

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВАСИЛЯ СТУСА
Кафедра загального та прикладного мовознавства і слов'янської філології

КРАВЧЕНКО ЯНА СЕРГІЇВНА

Допускається до захисту:
в.о. завідувача кафедри
загального та прикладного
мовознавства і слов'янської
філології, к.філол.н., доцент
_____ Гарбера І.В.
«_____» _____ 2024 р.

СЕНТИМЕНТ-АНАЛІЗ
ТЕКСТІВ КОНФЛІКТНОЇ КОМУНІКАЦІЇ В INSTAGRAM

Спеціальність 035 Філологія
Освітньо-професійна програма 035.10 Прикладна лінгвістика

Кваліфікаційна робота

Науковий керівник:
Гарбера І.В., в.о. завідувача кафедри
загального та прикладного
мовознавства і слов'янської
філології, к.філол.н., доцент

Оцінка: _____ / _____ / _____
Голова ЕК: _____

Вінниця 2024

АНОТАЦІЯ

Кравченко Я. С. Сентимент-аналіз текстів конфліктної комунікації в Instagram. Спеціальність 035 «Філологія», Освітня програма «Прикладна лінгвістика». Донецький національний університет імені Василя Стуса, Вінниця, 2024.

У кваліфікаційній роботі досліджено особливості мовних та графічних засобів у конфліктній комунікації в Instagram. Проаналізовано бібліотеки VADER та TextBlob як інструменти для автоматичного сентимент-аналізу негативних коротких текстів із соцмереж.

Ключові слова: емотикони, соцмережа Instagram, конфліктна комунікація, мовні маркери, негативна полярність, сентимент-аналіз, тональний аналіз тексту, цифрова комунікація.

Бібліографія: 50.

Kravchenko Y. S. Sentiment Analysis of Conflict Communication Texts on Instagram. Specialty 035 «Philology», Programme «Applied Linguistics». Vasyl' Stus Donetsk National University, Vinnytsia, 2024.

The qualification work investigates the features of linguistic and graphic means in conflict communication on Instagram. The VADER and TextBlob libraries are analyzed as tools for automatic sentiment analysis of negative short texts from social networks.

Keywords: emoticons, social network Instagram, conflict communication, linguistic markers, negative polarity, sentiment analysis, tonal analysis of text, digital communication.

Bibliography: 50.

ЗМІСТ

Вступ	5
Розділ 1. Дослідження мовних та графічних засобів на позначення емоцій або оцінки	8
1.1. Мовні маркери вираження прояву емоцій та почуттів людини	8
1.2. Емоційно-смісловий аналіз лексичних одиниць із залученням тонального словника	14
1.3. Графічні засоби передачі емоцій та їх роль у комунікації Instagram	17
Розділ 2. Огляд задачі сентимент-аналізу	19
2.1. Поняття, сутність та рівні аналізу тональності текстів	19
2.2. Класифікація систем сентимент-аналізу тексту	21
2.3. Існуючі системи сентимент-аналізу	22
Розділ 3. Реалізація та практичне застосування програмного коду сентимент-аналізатора текстів	25
3.1. Опис ідеї розробки проекту сентимент-аналізу коментарів у Instagram	25
3.2. Проектування та створення програмного коду для аналізу тональності тексту	28
3.3. Збір та підготовка даних для вибірки текстових коментарів	43
3.4. Експериментальне дослідження вибірки та оцінка результатів	48
Висновки	64
Список використаної літератури	66
Додатки	73

Додаток А. Вибірка негативних коментарів до посту новинного акаунту CNN в Instagram та результати сентимент аналізу за допомогою бібліотеки TextBlob	73
Додаток Б. Вибірка негативних коментарів до посту новинного акаунту CNN в Instagram та результати сентимент аналізу за допомогою бібліотеки VADER	75
Додаток В. Вибірка негативних коментарів до посту акаунту південно-корейської зірки в Instagram та результати сентимент аналізу за допомогою бібліотеки TextBlob	77
Додаток Г. Вибірка негативних коментарів до посту акаунту південно-корейської зірки в Instagram та результати сентимент аналізу за допомогою бібліотеки VADER	79
Додаток І. Вибірка негативних коментарів до посту спортивного новинного акаунту ESPN в Instagram та результати сентимент аналізу за допомогою бібліотеки TextBlob	81
Додаток Д. Вибірка негативних коментарів до посту спортивного новинного акаунту ESPN в Instagram та результати сентимент аналізу за допомогою бібліотеки VADER	83
Додаток Е. Апробація	85

