть збільшити активність почуття страху і гніву. Тим не менш, такі емоції, як печаль або задоволення, навпаки, знижують активність користувачів. Крім того, на обмін контентом в соціальних мережах також впливають ще одна

особливість психології користувачів - бажання привернути увагу своїх віртуальних друзів за допомогою «Like», «Love», «Care» і інших методів оцінки публікацій в Facebook, тому користувачі будуть ділитися тільки тією інформацією, яка ймовірно приверне його аудиторію. Популярність контенту також істотно залежить від особливостей, які залежать від самої соціальної мережі: години найбільшої активності користувачів в мережі, алгоритми ранжирування постів в стрічці новин та ін. Веб-аналітика показує, що в Facebook трафік починає рости з 9 ранку до 16 годин. Кращий час для публікації матеріалів з 13:00 до 16:00 по буднях. Як показують дослідження, найбільш ефективними публікаціями виявилися ті, які були розміщені в Facebook в середу в 15:00.

1.2. Особливості Facebook та його алгоритм ранжирування публікацій в новинній стрічці.

В додачу до всього, Facebook має власний алгоритм ранжирування публікацій в новинній стрічці, і це важливо враховувати при створенні вірусності. А ще Facebook часто змінює ці фактори та алгоритми ранжирування. Ось як вони змінювалися:

- У 2009 році Facebook представив першу подобу алгоритму: новинна стрічка почала формуватися в залежності від популярності кожного поста.
- У 2015 році Facebook представив функцію «Бачити першим», яка давала користувачам можливість вибирати сторінки, які повинні з'являтися вгорі новинної стрічки. Також платформа почала знижувати в стрічці сторінки, які постили дуже багато публікацій з рекламним підтекстом.
- З 2016 Facebook почав віддавати пріоритет постам від друзів і сім'ї, а також інформативному і розважального контенту. Також алгоритм почав вимірювати цінність поста на основі кількості часу, який користувачі проводили за його переглядом, навіть якщо вони не

ділилися ним і не лайкали. Прямі відеотрансляції також вийшли в пріоритет, так як вони отримували в три рази більше часу перегляду, ніж звичайні відеоролики.

- У 2017 алгоритм почав оцінювати реакції користувачів на пост (“Angry”, “Sad” і т.д.) вище звичайних лайків. Також до оцінки відео додався показник перегляду відео до кінця.
- У 2018 році Марк Цукерберг оголосив про те, що тепер алгоритм новинної стрічки буде віддавати пріоритет постам, які викликають обговорення і значущі взаємодії. Мета нововведення - вивести на перший план якість, а не кількість часу, який користувачі проводять онлайн, а також взяти більше відповідальності за те, як платформа впливає на своїх користувачів. У той же час у брендів з'явилися обґрунтовані побоювання, що їх органічний контент більше не буде так високо оцінюватися, як контент від сім'ї, друзів і груп. Так як алгоритм пріоритезувати пости, які збирали якісні взаємодії (коментарі, реакції, відповіді на коментарі, розшарювання в месенджері), виходило, що компаніям, щоб отримати охоплення для збільшення залученості користувачів, потрібно було вже мати цю залученість на високому рівні.
- У березні 2019 дослідження від компанії NewsWhip показало, що, в той час як залученість в Facebook зростає на 50% з року в рік, зміни в алгоритмі також збільшують чвари і образи в соцмережі, тому що вони сприяють просуванню постів, які хвилюють людей. У той же час алгоритм почав відстежувати неякісний контент і фейкові новини з ненадійних джерел, які знали, як обдурити систему, щоб отримати якомога більшу увагу користувачів.
- Зараз алгоритм показує в новинній стрічці пости таким чином, щоб кожен користувач бачив їх в тому порядку, в якому вони йому максимально сподобаються і будуть цікаві. І для цього система використовує певні чинники ранжирування. Фактори ранжирування - це

дані про минуле поведінці певного користувача і поведінки в принципі всіх користувачів платформи. Наприклад, як діляться люди цим постом з друзями? Як часто користувачі лайкають пости? Як часто він дивиться прямі трансляції? Яка у нього улюблена група? Скільки постів технічно для нього зараз є? Наскільки вони свіжі? Тобто алгоритм зважує безліч чинників перед тим, як вирішити, чи показувати вам те саме відео з маленькою пумою. Facebook згадує три великі категорії факторів ранжирування:

- З ким зазвичай взаємодіє користувач,
- Тип медіа в пості (відео, посилання, фото і т.д.),
- Популярність поста.

У березні 2019 Facebook анонсував інструмент, мета якого - дати користувачам більше прозорості і контролю над тим, що показується в новинній стрічці. Кнопка «Чому я бачу цей пост?» робить саме те, що обіцяє: допомагає користувачеві зрозуміти, чому алгоритм показує саме цей пост в новинній стрічці. А також він дозволяє людям «сказати» алгоритму безпосередньо, що їм важливо, а що дратує. Тобто користувачі можуть повідомити Facebook, що вони хочуть бачити менше постів від певної людини або бачити більше публікацій певної бізнес-сторінки.

У травні 2019 Facebook проводив призначені для користувача опитування, щоб дізнатися безпосередньо у своїй аудиторії, який контент їм важливий. В опитуваннях було наступне:

- Хто їх близькі друзі?
- Які публікації (посилання, фото, відео) вони вважають цінними?
- Наскільки важлива певна група, до якої вони належать?
- Наскільки їх цікавить контент певних бізнес-сторінок, на які вони підписані?

В ході обробки відповідей соціальна мережа виділяла певні патерни, які потім додавалися в алгоритм. Наприклад, сторінки і групи, які користувачі

позначали, як найбільш важливі, часто були тими, на кого вони давно підписані, часто взаємодіють, і тими, у кого було багато постів і активностей. Алгоритм Facebook можливо завжди буде в процесі вдосконалення. Тому брендам потрібно завжди бути напоготові щодо новинок від Facebook і постійно вдосконалювати методи для збільшення органічного охоплення.

1.3. Вплив і маніпулювання суспільною свідомістю.

Розглядаючи питання створення вірусного контенту, варто згадати не тільки інструменти, але і цілі поширення такої вірусної інформації. За допомогою ведення правильної вірусної кампанії у Facebook бренди можуть домогтися не тільки неймовірної популярності, а й кращих показників SMM, конверсії або продажі. Вірусність контенту дає можливість маніпулювати аудиторією. Неймовірна кількість інформації, що з'являється в стрічці новин в Facebook, ситість інформаційного простору змушує користувачів зосередитися на виборі людей, цікавитися змістом, а не шукати кращі пости в мережі на своїй власній стрічці. Як зазначає Дж. Бергер, отже, списки найбільш популярного контенту може генерувати теми для широкого обговорення. Наприклад, чим більше людей бачили і поділилися статтею про фінансову реформу, тим більша ймовірність, що інші отримали впевненість у тому, що фінансова реформа заслуговує на більшу увагу з боку держави, ніж реформа охорони навколишнього середовища, навіть якщо ядро фінансової реформа не так важлива, як екологічна.

Нав'язування аудиторії певних тем для обговорення або відволікання уваги від важливих, але проблемних питань за допомогою факторів вірусного контенту можна розглядати як метод маніпулювання суспільною свідомістю, що дозволяє досягти бажаної відповіді на інформацію, наприклад, з точки зору гібридної війни. Для визначення факторів вірусності контенту та сили впливу на суспільну думку в соціальних мережах я вирішив проаналізувати сторінки Facebook топ українських ЗМІ, які є найбільш популярними в цій мережі і які можна назвати брендами.

Бренд 24 каналу продовжує займати лідерські позиції, зокрема щодо роботи зі соцмережею Facebook. Протягом лютого 2021 року з медіа 24 каналу взаємодіяли 2,66 мільйона користувачів. Цей показник є найбільший серед усіх новинних ЗМІ України. У топ-5 також потрапили:

- ТСН – 2,45 мільйона взаємодій;
- "Радіо Свобода" – 921,66;
- "Обозреватель" – 846,90;
- "Громадське" – 619,24.

Щодо кількості підписників у фейсбуці, то у 24 каналу їх наразі 987,59 тисячі.

Тут видання посідає 3 сходинку після ТСН та "РБК-Україна" (рис 1.1)

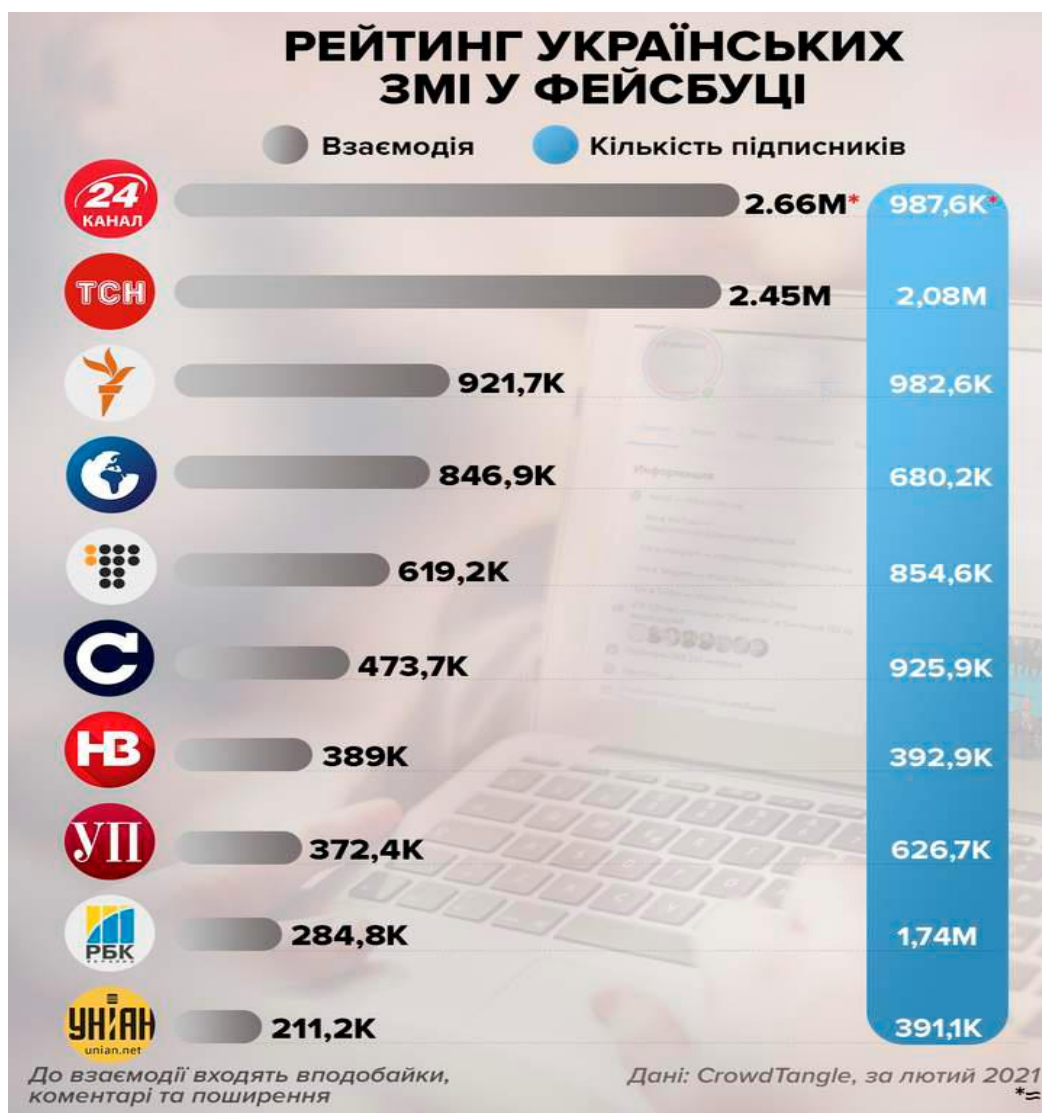


Рисунок 1.1 – Рейтинг українських ЗМІ у Фейсбуці

Результати аналізу розглядалися з точки зору факторів, таких як тип змісту, дата і час публікації і тема області повідомлень. Тип контенту, яким діляться в соціальних мережах вважається одним з найважливіших чинників популярності поста. Сторінки аналізованих ЗМІ найчастіше містили опубліковані посилання на сайт з коротким повідомленням про зміст статті. Таким чином, наприклад, всі матеріали на сторінках «Української правди» були посиланнями на сайт з повним текстом; сторінку «ТСН» крім публікацій з посилання, кілька разів викладалися тільки фотографії; «Громадське ТБ» містить різний контент: фото, відео, посилання на сайт, gif-анімацію і т. д.; на сторінці 24 каналу також міститься різноманітний тип контенту, який у всіх випадках містить посилання на повну статтю на сайті, а в коментарях під кожною публікацією міститься посилання ще й на телеграм 24 каналу.

І як виявилось, одна й та ж інформація на каналах з майже однаковою кількістю аудиторії може отримувати різне число реакцій від користувачів, якщо ця інформація була опублікована в різний час. Наприклад, публікація про вихід української групи Go_A у фінал Євробачення 2021 опублікована на 24 каналі в 8:43, отримала 10635 лайків, 425 коментарів і 309 репостів. Та сама новина опублікована на сторінці «Радіо Свобода» в 12:07, отримала 2142 лайки, 91 коментар і 296 репостів. В результаті оперативнішого публікування інформації 24 канал отримав в 4,5 разів більше реакцій чим «Радіо Свобода». Також на кількість результатів впливає і тип контенту.

Наприклад, якщо опублікувати сам ролик із подією і ще написати опис до нього на своїй сторінці в Facebook, а не просто посилання на сайт з цією новиною, то вдасться отримати в десятки разів більше взаємодій. Мобільність користувачів на соціальній мережі і специфіка розважальних сервісів (спілкування, перегляд фото, відео, прослуховування музики)

викликають велику популярність контенту, який можна відразу переглянути в стрічці новин. У цьому причина великої популярності інших подібних постів на різні теми (фотографії В. Кличко на велосипеді, фотографії з Євробачення, відео Києва з висоти пташиного польоту, фото тюльпанів

виставка в Києві, відео, присвячене Дню вишиванки і т. д. - всі вони отримали від 2000 до 10 000 лайків). Крім формату матеріалу, визначальним фактором популярності контенту в соціальних мережах є тематика інформації.

Відповідно до аналізу, найбільш популярні пости в ЗМІ

сторінки в Facebook наступні:

- 24,7% - про політику і публічне управління;
- 16,8% - війна на сході України;
- 16,2% - медицина, здоров'я, Covid-19;
- 12,2% - культура, освіта, матеріальне благополуччя, поліпшення міста і ін.;
- 11,3% - зі світової політики;
- 11,3% - пізнавальна інформація, спорт, космос і ін.;
- 7,5% - трагічні події в Україні та світі.

Згідно з міжнародними дослідженнями, найбільш вірусним є контент, який викликає позитивні емоції. Однак в Україні найвищий відсоток публікацій присвячений війні і політиці, а також поточним умовам життя, такі пости не викликають почуття радості і задоволення аудиторія. Переважання негативних емоцій в ЗМІ соціальних мереж можна пояснити трансформацією сенсу матеріалів по заданих темах в сфері життєвих потреб (по піраміді Маслоу). В умовах гібридної війни більшість людей бачить навіть політику через призму виживання, безпеки, можливості забезпечити себе їжею, теплом і іншими терміновими ресурсами. Серед найпопулярніших постів - ті, які коментують найактивніші користувачі і вони теж в більшості присвячені політиці. А оскільки згідно з алгоритмом ранжирування інформації в стрічці новин Facebook, коментар для публікації цінніший, ніж лайк, то інформація таких тематики стає найпопулярнішою. В результаті впливу Facebook на українських користувачів, скалася ситуація, що в українському контексті з урахуванням політична ситуація і гібридна війна, найпопулярніший контент присвячений актуальним темам - політиці, війні,

основним життєвим потребам; у вигляді відео, фото або короткий текст без посилань на зовнішні ресурси для отримання повної інформації; відправлено днем або пізніше, спонукає людей до обговорення і коментування (включаючи питання або заклики в тексті, щоб мати трохи суперечливий сенс), відповідають тематична область популярних тем для українців, а для обміну матеріалом - дати можливість користувачу отримати реакцію друзів на свій «репост», наприклад, завдяки корисності матеріалу, емоціям радості, гордості, екзальтації і важливість інформації для широкої аудиторії.

1.4. Необхідність і способи моніторингу ставлення користувачів Facebook до брендів.

Соціальний моніторинг - невід'ємна частина будь-якої стратегії соціального обслуговування клієнтів. Це може привести до більш швидкого реагування, запобігти PR-криз і виявити так званих захисників бренду (публічних або впливових фанатів бренду). Було досліджено, що реагування на запити клієнтів веде до підвищення лояльності клієнтів до бренду.

Що ж собою являє моніторинг Facebook? Моніторинг Facebook має на увазі відстеження згадок вашого бренду в соціальній мережі і реагування на будь-яку взаємодію з ним. Це стратегія негайного реагування, що відноситься до всіх розмов про ваш бренд: оглядам, питань про продукти і навіть скарг на ремонтні роботи. Що ще більш важливо, моніторинг Facebook - це не тільки прямі згадки або коментарі до ваших публікацій. Він поширюється на публікації або повідомлення, в яких згадується ваш бренд, але вас не відзначають. Такі повідомлення не з'являються в звичайних мережевих повідомленнях, тому важливо налаштувати інструменти соціального моніторингу. До моніторинга Facebook також варто віднести і соціальне прослуховування. Коли ви моніторите, ви відповідаєте відповідно. Але коли ви практикуєте прослуховування соціальної мережі, ви в цілому аналізуєте згадки бренду і галузеві тенденції, щоб краще інформувати про майбутні маркетингові рішення. І те й інше може бути надзвичайно цінним для вашого

бренду в індивідуальному порядку. Моніторинг Facebook може принести бренду багато користі. Найбезпосереднішою перевагою моніторингу соціальних мереж є те, що він змушує ваших клієнтів відчувати, що їх чують і підтримують. Соціальний індекс Sprout: Empower & Elevate показав (рис 1.2), що:

- 59% споживачів звертаються до брендів в соціальну мережу, тому що у них є хороший досвід;
- 47% споживачів звертаються до брендів в Facebook, тому що у них є питання про продукти або послуги;
- 40% споживачів звертаються до брендів в соціальну мережу, тому що у них був поганий досвід;



Рисунок 1.2 – Графік «Чому користувачі звертаються до брендів в соціальній мережі?»

Можу припустити, що ви коли-небудь писали або ж залишали коментар якійсь компанії на її сторінці у Facebook і не отримували жодної відповіді. Це, мабуть, засмучувало і вплинуло на те, як ви зараз сприймаєте компанію, незалежно від того, усвідомлювали ви це в той час чи ні. 78% опитаних Sprout маркетологів відповіли на запити клієнтів протягом 12 годин. Коли ви відповідаєте споживачам в соціальних мережах, вони на 21% з більшою ймовірністю здійснять покупку саме у вас. Моніторинг допомагає зрозуміти сприйняття вашого бренду і настрою.

Моніторинг Facebook також допомагає розібратися в шумі і пропускати тільки важливі повідомлення. Замість того, щоб прокручувати свою тимчасову стрічку в Facebook в надії знайти згадку про ваш бренд, або повторювати трудомісткий ручний пошук, який може пропустити варіанти термінів бренду, краще використати стратегію моніторингу, яка буде попереджати бренд всякий раз, коли про нього говорять. У ситуаціях, коли ключові впливові особи скаржаться в Інтернеті на бізнес, більш швидка відповідь може запобігти PR кошмарам. Це також допомагає брендам стежити за захисниками бренду, з якими ви хочете взаємодіяти. Якщо бренд не знав про цих інфлюенсерів раніше, то може знати зараз завдяки моніторингу і можуть пам'ятати про це в майбутніх маркетингових кампаніях.