ВСТУП

Сучасна комунікація, особливо в цифровому середовищі, все частіше стає ареною для конфліктів. Подібні дискурси проявляються як у міжособистісних взаємодіях, так і в публічному середовищі. Мовні та графічні засоби стали трансляторами емоцій та настроїв, які можуть сприяти взаєморозумінню, або ж загострювати протистояння. Розуміння того, як вищезгадані засоби впливають на характер комунікації, стає надзвичайно важливим у контексті цифрової грамотності та запобігання ескалації конфліктів.

Актуальність дослідження. Активний розвиток штучного інтелекту та провідних ІТ-технологій є тісно пов'язаним із обробкою природної мови. Саме це стало фундаментом для нових напрямків досліджень у прикладній лінгвістиці. Різноманітні методи та алгоритми обробки природної мови сприяють результативному аналізу та ефективній інтерпретації текстів.

Стрімке зростання соціальних мереж сприяє розростанню комунікаційних конфліктів, адже користувачі можуть не тільки поширювати свої фото, але й висловлюватися, у позитивному чи негативному світлі, щодо певних новинних публікацій, поглядів інших людей тощо. Завдяки розвитку автоматизованих технологій, таких як сентимент-аналіз, стало можливим дослідження емоційного забарвлення повідомлень, їхнього впливу на розвиток конфлікту та способів його вирішення. Аналіз мовних та графічних засобів дозволяє зрозуміти, як формується конфліктний наратив та які механізми використовуються для його інтенсифікації чи пом'якшення.

Об'єктом дослідження є процес автоматизації сентимент-аналізу тексту.

Предметом дослідження є емоційне забарвлення текстів, мовні та графічні засоби на позначення емоції чи оцінки.

Метою дослідження є вивчення мовних та графічних засобів конфліктної комунікації, огляд типів систем сентимент-аналізу тексту та

аналіз емоційного чи оцінного настрою коментарів користувачів соціальної мережі Instagram.

Для реалізації поставленої мети необхідно вирішити **завдання**:

- дослідити поняття та сутність тонального аналізу тексту;
- проаналізувати класифікацію систем sentiment-аналізу тексту;
- розробити sentiment-програму аналізу відгуків;
- зібрати вибірку коментарів у Instagram, які містять мовні засоби на позначення негативних оцінки або емоцій;
- проаналізувати результати автоматизованого sentiment-аналізу оцінки або емоційності в зібраному матеріалі;
- з'ясувати, які з інструментів є більш точними для автоматичного тонального аналізу коментарів у соцмережах;
- проілюструвати результати дослідження вибірки зразків конфліктної комунікації.

Під час виконання вищевказаних завдань доцільно було застосувати основні **методи досліджень**, а саме структурний метод (застосовувався для дослідження внутрішньої організації негативних коментарів, як-от аналіз частотності використання ключових лексичних одиниць у конфліктній комунікації), функціональний метод (практикувався для вивчення комунікативної ролі негативних коментарів у різних контекстах публікацій) та конструктивний метод (побудова алгоритмів автоматизованого тонального аналізу з використанням бібліотек TextBlob і VADER).

Наукова новизна дослідження полягає у кількох аспектах. По-перше, проведення дослідження особливостей негативного дискурсу в трьох типах публікацій: політичних та спортивних новинних акаунтах, та особистих сторінках – буде проведено вперше. Така вибірка тем дозволяє виявити те, якою є специфіка публікацій та як вона впливає на характер та інтенсивність негативу. По-друге, порівняльний аналіз двох популярних бібліотек для автоматичного sentiment-аналізу TextBlob і VADER у контексті негативного

дискурсу: вперше оцінюється їхня ефективність у виявленні негативу в коротких коментарях, характерних для соціальних мереж.