Однак моніторинг Facebook не обмежується тільки своїм власним брендом. Можна також використовувати моніторинг як частина аналізу конкурентів: подивитися, що люди говорять про конкурентів і які їхні основні компліменти і скарги. Коли ваш бренд про це попросить, ви можете почати розмову і запропонувати свої продукти в якості альтернативи тому, що використовують конкуренти.

Моніторинг Facebook включає в себе відстеження згадок бренду і подальшу реакцію на них. Моніторинг тримає бренд в курсі повідомлень, перш ніж вони можуть стати вірусними по неправильним причин. Вони також допомагають підтримувати позитивний імідж бренду серед споживачів і виявляти потенційних впливових осіб.

Моніторинг облікових записів в соціальній мережі дає власникам брендів уявлення про те, що обговорюють їхні клієнти, тим самим допомагаючи задовольнити їхні потреби і підвищувати рівень задоволеності. Facebook змінив спосіб взаємодії компаній зі своїми клієнтами. Незалежно від того, що користувачі говорять про продукти і послуги компанії, вони можуть звертатися до соціальних мереж, щоб висловити свої думки і думки своїх друзів. Що, якщо бренда немає поруч, щоб почути їх коментар і відповідати на нього?

Хоча теоретично немає обмежень на кількість і типи речей, які можете відстежувати в ширшій бесіді в соціальних мережах, але ось деякі з найбільш важливих показників, які потрібно відстежувати:

- Брендіві терміни і згадки бренду: найпростіший спосіб залишатися в курсі розмов про бізнес в Facebook - це відслідковувати всі згадки бренду, які компанія отримує в соціальній мережі.
- Настрій клієнтів: хороший інструмент для моніторингу Facebook дасть уявлення не тільки про те, як часто згадується бренд, а й про те, що про нього говорять люди.
- Огляди: інструменти моніторингу Facebook можуть допомогти відстежувати, що люди говорять про бізнес на основних сайтах оглядів соціальних мереж. Бренд повинен відповідати на будь-які відгуки, які заслуговують на увагу, особливо на негативні.
- Хештеги: потрібно відстежувати будь-які хештеги, що відносяться до бренду або галузі, щоб зрозуміти, як цільова аудиторія сприймає компанію.
- Конкуренція: так само, як люди говорять про бренд в Facebook, змагання також є частиною розмови. Потрібно використовувати інструменти управління соціальною мережею, щоб дізнатися, як конкуренти досягають більшого успіху, а потім діяти у відповідності зі своїми відкриттями.

Моніторинг соціальної мережі став більш корисним, ніж будь-коли, оскільки компанії прагнуть зрозуміти швидко мінливі настрої споживачів після епідемії Covid-19. Щоб отримати всі переваги моніторингу соціальних мереж, організації повинні розділити відділи і обмінюватися інформацією між декількома відділами. Малі підприємства, некомерційні організації та державні установи можуть отримати від прослуховування соціальних мереж ті ж переваги, що і великі корпорації.

Багато керівників продажів і маркетингу пророкують, що продажі у Facebook стануть домінуючим напрямком продажів B2B. І більшість говорить, що прослуховування соціальних мереж - це важливий перший крок в процесі пошуку потенційних клієнтів і клієнтів в соціальній мережі.

Споживачі часто публікують фотографії, відео і хвалять продукти в Facebook. Соціальний моніторинг дозволяє PR і маркетингу знаходити цей контент. Контент, створений користувачами, може бути надзвичайно ефективною - але доступною - маркетинговою стратегією. Дослідження показують, що споживачі довіряють рекомендаціям інших споживачів, особливо знайомих і впливових осіб, більше, ніж рекламі бренду. Пошук, розширення і повторне використання користувацького контенту доступні за ціною і ефективні. Це дозволяє невеликим організаціям конкурувати з компаніями з великими кишнями, щоб платити за дорогу рекламу.

Повідомлення в Facebook тепер впливають на репутацію бренду і покупки так само сильно, як і основні новини. Досвідчені фахівці зі зв'язків з громадськістю розуміють, що моніторинг Facebook, а також блогів і дощок оголошень тепер так само важливий, як і моніторинг традиційних новинних агентств. За даними Pew Research Center, все більше населення Землі отримують новини через Facebook, ніж будь-коли, і ця цифра продовжує зростати, а соціальна мережа вже перевершила друковані газети в якості кращого джерела новин для американців. Споживання новин у Facebook більше не обмежується молоддю або інтернет-кмітлівістю. Тому важливо

відстежувати згадки в соціальній мережі бренду, продуктів, керівників компанії і інших ключових слів компанії.

Моніторинг Facebook - життєво важливий інструмент кризової комунікації до, під час і після PR-кризи. У розпал кризи моніторинг соціальної мережі може надати команді PR-криз інформацію, близьку до реального часу, яка часто буває більш актуальною, ніж новинні агентства. У міру розвитку кризи моніторинг може виявити думки ключових груп споживачів і важливих впливових осіб, таких як журналісти, блогери та інвестори. Крім повідомлення про те, що впливові особи говорять про брендах, і оцінки настрою їх коментарів, прослуховування в соціальних мережах може виявляти і ранжувати впливових осіб, щоб допомогти PR пріоритезувати свої відповіді. Відстежуючи, як кількість негативних, позитивних і нейтральних згадок змінюється з плином часу, фахівці зі зв'язків з громадськістю можуть дізнатися, як їхня стратегія впливає на громадське обговорення з плином часу. Виявлення і аналіз сплеску негативних згадок або іншої незвичайної активності в Facebook також може виявити зародження PR-кризи і дати організації час діяти до того, як проблема приверне увагу основних засобів масової інформації.

Відстеження настрою співробітників. Багато співробітників можуть страждати від соціальної ізоляції і турбуватися про збереження роботи, своє здоров'я і безпеки повернення в офіс через кризу Covid-19. Експерти з антикризового управління рекомендують бізнес-лідерам поєднувати традиційні методи з прослуховуванням інформації в Facebook, щоб оцінити психологічне благополуччя співробітників і підготуватися до швидкому коригуванню стратегій. «Моніторинг та отримання інформації про реакцію робочої сили на урядові заходи в режимі реального часу життєво важливі для підтримки бізнесу на плаву і ефективного планування на невизначене майбутнє», - заявляють експерти ETN Zurich (Швейцарський федеральний технологічний інститут у Цюріху). Також використання аналітики соціальної мережі, дає можливість відстежувати і розуміти колективні настрої робочої

сили, оскільки, на їхню думку, почуття можуть швидко поширюватися через Facebook.

Думки, які люди висловлюють в Facebook, дають реальне уявлення про те, як вони ставляться до організації та її продуктів. Зазвичай це перший ключ до вирішення проблем. Щоб зрозуміти це, важливо дійсно читати думки споживачів, а не просто об'єднувати їх в дані.

Висновки до розділу 1

В українському контексті з урахуванням політичної ситуації та гібридної війни, найпопулярніший контент в Facebook присвячений актуальним темам - політиці, війні, коронавірусу, основним життєвим потребам; у вигляді відео, фото або коротких текстів без або з посиланнями на зовнішні ресурси для отримання повної інформації; створені днем або пізніше. Найчастіше публіці подобається контент, що викликає позитивні емоції, однак, згідно з принципами ранжирування інформації в стрічці новин Facebook, пости отримують найбільше охоплення аудиторії, коли ними діляться і коментують користувачі. Таким чином, для підвищення інформаційної вірусності чи популярності в Facebook публікація повинна спонукати людей до обговорення і коментування (включаючи питання або недоговореність в тексті, щоб мати трохи суперечливий сенс), відповідати тематичним областям популярних тем для українців, а для обміну матеріалом - дати можливість користувачу отримати реакцію друзів на свій репост, наприклад, завдяки корисності матеріалу, емоціям радості, гордості, шоку, несправедливості чи загрози і важливістю інформації для широкої аудиторії.

Аналіз соціальних мереж для дослідження ринку може доповнити або, можливо, навіть замінити традиційні дослідження ринку. Традиційні дослідження ринку - опитування, фокус-групи та глибинні інтерв'ю - відображають почуття невеликої групи людей в певний момент часу. Зібрана інформація обмежена за часом, вибіркова і спотворена. Аналітика та

моніторинг Facebook може виробляти дані за години або дні, на відміну від тижнів або місяців для традиційних методів дослідження. Крім того, оскільки думки виражаються без підказки, вони з більшою ймовірністю розкривають справжні почуття клієнтів.

Прослуховування чи моніторинг Facebook дає організаціям переваги, що виходять за рамки маркетингу і PR. Компанії можуть знаходити нових клієнтів, покращувати продукти, досліджувати свої ринки, збільшувати продажі, оцінювати добробут співробітників і покращувати свої PR-кампанії.

РОЗДІЛ 2. МЕТОДИ АНАЛІЗУ ІНФОРМАЦІЙНИХ ПОТОКІВ

Останнім часом сервіси соціальних мереж, такі як Facebook, все частіше стають онлайн-платформою, на якій окремі особи і організації розміщують і обмінюються інформацією, такий як новини, думки і рекламні оголошення, швидким і простим способом. Оригінальні повідомлення повідомляються передплатникам сайтів, а коментарі до повідомлення дозволяють знову доставляти інформацію друзям. Оскільки коментарі до повідомлення відображають, як інформація насправді поширюється від людей до людей, виявлення прихованих інформаційних потоків, в основі яких лежать контексти на основі коментування поведінки людей, є важливою проблемою аналізу поширення інформації в соціальних мережах. Складне розуміння механізмів поширення інформації, якими насправді користуються люди, може бути досягнуто за допомогою інформаційних потоків та контекстів, які прискорюють роботу багатьох соціальних методів, таких як вірусний маркетинг, соціальне маніпулювання і аналіз впливу.

Зокрема, це дослідження фокусується на фан-сторінках (FFP - Facebook Fun Pages) Facebook, які покликані об'єднувати людей з однаковими інтересами. FFP включає в себе безліч повідомлень, кожен з яких складається з текстового опису та коментарів, як показано на рисунку інформаційного потоку та контексту допису на сторінці фанатів у Facebook (рис. 2.1)

Опис повідомлення представляє його конкретний контекст, в той час як його коментарі містять інформаційний потік контексту, заснований на людях. взаємодії в хронологічному порядку. Тут послідовність упорядкованих коментарів до конкретної публікації називається слідом коментарів до публікації.

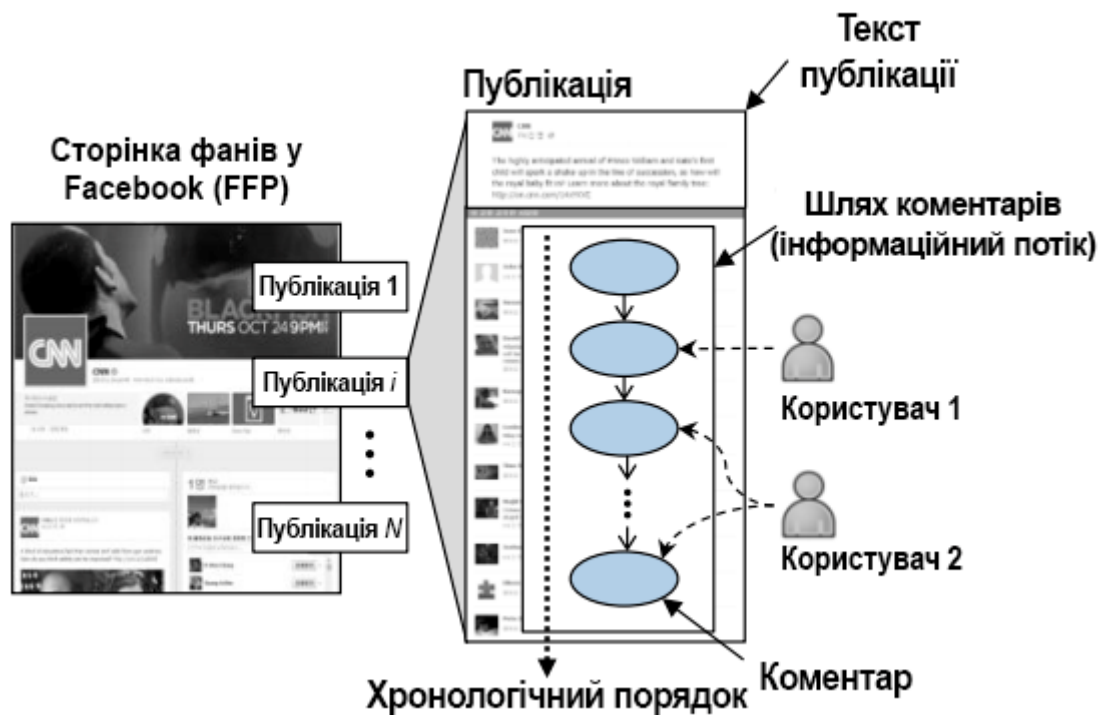


Рисунок 2.1 - Інформаційний потік та контекст допису на сторінці фанатів Facebook.

Існує безліч досліджень інформаційних потоків, заснованих на соціальних взаємодіях, таких як розрахунок впливів, аналіз структури соціальної мережі Facebook і виявлення процесів розповсюдження інформації. Однак їх головний недолік - відсутність можливості запропонувати відповідний контекст з інформаційним потоком. Більш того, попередні методи не можуть вловити динамічну природу інформаційних потоків, оскільки вони в значній мірі покладаються на статичні зв'язки між людьми, так звані дружбою, а не на динамічне поширення інформації, встановлені людьми для контексту. Далі хочу представити інтегрований метод аналізу контексту на основі коментарів до FFP, а також виявлення значущих інформаційних потоків від FFP. У той час як попередній метод мав справу тільки з виявленням інформаційного потоку на основі поведінки людей при повторній публікації, інтегрований метод, спочатку пропонує застосування кластеризації трасування для пошуку кількох значущих інформаційних потоків, а потім

розкриває основний контекст кожного потоку на основі часто використовуваних ключових слів. Таким чином, використовуючи коментарі людей до повідомлень, запропонований метод може виявляти інформаційні потоки, які одночасно мають на увазі фактичні відносини між людьми, а також їх контекст. Зокрема, методи інтелектуального аналізу даних і текстового аналізу використовуються для розкриття мети цього метода.

Результати експерименту, засновані на коментарях, зібраних в FFP, демонструють, що цей метод може запропонувати ще одну точку зору для розуміння поширення інформації серед людей шляхом виявлення прихованих інформаційних потоків і їх контекстів.

2.1. Аналіз інформаційних потоків і контекстів

Запропонований метод складається з п'яти кроків, як показано на рис. 2.2, де:

- Comments gathering by using Facebook Graph API - збір коментарів за допомогою Facebook Graph API;
- Agglomerative trace clustering - агломеративна кластеризація трасування;
- Information flow discovery for a trace cluster - виявлення потоку інформації для кластера трасування;
- Data acquisition - збір даних;
- Trace clustering - кластеризація трасування;
- Information flow discovery - виявлення потоку інформації;
- Context analysis - аналіз контексту;
- Information flow visualization - візуалізація потоку інформації;
- Description based keyword suggestion - Пропозиція ключових слів на основі опису;

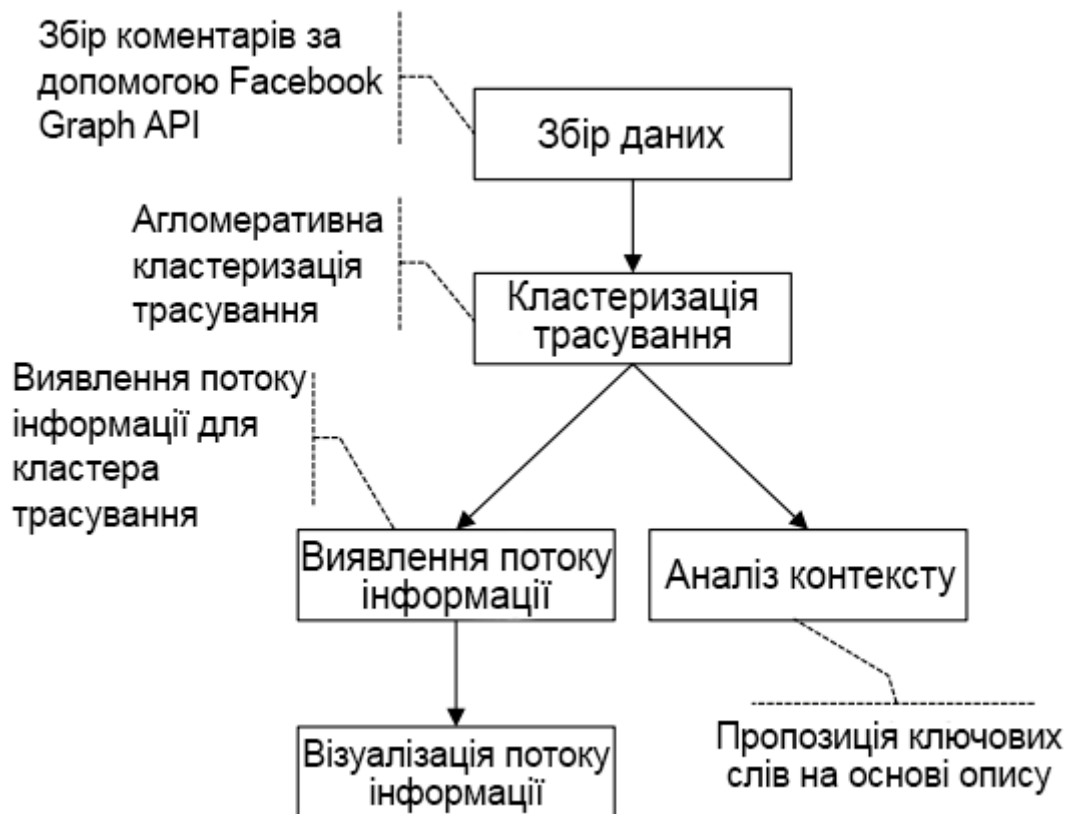


Рисунок 2.2 - Блок-схема запропонованого способу

На основі трасування FFP, отриманих на етапі збору даних, аналогічні траси групуються в кластер трас на етапі кластеризації траси. Потім для кожного кластера трасування його інформаційний потік виявляється шляхом виявлення послідовностей, які часто з'являються в коментарях людей, пов'язаних з інформаційним потоком, і виявлений інформаційний потік візуалізується в графічному вигляді. Тим часом, контекст кожного інформаційного потоку досліджується за допомогою ключових слів на основі описів публікацій, пов'язаних з інформаційним потоком.

2.2. Виявлення інформаційного потоку

Перш ніж виявляти інформаційні потоки, схожі сліди необхідно згрупувати, оскільки слід окремого повідомлення виявляє тільки конкретний випадок поширення інформації, а не радикальні потоки, неодноразово

спостерігаються між повідомленнями. Відповідно, проводиться кластерний аналіз слідів, використовуючи метод агломеративної ієрархічної кластеризації, який шукає наближені кластери в ієрархічному дереві. Траса представлена як вектор довічних значень, що означає послідовність появи людей на трасі, а схожість між трасами вимірюється за допомогою коефіцієнта Жаккара. На початковому етапі кожна траса налаштовується як сам кластер трасування, і два кластери трасування послідовно об'єднуються в один кластер траси, якщо вони є найбільш схожою парою, до тих пір, поки у них не залишиться жодного кластера траси для об'єднання.