Практичне значення отриманих результатів. Вивчення методів сентимент-аналізу текстів може допомогти науковцям розв'язати питання, що пов'язані із машинним навчанням, адже досі не розв'язана проблема того, аби навчити комп'ютер сприймати живу природню мову на людському рівні. Також аналіз тональності текстів дає змогу розв'язати безліч інших проблем, які пов'язані із знаходженням істинного емоційного забарвлення висловлюваного.

Ілюстрування аналізу вибірки дає змогу обґрунтувати вибір оптимального інструменту для аналізу текстів такого типу, як короткі коментарі в соцмережах. Виділені мовні маркери негативу можуть бути інтегровані в алгоритми автоматичного виявлення токсичних коментарів. Це дозволить модераторам соціальних мереж, як-от Instagram, ефективніше ідентифікувати образливі або агресивні висловлювання, враховуючи їхню специфіку в цифровій комунікації.

Структура роботи. Робота складається із вступу, трьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг роботи становить 85 сторінок комп'ютерного тексту. Робота містить 8 таблиць, 35 рисунків, 7 додатків; список використаних джерел містить 50 найменувань.

Розділ 1. ДОСЛІДЖЕННЯ МОВНИХ ТА ГРАФІЧНИХ ЗАСОБІВ ДЛЯ ПОЗНАЧЕННЯ ЕМОЦІЙ АБО ОЦІНКИ

1.1. Мовні маркери вираження прояву емоцій та почуттів людини

Українська мова надзвичайно багата на експресивні мовні засоби, які становлять фонд мовних одиниць, що виступають засобами вираження ставлення мовця до певного суб'єкта. Емоційні переживання людини важливо передавати в комунікації, адже саме вони відображають дійсність.

Сучасна лінгвістика починає розвивати дослідження емотивних мовних маркерів, адже цей напрям сформувався порівняно недавно. Лінгвістика емоцій як окрема наукова галузь почала розвиватися у ХХ столітті, спираючись на здобутки психології та традиційного мовознавства. До середини 70-х років питання вираження емоцій у мовленні не було центральним об'єктом досліджень, а праці, що розглядали цю тематику, з'являлися епізодично та не викликали широкого зацікавлення.

Загальна картина мовного світу демонструє, що у більшості мов кількість емотивних одиниць із негативною оцінною семантикою значно перевищує кількість одиниць із позитивною оцінкою. Це свідчить про певну асиметрію в репрезентації емоційного досвіду засобами мови.

Серед дослідників, які зробили вагомий внесок у розвиток теорії емотивності, важливе місце посідає В. І. Шаховський. Учений розглядає емотивність як семантичну характеристику мови, що дозволяє виражати емоції — як соціальні, так й індивідуальні. На його думку, емотивність є інтегрованою властивістю мовної системи, яка проявляється у значеннях мовних одиниць і відображає психологічні аспекти емоційного досвіду. Під час дослідження емотивності в художніх текстах важливо враховувати як особливості текстової семантики, так і природу людських емоцій. Основними характеристиками емоцій, які знаходять своє відображення у мовленні, є їхня суб'єктивність та зв'язок із конкретними предметами чи ситуаціями. Таким

чином, вивчення емотивності в лінгвістиці відкриває нові горизонти для аналізу не лише мови, а й людської психіки в цілому [16].

Деяким масивам лексичних одиниць притаманні ознаки експресії та емоції. Наприклад, експресією є саме сила прояву почуттів чи переживань, тобто така інтенсивність завжди має оцінний характер. Такі слова можуть бути синонімами до звичайних нейтральних слів: пика – обличчя, воїн – військовий, голий – придатний, дівчинонька – дівчина, глава – керівник, панькатися – догоджати. Слова, що відносяться до категорії, яка виражає емоційну складову сказаного, окрім об'єктивного лексичного значення, мають суб'єктивне забарвлення, адже є засобом вираження особистого ставлення мовця до події чи явища. Людина завжди користується експресією, аби висловлювати свої позитивні чи негативні почуття, адже саме сила вияву супроводжує вияв емоцій. Тобто, емоційна та експресивна лексика мають швидше умовне розмежування та можуть належати до однієї категорії слів, а саме емоційно-експресивної лексики.