Згодом виявляється інформаційний потік кластера трасування, який успішно пояснює моделі взаємодії між людьми через повідомлення в кластері трасування. Оскільки всі відносини між людьми дуже складні, інформаційні потоки бажано абстрагувати, зосередивши увагу на деяких важливих потоках. Отже, ми застосовуємо нечіткий аналіз, який виявляє значущі послідовності дій, використовуючи спрощення адаптивного графа, особливо для менш структурованих потоків. Значимість між людьми досліджується на основі того, наскільки дві людини послідовно корелюють один з одним. Згідно значущості людей деякі групуються, якщо вони менш значущі в порівнянні з іншими.

2.3. Аналіз контексту

Щоб витягти ключові слова відповідно до інформаційними потоками, текстові описи повідомлень, пов'язаних з інформаційним потоком, об'єднуються в один текстовий документ. Потім важливість слова в текстовому документі досліджується з точки зору частоти терміна (TF) і зворотної частоти документа (IDF). Зокрема, в якості показника важливості слова використовується схема зважування TF-IDF, яка широко застосовується в багатьох додатках інтелектуального аналізу тексту, таких як пошук інформації та узагальнення тексту.

У цьому дослідженні слово вважається ключовим словом інформаційного потоку, якщо воно часто зустрічається в описах, пов'язаних з інформаційним потоком, і рідко з'являється в описах, пов'язаних з іншими інформаційними потоками. Після обчислення значень важливості всіх слів для кожного інформаційного потоку, слова з найвищим рейтингом розглядаються як ключові слова інформаційного потоку.

2.4. Результати експерименту.

Для експериментів ми використовували FFP CNN (<https://www.facebook.com/cnn>), який публікує новинні статті, і більше 0,2 мільйона осіб діляться думками, коментуючи повідомлення в FFP. Пости, опубліковані в період з 1 січня 2021 року до 15 травня 2021 року, були зібрані з використанням розробленого веб-сервісу. Щоб зменшити шум, ми видалили коментарі користувачів, які з'являлися менше восьми разів, і в цілому В результаті в експерименті було залучено 297 повідомлень і 22 297 коментарів. Для кластеризації трасування і виявлення інформаційних потоків використовувався інструмент аналізу процесів ProM.

На рисунку 2.3 показана дендрограма кластеризації трасування для зібраних повідомлень і коментарів, де:

- Clustering threshold - поріг кластеризації;
- Traces - сліди, траси;

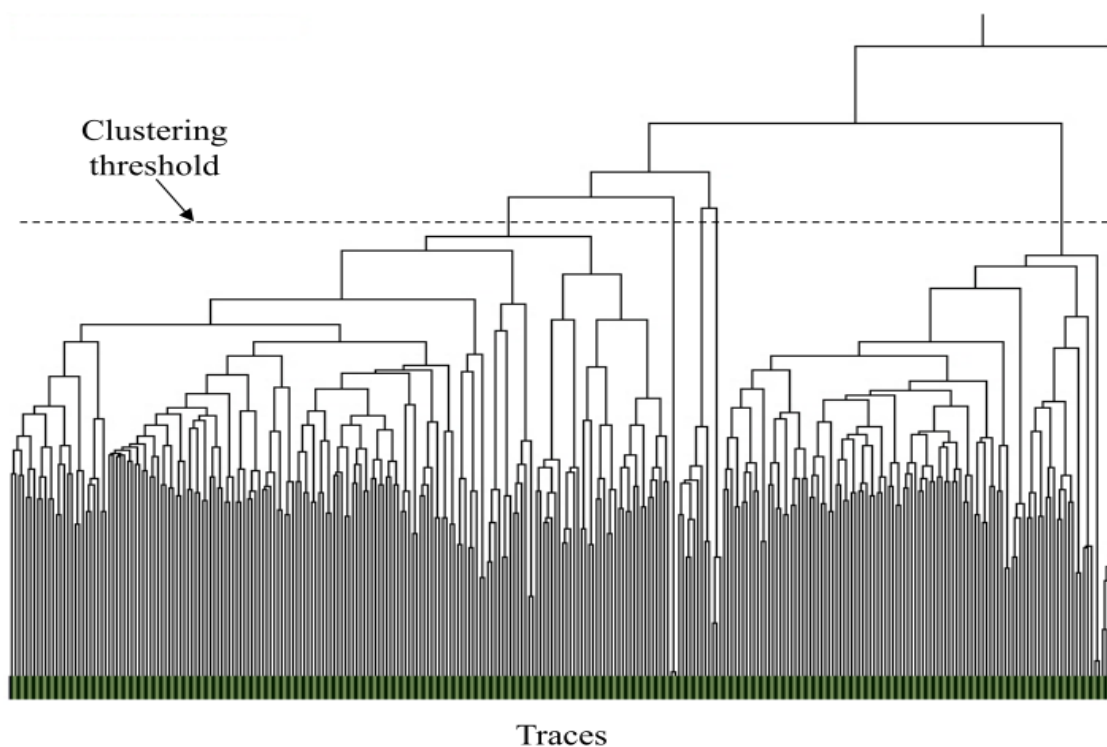


Рисунок 2.3 - Дендрограмма трас, витягнутих при кластеризації трас.

Графік візуалізує схожі повідомлення з точки зору інформаційних потоків їх коментарів, а лінії на графіку представляють результати етапів ітеративної кластеризації. Поріг кластеризації, який визначає кількість кластерів трасування, був обраний емпірично. Після видалення кластерів трас з менш ніж трьома трасами залишилося чотири кластери трас. Кластери А, В, С і D містили 179, 10, 104 і 4 траси відповідно.

Для кожного кластера трасування було виявлено інформаційний потік, який показує, як інформація поширювалася серед людей в FFP. З точки зору інформаційного потоку можна знайти три типи ролей, таких як:

- джерело інформації, яка ініціює інформаційні потоки;
- інформаційний посередник, який пов'язує між людьми або групами;
- приймач інформації, який в основному споживає інформацію і майже не вводить наступні реакції інших.

Нарешті, ключові слова кожного кластера, які представляють його основний контекст, були проаналізовані на основі агрегованих описів повідомлень, пов'язаних з його інформаційним потоком. Було видалено стоп-слова, щоб усунути шуми. 5 основних ключових слів для чотирьох кластерів описані в таблиці 1, в якій їх значення TF-IDF представлені в дужках. Крім того, виявлений контекст може бути візуалізованим в хмарі слів на основі значень TF-IDF відповідних ключових слів.

Наприклад, ключові слова кластера C в таблиці вказують на те, що його інформаційний потік стосувався в основному жертв бомбардування Бостона. Очікується, що інформація, пов'язана з бомбардуванням Бостона, буде поширюватися серед людей. По-перше, було виявлено, що такі люди, як «Ендрю Картер» і «Елізабет Сагг», в основному ініціювали потоки інформації про бомбардування Бостона. По-друге, такі люди, як «Ланора Макнейл» і «Моллі Діас», використовували пов'язану інформацію з іншими, зв'язуючись між групами людей. Нарешті, було відмічено, що людина, «Гленда Дін Пол», часто завершував інформаційний потік про вибух в Бостоні, вказуючи на те, що люди рідко реагували на її коментарі з приводу цієї події.

Висновки до розділу 2

У цьому дослідженні був проаналізований метод, який створений не лише для виявлення інформаційних потоків на основі кластерів трасування за допомогою об'єднання подібних трасів, але та на їх аналіз базових контекстів. Результати експерименту показані, як запропонований метод може бути застосований до розуміння поширення інформації постів на FFP. Я вважаю, що цей метод дозволяє веб-аналітику виявити, хто зацікавлений у повідомленнях на сайті, і як інформація з цього повідомлення розповсюджується в Facebook на основі коментарів до FFP.

РОЗДІЛ 3. РЕАЛІЗАЦІЯ МЕТОДІВ АНАЛІЗУ ТА МОНІТОРИНГУ СТАВЛЕННЯ КОРИСТУВАЧІВ FACEBOOK ДО БРЕНДІВ.

Основна проблема, яку я хотів вирішити, - це створення техніки, яка допомагає зрозуміти і передбачити ставлення користувачів до брендів, виражене в коментарях Facebook на сторінках цих брендів, тобто на FFP.

У веб-сервісі я відтворив наступний метод для моніторингу та аналізу даних:

- використовуючи інформацію зібрану з даних Facebook, буде проводитися пакетний аналіз, щоб зробити точну модель прогнозування для виявлення і прогнозування негативних і позитивних моделей поведінки користувачів;
- використовуючи обробку тексту в реальному часі, буде виявлятися закономірність і розуміємо, як саме виглядає поточна ситуація і поточний настрій;
- зрозуміти, як ця ситуація буде моделюватися на основі передбачення патернів;
- виконати дії зі зміни настрою, якщо це необхідно.

Перед тим як приступити до опису самої реалізації веб-сервіса із аналізу ставлення користувачів Facebook до брендів необхідно ознайомитися із інструментами та засобами, які були використані для побудови цього веб-сервіса.

API Graph - це основний інструмент для завантаження даних на платформу Facebook і їх отримання звідти. Він являє собою API на базі HTTP, за допомогою якого додатки можуть програмним шляхом запитувати дані, публікувати нові історії, управляти рекламою, завантажувати фото і виконувати безліч інших завдань. Назва API Graph підкреслює зв'язок цього API з "соціальним графом" - системою подання інформації на Facebook. API Graph складається з наступних елементів:

- вузли - окремі об'єкти (наприклад, Користувач, Фото, Сторінка або Коментар);
- кордону контексту - зв'язку між підбіркою об'єктів і окремим об'єктом, наприклад фото на Сторінці або коментарі до фото;
- поля - дані про об'єкт, наприклад день народження користувача або назву Сторінки.

За допомогою вузлів можна отримувати дані про конкретний об'єкт, межі контексту дозволяють отримувати добірки об'єктів, пов'язані з окремим об'єктом, а поля - отримувати дані про окремий об'єкт або про кожен об'єкт в добірці. API Graph заснований на протоколі HTTP і працює з усіма мовами, в яких є бібліотека HTTP (наприклад, cURL, urllib). Іншими словами, API Graph можна використовувати прямо в браузері.

Webhook в веб-розробці - метод розширення або зміни поведінки веб-сторінки або веб-додатки за допомогою зворотних викликів. Ці зворотні виклики можуть підтримуватися, змінюватися і управлятися сторонніми користувачами та розробниками, які не обов'язково пов'язані з вихідним веб-сайтом або додатком. Основний формат - JSON. Запит виконується як HTTP POST запит. Webhook також використовується для того, щоб різні системи могли обмінюватися один з одним інформацією. Ось тільки «спілкування» тут відбувається за іншим принципом. Цей механізм спеціально створювався для того, щоб спростити процедуру повідомлення про різні події - зміни налаштувань, додаванні нового користувача, видаленні повідомлення і т.п.

Natural language processing (NLP) – це область, яка зосереджена на тому, щоб зробити природну мову людини придатною для використання за допомогою комп'ютерних програм.

NLTK, or Natural Language Toolkit - набір інструментів для природної мови - це пакет Python, який можна використовувати для NLP. Багато даних, які б можна було аналізувати, є неструктурованими даними і містять текст, який пишеться або читається людиною. Перш ніж вдасться програмно

проаналізувати ці дані, спочатку необхідно їх обробити. **NLTK** використовують для:

- знаходження текст для аналізу;
- попередня обробка тексту для аналізу;
- аналіз тексту;
- створюйте візуалізації на основі аналізу;

JupyterLab - це інтерактивне веб-середовище розробки для блокнотів, коду і даних Jupyter. JupyterLab відрізняється гнучкістю: налаштовуйте і упорядковуйте призначений для користувача інтерфейс для підтримки широкого спектра робочих процесів в області науки про дані, наукових обчислень і машинного навчання. JupyterLab є розширюваним і модульним: пишуть плагіни, які додають нові компоненти і інтегруються з існуючими.

The Jupyter Notebook - це веб-додаток з відкритим вихідним кодом, яке дозволяє створювати і обмінюватися документами, що містять живий код, рівняння, візуалізації і розповідний текст. Області застосування: очищення і перетворення даних, чисельне моделювання, статистичне моделювання, візуалізація даних, машинне навчання і багато іншого.

Python - високорівнева мова програмування загального призначення з динамічної строгою типізацією і автоматичним управлінням пам'яттю, орієнтований на підвищення продуктивності розробника, читання коду і його якості.

3.1. Безпосередньо сама реалізація

Перший підхід - аналіз коментарів в реальному часі. Це рішення дозволяє аналізувати і оновлювати результати вчасно, коли дані генеруються користувачами Facebook. Другий підхід аналізує збережені дані в пакетному режимі. Це кластерний підхід для виявлення і прогнозування моделей негативного / позитивного ставлення. Для оцінки ефективності прогнозування Використовується медіанна міра абсолютного відхилення. Аналіз тональності

поток в реальному часі і генерація подій. Рішення для потокової обробки в реальному часі постійно отримує дані з сервера Facebook, а потім обробляє пакет даних за невеликий проміжок часу. обробка в реальному часі. Бібліотека NLTK використовується для аналізу тональності тексту.

Результати обробки даних перевіряються заздалегідь заданими користувачем умовами. Якщо він задовольняє умовам, рішення створює подію для поновлення статусу інформаційної панелі. Крім того, рішення в реальному часі включає в себе процедуру включення в події. Якщо відбулася певна подія, панель буде оновлена. На рисунку 3.1 представлена запропонована схема обробки поток в реальному часі, де:

- Get data – отримані дані;
- Storage data – дані для зберігання;
- Check for generation event - перевірка генерації;
- Process data – обробка даних;
- Event for new data – події для нових даних;
- Event observation – спостереження за подіями;
- Dashboard – панель відображення результатів;

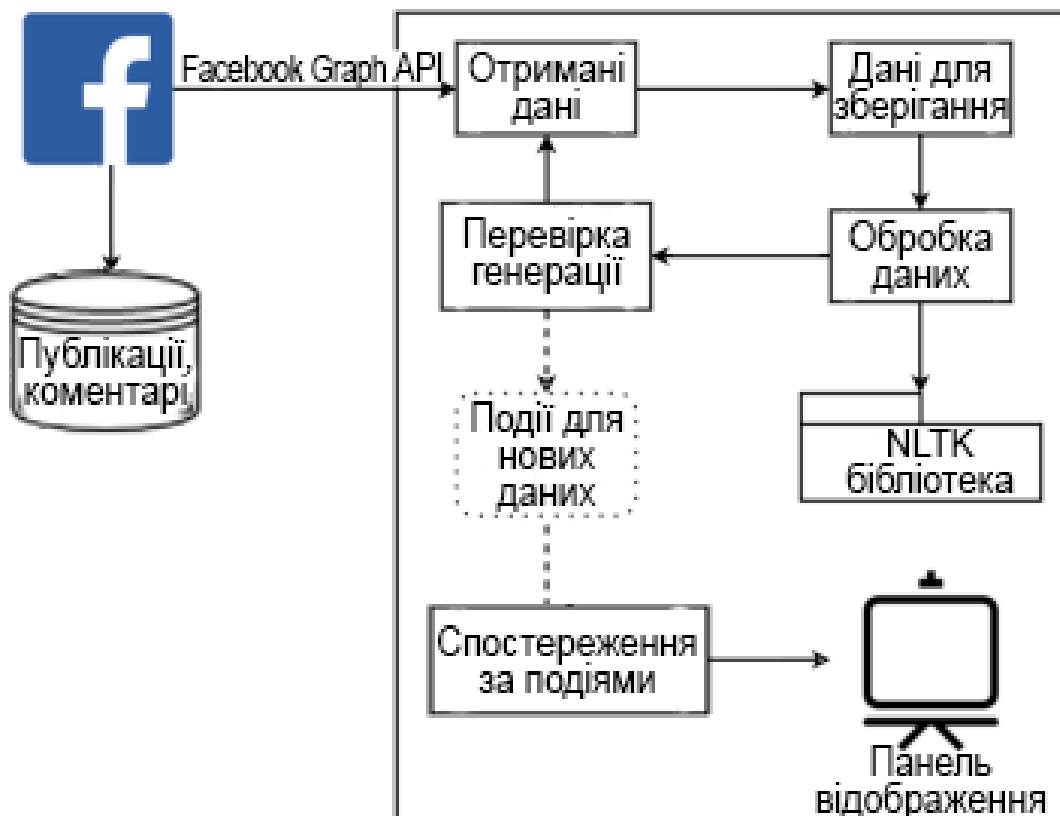


Рисунок 3.1 - Схема обробки потоку в режимі реального часу

Отримання даних. Для поновлення набору коментарів використовується оператор циклу. Використовується для отримання даних коментарів від Facebook час (T_u) між двома петлями. Виписки визначається заздалегідь. Було встановлюємо $T_u = 0,1$ с, так як це майже в реальному часі і дозволяє зберігати постійно оновлювані дані. Однак цей параметр може при необхідності можна змінити.

Graph API Facebook використовується на кожному етапі циклу для отримати всі коментарі до вибраного посту і передати дані для зберігання і подальший аналіз. Такий підхід гарантує, що зібрані дані актуальні. Генерація подій. У внутрішньому циклі, коли дані передаються для подальшого опрацювання, порівняння нового сегмента даних, який був отриманий з Facebook недавно з кеш даними, зроблено. Це дуже важливо, тому що допомагає нам дізнатися зміни в даних. Якщо виявляються зміни в даних, подія вимагає бути згенерували. Спостерігачі отримають цю подію в

подальшому. Спостереження за подією. Для оцінки ефективності методу програма реалізації включає процедуру, що дозволяє спостерігати будь-які згенеровані події. У разі подій, що відображають зміну даних, рішення оновлюється, система автоматизує повторний аналіз даних настроїв. Це пов'язано зі зміною статусу приладової панелі. Крім того, оновлюються дані кеша, щоб переконатися, що у нас є новітні дані з сервера.

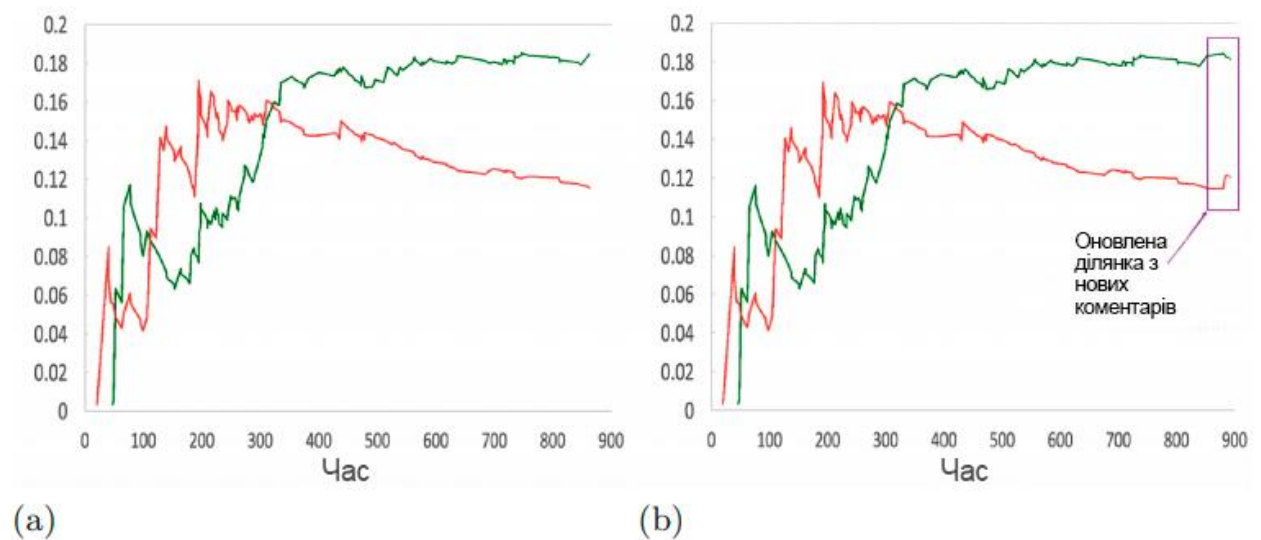


Рисунок 3.2 - негативне і позитивне ставлення в різні позначки часу

На рисунку 3.2 показана панель, що відображає негативне і позитивне ставлення в різні позначки часу. Червона лінія представляє значення негативних настроїв. Зелена лінія - позитивний. Ці значення потрапляють в інтервал $[0; 1]$. Кожен в момент t , у нас є сума від'ємного, позитивного і нейтрального, рівна 1:

$$V_t^{(p)} + V_t^{(n)} + V_t^{(u)} = 1$$

де, $V(p)$ - позначає позитивну оцінку, $V(n)$ - негативну оцінку, $V(u)$ - позначає нейтральну оцінку. У цій статті ми не ставимо нейтральні оцінки.