У сучасному світі є потреба не тільки ділитися своїми емоціями, але й існує бажання поширювати особисті оціночні судження щодо певних товарів, послуг чи подій. Загалом, об'єктом оцінки для мовця може стати будь-який предмет чи особа, які викликають у нього певні позитивні чи негативні емоції.

У сучасній лінгвістиці, категорію оцінки співвідносять як процес позитивної чи негативної кваліфікації мовцем предмета думки, відношення особи до нього, наприклад, схвалення/несхвалення, бажане/небажане, корисне/некорисне, цікаве/нецікаве, потрібне/непотрібне, якісне/неякісне, тощо. Результатом цього процесу є оцінка. Оцінний аспект полягає у тому, що об'єктивний світ мовця поділяється умовно на добро і зло, користь і шкоду, прийнятність та неприйнятність. Такі судження є соціально зумовленими та фіксуються за допомогою мовних видів писемності. Тобто, категорія оцінки взаємодіє із текстом та оцінка має повну реалізацію у ньому.

Оцінні процеси формуються на основі протиставлення позитивних і негативних якостей, таких як «добре/погано», що інтуїтивно сприймаються

людиною. Оцінка відіграє ключову роль у пізнанні навколишньої дійсності, оскільки саме, через оцінювання, визначається цінність різних предметів, дій або характеристик у загальній картині світу суспільства.

Залежно від ставлення мовця до об'єкта оцінювання, оцінка поділяється на три основні типи: позитивну, негативну та нейтральну. Кожен із цих видів відображає різний емоційний та когнітивний підхід до сприйняття й інтерпретації дійсності, формуючи основу для комунікації та взаємодії [11].

Позитивна оцінка виражає схвальне ставлення до певного об'єкта, який оцінюють, а от негативна оцінка має несхвальне ставлення. Нейтральна оцінка може бути виражена як синонім до слова «байдуже», тобто суб'єкт не викликав у носія мови якихось емоцій чи переживань.

Існує думка, що оцінка може складатися у мовця під впливом суспільства. Особа існує у соціумі, тому має властивість опиратися на досвід інших людей. Досвід можна отримати тільки зазнавши певного емоційного збудження, тобто немає пізнання без емоцій. Так чи інакше, оцінка корелює із емоціями, вони реалізуються у емоційно-оцінній лексиці.

Емоційно-оцінна лексика є лінгвістичним феноменом, особливим компонентом, за допомогою якого виражається психічний стан мовця, його ставлення до предмета, об'єкта, адресата мовлення, тобто все те, що складає мету емотивної функції слова. Науковці виділяють такі основні риси та ознаки цієї категорії слів:

- полікомпонентність семантичної структури;
- наявність формальних ознак та неформальних показників експресивності слова;
- особливості контекстуального оточення [9].

Оцінна лексика має тісний зв'язок із емоційною, але зберігає свої унікальні особливості. Лінгвісти зазвичай виділяють два основні типи одиниць оцінної лексики:

- слова, що містять оцінний визначальний компонент (наприклад, красуня – дуже гарна дівчина);

- слова, в яких визначальний компонент є нейтральним щодо оцінки (наприклад, співак – той, що співає).

Емоційна лексика, у свою чергу, збагачує оцінні висловлювання емоційними відтінками й умовно поділяється на три категорії:

- слова, для яких емоційне забарвлення є частиною їхнього основного значення (наприклад, цей *багатий* походить із родини депутатів);
- слова, де емоційне значення виникає через метафоричне вживання (наприклад, «*залізна леді*» жінка, яка завжди залишатиметься вірною своїм принципам);
- слова, які набувають емоційного забарвлення завдяки додатковим елементам, як-от суфіксам (наприклад, одягну *хустиноньку* та піду в *садочок* співати).

Слова, що передають емоції чи оцінки, називають маркерами. Маркована лексика є яскравим проявом експресивної функції мови, яка реалізується через специфічні ознаки предметів і явищ, відображені в значеннях слів. Такі особливості закріплюються за окремими мовними формами та стають основою для вираження оцінних суджень і емоційних станів мовців [8].