в графіку. Шкала «час» представляє час публікації коментаря у вигляді інтервал від створення поста в секундах. Шкала «аналізу настроїв» - це

полярність коментарів. Перший графік пояснює поведінку через кожні 780 секунд життя поста, а другий - реакцію на опублікований синтетичний негативний коментар. Аналіз в реальному часі дозволяє виявляти поточні патерни і порівнювати отриманий патерн з очікуваним або потрібно. Однак пошук і коригування цих шаблонів посилянь залежить від при втручанні фахівця (людини) і через високу швидкість і різноманітності даних, ця процедура дуже дорога.

Пакетна обробка даних. Оскільки метод включає моделі виявлення і прогнозування шаблонів, інший компонент нашої системи реалізує виявлення шаблонів в пакетному режимі за допомогою обробки великих обсягів даних з певних тем, зібраних за тривалий період часу. Коментарі у вигляді екстового вмісту по темі збираються з декількох популярні сторінки. Кількість постів вибирається довільно (100 постів 1000 постів, або навіть більше). Наше рішення використовує NLTK, а результати використовуються для оцінки настроїв. Виявлення і передбачення закономірностей. Збір даних. Реалізація пакетної обробки даних має сенс у випадку великих обсягів даних. По-перше, ми вибрали популярну останнім часом тему. Для кожного поста з використанням Facebook Graph API були зібрані всі коментарі протягом перших 30000 с. Дані зберігаються в форматі плоскою таблиці (наприклад, файл CSV), який легко зберегти в розподіленій файлової системи. Заголовок файлу CSV містить такі стовпці: [Дата і час] [Тема] [Повідомлення] [Коментар] [Позитивне] [Негативне]. Перший стовпець містить значення в секундах, коли з'явився коментар, який відраховується з моменту появи вихідного повідомлення. Тип даних значення – це ціле число. Другий стовпець містить назву теми. Третій стовпець - це пост назву. Четвертий стовпець - це зміст тексту коментаря. Останні два стовпці представляють оцінки аналізу настроїв (негативні і позитивні) коментаря на пошті. Тип даних оцінок - це число з плаваючою комою. Дані аналізу. Як уже згадувалося вище, аналіз настроїв NLTK був подав заявку на кожен коментар до публікації. В результаті отримуємо тимчасовий ряд який описує негативний $V(n)$ і позитивний $V(p)$

бал за цей час. На рисунку 3.3 показані позитивні тимчасові ряди оцінок (зліва) для набору постів і негативний (права сторона).

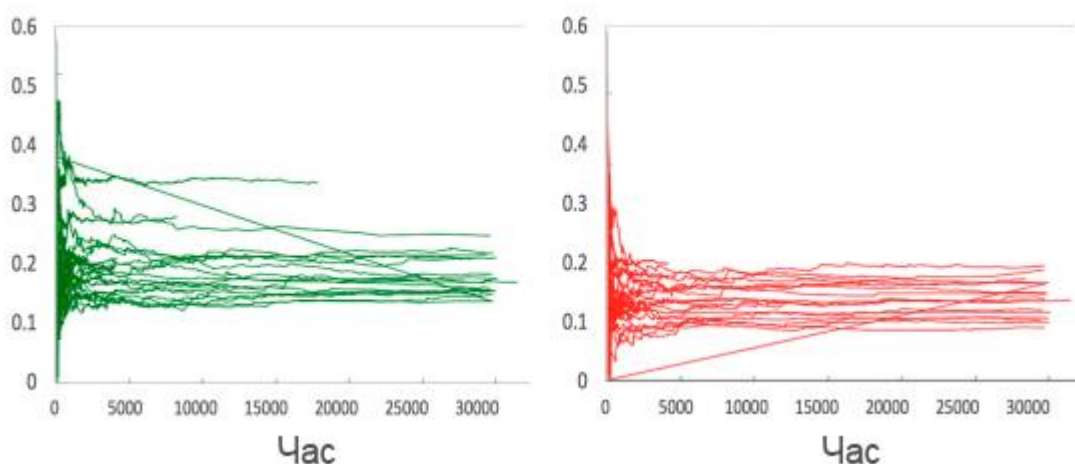


Рисунок 3.2 – позитивні і негативні ряди оцінок

Кластеризація даних. Результатом попереднього кроку є набір часових рядів і кількість часових рядів дорівнює кількості перевірених постів. Ми використовуємо той факт, що деякі тимчасові ряди мають схожий зразок, і вони можуть бути організовані в 3 або 4 різні групи. Це дозволяє визначити «типове» поведінку люди з точки зору настроїв. Наприклад, певний пост міг мати сильний негативний вираз спочатку і поступово згасати негативним. Добре відомим методом неконтрольованої угруповання є кластеризація. У цій статті для кластеризації використовувалися два підходи машинного навчання: k-середніх і МВ-середніх. настроїв постів. Зайве говорити, що k-means вивчалася і застосовувалася в широкий спектр доменів, наприклад аналіз транспортної мережі, інформація безпеки, розпізнавання образів, класифікація тексту і багато інших доменів.

Реалізована кластеризація, при якій вихідні тимчасові ряди розміщуються в трьох різницевих групах відповідно до їх характеристиками. Для виконання кластеризації кожен повідомлень, представлених у вигляді вектора з 20 компонентами (функціями): [id] [Публікація] [Період 1] [Період 2] ... [Період 20]. Ці особливості описують наші перші 30 000 постів. Значення 30 000 з були обрані довільно і при необхідності можуть бути змінені. Кожна

функція має час того ж періоду, що становить близько 1500 с. Цінність кожної функції - середня вартість всієї вартості в цей період. Всі тимчасові ряди (або вектори) були розділені на три кластери з використанням алгоритму k-середніх. Кількість зібраних кластерів за попереднім аналізом. Для порівняння з методом k-середніх ми також використовуємо інший алгоритм для кластеризації, який називається MB-means. Кластеризація дозволяє визначати типові закономірності в поведінці людей. Наступне завдання - прогнозування тенденції розвитку відносини людей до посту з використанням даних про першу реакції (або 5 значень у векторному оданні).

Прогнозування виконується в два етапи. Перший крок, включаючи кластер процедура виявлення. На основі перших 5 ознак (або 5 значень у векторному поданні) даного шаблону, який необхідно передбачити, ми вибираємо відповідний кластер. Обраний кластер - це кластер, який має середній профіль для цих 5 особливості з мінімальним відхиленням від нашого поста (з точки зору евклідової відстані).

На другому етапі ми прогнозуємо тенденцію розвитку відносини людей до цього поста. Найближчий (по віддаленості) тимчасовий ряд від певного кластера дорівнює знайденому і використовуються в якості прогнозу для тимчасового інтервалу відпочинку. середній Абсолют Помилка (MAE) використовується для оцінки ефективності методу прогнозування.

3.2. Результати

Для оцінки підходу було розроблено і впроваджено веб-сервіс з використанням Python, Facebook Graph API, HTML, CSS, JavaScript Для експериментів було використано тему «Президентські вибори в Штатах 2020». Всі дані були отримані від двох відомих нових канали: новини BBC та CNN на Facebook. Було зібрано 200 постів про згадану вище тему. Для кожного повідомлення з використанням Facebook Graph API всі коментарі були зібрані протягом перших 30 000 с. Приблизно загальна кількість коментарів складає близько 100 000 коментарів. Результати аналізу настроїв в реальному часі

показують емоції користувача і думки в потоці в реальному часі. На основі моніторингу панелі відображення ми можемо зрозуміти, що співвідношення позитивних настроїв вище, ніж негативних. Це означає, що частка людей, що підтримують новини, становить більше, ніж частка хто проти. У той час, коли перетин червона лінія і зелена лінія спостерігаються, це вказує на те, що тенденція відносини змінено. Можливо, справа в тому, що, можливо, люди змінюють ставлення до теми. Грунтуючись на цих моментах, ми можемо створити лінію, яка є спільною тенденція відносини людей. Також можна налаштувати тригери, які вказують, коли відбувається зацікавлена подія.

Використовуючи методи кластеризації, ми можемо виявити найбільш типова поведінка користувачів і описує їх. Наприклад, ми спостерігаємо, що негативні або позитивні оцінки асимптотично наближаються до певного рівня і ніколи не перевищують поріг. Негативний і позитивне ставлення з часом зникає, і ми вміють оцінити час популярності поста і поради дії, щоб підтримати популярність. Також наша методика дозволяє прогнозувати розвиток тренда ставлення людей. Це може бути фреймворк для виявлення викидів в коментарі спільноти Facebook. Для оцінки ефективності прогнозування було проведено MAE - застосовується як вимір похибки. Щоб уникнути випадкового отримання результатів, ми розробили перехресну перевірку (зі складками = 10) і отримали середнє значення MAE = 0,008.