Мовними маркерами вираження прояву емоцій та почуттів людини є групи слів, які означають назви певних почуттів і є емоційними вже за своїм лексичним значенням. Наприклад, любов, щирість, гнів, злість, ніжність, чуйність, жаль, радість, страх, сум, блаженство та інші. В українській мові є чимало слів із емоційним забарвленням, адже їх творенню сприяють суфікси пестливості, зневаги, тощо. Наприклад, личко, дівчинонька, хлопчисько, ледацюга. Емоційною лексикою є не тільки слова, які означають назви певних почуттів та емоцій, а й ті слова, які виражають позитивну чи негативну оцінку явищ і предметів дійсності або почуття й стан людини. Наприклад, добрий, сумно, любити, ненавидіти, подобатися, засмучений. Мовними маркерами оцінки або емоційності є й слова, які мають нейтральний синонім-відповідник, наприклад, бахкати – стукати, верещати – кричати, попертися – піти. Тобто, з

погляду на походження, емоційно-експресивна лексика сучасної української мови може бути неоднорідною.

У кожній розвиненій національній мові науковці виділяють одиниці або форми, які виконують переважно емоційно-оцінні функції. Українська мова має багатий арсенал таких засобів, які найбільш яскраво проявляються на лексичному рівні. Серед них:

- лексеми, що безпосередньо виражають почуття та настрої (наприклад, любов, ненависть, радість, смуток);
- афективна лексика, яка містить оцінку в своєму значенні (чудовий, жахливий, прекрасний тощо);
- якісні прикметники та прислівники з оцінними відтінками (мудріший, найпідліший, щонайкраще тощо);
- слова-інтенсифікатори, що підсилюють значення (наприклад, дуже, надзвичайно).

На рівні словотвору емоційно-оцінна функція реалізується через суб'єктивно-оцінні форманти (серденько, грошики, дівчисько, бородище тощо). Мовознавці традиційно розглядають категорію оцінки як позитивну або негативну характеристику об'єкта, яка безпосередньо пов'язана з відповідними емоціями [8].

Слід зазначити, що не тільки повноцінні слова можуть бути засобом вираження емоційної напруженості мовця, адже існують випадки, коли короткі вигуки, можуть бути ідентифікатором позитивних чи негативних почуттів. Вигуки служать засобами виразності не тільки стилістично, але й допомагають виражати почуття, називаючи чи не називаючи їх. Наприклад, вигук «тьху» може сигналізувати про негативне забарвлення думки об'єкта щодо певного суб'єкта, а вигук «вау» означає позитивний характер сказаного. Дуже частими є випадки, коли функцію вигуків у мовленні персонажів виконують повнозначні частини мови, які можуть бути прямими ідентифікаторами емоції, яку людина переживає зараз. Наприклад, «чудово», «супер», «лишенько».

Емоції можна виражати невербально, а саме через жести, міміку, пози тіла чи навіть інтонацію голосу. Ці способи дозволяють передавати відтінки настрою, не вдаючись до слів. Однак вербальні засоби комунікації займають не менш важливе місце, оскільки мова дає змогу називати емоції та почуття прямо, створюючи повноцінну картину стану, який переживає мовець у певний момент. Українська мова багата на лексичні засоби для вираження емоцій, що дозволяє створювати живе та емоційно насичене спілкування.

Використання емотивних маркерів є ефективним на будь яких рівнях мови, тому комунікація – це один із найкращих та найпростіших способів репрезентувати та донести свої емоції чи почуття до інших.

Отже, використання мовних засобів для вираження емоцій є не лише природним процесом, а й надзвичайно важливим елементом людської взаємодії, який глибоко впливає на формування стосунків між людьми. Мова дозволяє не просто передати інформацію, а й створити емоційний зв'язок, який сприяє більш ефективному розумінню, співпереживанню та взаємодії.

1.2. Емоційно-смісловий аналіз лексичних одиниць із залученням тонального словника

Опрацювання природної мови у напрямку тонального аналізу все більше стає поширеним не тільки суто у лінгвістичних працях, але й охоплює маркетингові чи соціологічні проєкти. Наприклад, були проведені дослідження емоційного забарвлення відгуків користувачів про готелі, фільми, продукти, послуги та ін. Лексичні одиниці у коментарях набувають набагато більшого сенсу, коли розглядаються під призмою емоційного забарвлення мовного потоку.