На рисунку 3.4 представлені результати прогнозування тенденцій. Зелена лінія це реальне думка про коментарях до цього посту. І червона лінія - це передбачення лінія розвитку відносини людей до пошти.

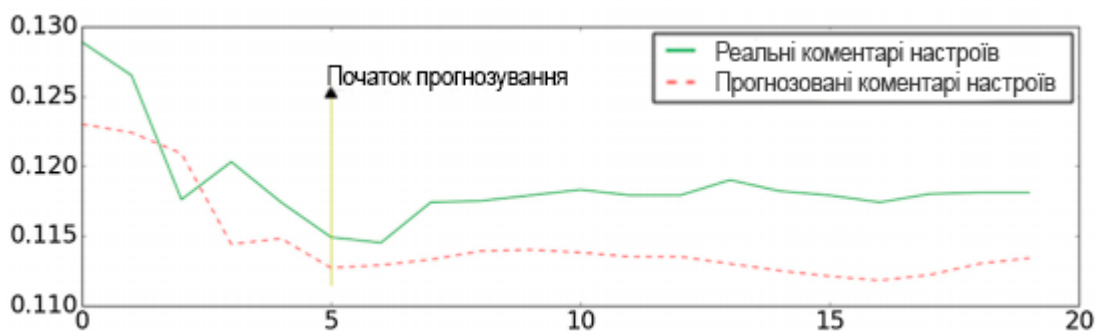


Рисунок 3.4 - Прогнозування тенденції розвитку позитивних настроїв

Висновок до розділу 3

У цьому дослідженні були виконані дії, щоб зрозуміти переваги і ставлення користувачів на основі аналізу настроїв коментарів Facebook і додатки машини методи навчання. Виявлення законів і послідовних закономірностей в коментарях користувачів, опублікованих у встановлені терміни, дозволяє брендам послуг і продажу реагувати в режимі реального часу і бути активнішими, використовуючи прогнозування тенденцій. За результатами було виявлено три типи поведінки користувачів та їхню думку.

ВИСНОВКИ

В результаті виконання даної роботи було розроблено веб-сервіс, який дає можливість власникам брендів аналізувати ставлення користувачів Facebook до свого бренду і до бренду конкурентів, а це в свою чергу дозволило досягнути поставленої мети цієї роботи, а саме – збільшити швидкість реагування на ризики для бізнесу, ефективно справлятися із виявленими проблемами, розширити можливості моніторингу і аналізу ставлення користувачів до бренду.

Отриманий веб-сервіс доцільно використовувати для автономного онлайн моніторингу Facebook, щоб моделювати та аналізувати розвиток подій, завчасно реагувати на виникнення можливих ризиків та правильно прогнозувати подальший розвиток ситуації, а головне зменшити вплив негативних відбитків на власний бренд і налагодити кращі зв'язки зі своїми клієнтами.

Для створення даного веб-сервісу були використані такі інструменти для розробки, як HTML, CSS, JavaScript для побудови візуального відображення розробленого веб-сервісу, Facebook Graph API для отримання даних із соціальної мережі Facebook, NLP та NLTK бібліотеки для обробки та аналізу текстових даних природної мови, мова програмування Python для побудови необхідних функцій, JupyterLab та Jupyter notebook для попереднього відтворення і запуску програмного коду.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. 20 найцінніших статистичних даних Facebook, грудень 2020 р.
<https://zephoria.com/top-15-valuable-facebook-statistics/>
2. Метод TF-IDF на Python – Режим доступу: <http://nlpx.net/archives/57>
3. Facebook Graph API – довідник для розробників – Режим доступу:
<https://developers.facebook.com/docs/graph-api/>
4. Natural Language Processing with Python | Bird Steven, Klein Ewan
5. Основы Natural Language Processing для текста – Режим доступу:
<https://habr.com/ru/company/Voximplant/blog/446738/>
6. Modified K-means Combined with Artificial Bee Colony Algorithm and Differential Evolution | Mohammad Babrdel Bonab
7. BeautifulSoup – парсинг HTML в Python на примерах – Режим доступу:
<https://python-scripts.com/beautifulsoup-html-parsing>
8. Facebook Business Help Center – Режим доступу:
https://www.facebook.com/business/help?ref=mobile_logo
9. 500 Social Media Marketing Tips | Andrew Macarthy
10. Social Media Marketing Workbook | Jason McDonald
11. Facebook Marketing Step-by-Step | Bryan Bren
12. Facebook Marketing: World Class Strategies For Optimizing Your Page, Getting Lots of Likes and Creating Compelling Facebook Ads That Produce Powerful ... Strategies for Business Advertising | Susan Hollister
13. Facebook Application Development For Dummies | Jesse Stay
14. Facebook Application Development | Nick Gerakines
15. The key word in "Data Science", Simply Statistics / J. Leek.
16. Метод определения эмоций в текстах | Пазельская А. Г., Соловьев А. Н.
17. Інтернет-технології як спосіб просування сучасних брендів | Скубченко А.О.
18. Дослідження. Для чого насправді Facebook українським ЗМІ, хто використовує його ефективніше за інших та чи постраждають медіа від

- змін Цукерберга. Автор: Євгенія Кузнєцова. Ресурс доступу:
<https://voxukraine.org/longreads/fb-smi/index.html>
19. Бренд как социокультурный феномен | Костылева Н. В
 20. Оманлива реклама: Світова практика контролю | Головкина Н.В.
 21. Использование интернет-технологий в создании и развитии образа бренда – Режим доступу: <https://cyberleninka.ru/article/n/ispolzovanie-internet-tehnologiy-v-sozdanii-i-razvitiibraza-brenda>
 22. Как узнать о лояльности клиента к бренду– Ресурс доступу:
https://gb.ru/posts/brand_assessment
 23. 35 + Facebook статистика и факты для 2021 – Ресурс доступу:
<https://www.websitehostingrating.com/ru/facebook-statistics/>
 24. 2020 social media marketing industry report. How Marketers Are Using Social Media to Grow Their Businesses | Michael A. Stelzner – Ресурс доступу:
<https://alltomalmhult.se/wp-content/uploads/2020/05/IndustryReport-2020.pdf>
 25. The new Digital 2020 October Global Statshot Report – Ресурс доступу:
<https://wearesocial.com/blog/2020/10/social-media-users-pass-the-4-billion-mark-as-global-adoption-soars>
 26. 10 лучших сервисов для мониторинга социальных сетей – Ресурс доступу:
<https://www.unisender.com/ru/blog/idei/monitoring-sotsialnyh-setey/>
 27. Why brands need a social media monitoring strategy – Ресурс доступу:
<https://sproutsocial.com/insights/social-media-monitoring/>

ДОДАТКИ

Додаток А

Код алгоритму входу в користувацький акаунт Facebook:

```
def login(fb_login, fb_password):
    #access website
    try:
        driver.get('https://www.facebook.com/')

        #Accessing Login frame
        form=driver.find_element_by_id('login_form')
        form.click()

        #Entering email details
        email = form.find_element_by_id('email')
        email.send_keys(fb_login)
        #time.sleep(1)

        #Entering password details
        pwd = form.find_element_by_id('pass')
        pwd.send_keys(fb_password)
        time.sleep(1)

        #Clicking the login button
        button=WebDriverWait(driver,
1000).until(EC.element_to_be_clickable((By.ID, 'loginbutton')))
        button.click()

    except Exception as e:
        print('Exception encountered during Login')
        print(e)
```

Додаток Б

Код алгоритму зчитування відгуків користувачів з FFP:

```
def scrapData(url,html):
    try:
        driver.get(url)
        time.sleep(3)

        # Scrolling to the bottom
        lenOfPage = driver.execute_script("window.scrollTo(0,
document.body.scrollHeight);var
lenOfPage=document.body.scrollHeight;return lenOfPage;")
        noOfPageScrolls=0
        while(noOfPageScrolls < 50):
            print("noOfPageScrolls: ",noOfPageScrolls)
            time.sleep(3)
            lenOfPage = driver.execute_script("window.scrollTo(0,
document.body.scrollHeight);var
lenOfPage=document.body.scrollHeight;return lenOfPage;")
            print(lenOfPage)
            noOfPageScrolls=noOfPageScrolls+1

        source_data = driver.page_source

        # Passing page source into BeautifulSoup to start parsing
        bs_data = bs(source_data)
        except Exception as e:
            print('Помилка отримання джерела сторінки')
            print(e)
        try:
            with codecs.open(html,'w',encoding='utf8') as fw:
                fw.write(str(bs_data))
            fw.close()
        except Exception as e:
            print('Помилка записання отриманих даних до файлу')
            print(e)
```

Додаток В

Код алгоритму аналізу настроїв (позитивні, негативні) користувачів:

```
def sentiment_analysis(text, positive_words, negative_words):

    sentiment=None

    posWordCount = 0
    negWordCount = 0

    tokens = nltk.word_tokenize(text)

    print("\n")

    for idx, token in enumerate(tokens):
        if token in positive_words:
            if idx>0:
                # - Positive words:
                # * a positive word not preceded by a negation word (i.e. not,
                n't, no, cannot, neither, nor, too)
                if tokens[idx-1] not in negative_words:
                    posWordCount += 1
                else:
                    #- Negative words:
                    # * a positive word preceded by a negation word
                    negWordCount += 1
            else:
                posWordCount += 1
        elif (token in negative_words):
            if idx>0:
                #- Negative words:
                # * a negative word not preceded by a negation word
                if tokens[idx-1] not in negative_words:
                    negWordCount += 1
                else:
```

```
# - Positive words:
# * a negative word preceded by a negation word (ex -not bad)
    posWordCount += 1
    else:
        negWordCount += 1

if(posWordCount > negWordCount):
    sentiment = "positive"
elif(posWordCount <= negWordCount):
    sentiment = "negative"

return sentiment
```