Задля опрацювання та визначення мовних маркерів емоційності тексту, потрібно укласти такий словник, що буде містити в собі певну оцінку для кожної лексичної одиниці. Тональний словник, у найпростішому його вигляді, становить список слів і словосполучень зі значенням тональності кожного слова. Це може бути як числове значення в межах певної шкали (наприклад, від 1 до 10, де 1 – негативне слово, а 10 – позитивне), так і категорія (наприклад, позитивне чи негативне слово). Кожному слову приписується його частиномовна характеристика та початкова форма у випадку, коли словник містить різні словоформи одного слова [17].

Тональні словники можуть бути створеними лише на основі самостійних частин мов, в той час, коли службові частини мови будуть вважатися «шумом» для аналізатора та не будуть відібрані до обробки. На практиці вже є створені праці на базі подібних словників, тобто при дослідженні науковці зосереджувалися на найчастотніших іменниках, прикметниках, прислівниках та дієсловах, що були в досліджуваному матеріалі та мали відповідник у тональному словнику. Загалом, до кожної зазначеної у словнику лексичної одиниці є записані дані щодо її частиномовної приналежності, тональності та сили тональності за шкалою від 1 до 3. Існує тенденція поділяти прислівники та прикметники на такі групи:

- позитивні;

- негативні;
- ті, що слугують для підсилення тональності.

Іменники є умовно поділеними на:

- позитивні;
- негативні;
- потенційно позитивні (якщо знаходяться в позитивному оточенні);
- потенційно негативні (якщо знаходяться в негативному оточенні).

Дієслова ж поділяють в залежності від їх ролі та оточення в тексті та виділяють окремо дієслова-зв'язки. Існують лексичні одиниці, тобто слова-інвертори, які мають здатність змінювати тональність сусідніх компонентів на протилежну тональність.

Словники тональності, які створені на базі лексичного підходу, можуть бути створені вручну, або автоматично, за використання коренів слів розширитися [41].

Тональні словники можна укладати вручну, аби забезпечити полісемантичну тематику слів, але якщо немає потреби у багатогранності лексичних одиниць, то можна згенерувати тональний словник за допомогою машинного навчання на певну вузьку тему автоматично. У таких випадках застосовуються онтології. Як правило, лінгвістична онтологія є терміном, який вживається в інформатиці. Цей термін розуміють як спеціалізований інформаційно-пошуковий тезаурус для автоматичної обробки текстів, тобто це є різновидом програмного забезпечення. На жаль, є певні недоліки такого способу, тому варто залучити анотовані корпуси. Ефективність генерації за допомогою корпусів значно зростає, адже вона базується на об'єктивній думці експерта-лінгвіста.

По-перше, сентимент-анотований корпус є джерелом контекстів для слів, які потраплять до тонального словника, що дає можливість краще визначити емоційне забарвлення кожного слова, а в онтології переважно

містять ізольовані слова. По-друге, корпус подає велику кількість нелітературних слів, які важливі для визначення емоційного забарвлення повідомлення, але відсутні в онтологіях [17].

Сентимент-анотований корпус – це корпус текстових повідомлень, у якому кожному повідомленню присвоєно значення емоції, яку воно передає. Створення сентимент-анотованого корпусу – це невід’ємна частина реалізації системи емоційно-сміслового аналізу [17].

Такий корпус надає розуміння машині того, як саме людина виражає свої почуття, емоції, оціночні судження щодо певного об’єкта, тобто сентимент-анотований корпус виділяє не тільки слова маркери емоційного стану, але й роз’яснює емоційне забарвлення емотиконів, вживання тих чи інших розділових знаків у реченні. На базі цього, машина може навчатися генерувати з більшою ймовірністю правильне визначення тональності слів.

У сентимент-анотованих корпусах до кожного слова прикріплюють кілька атрибутів. Такими атрибутами можуть бути:

- частина мови (наприклад визначають, іменник, прикметник, дієслово, прислівник, займенник, сполучник, прийменник, частка, числівник, вигук та інше. Інше може бути – емотикони та інші графічні позначки, які містять у собі емоційне забарвлення);
- емоція (наприклад, радість, злість, сум);
- емоційне забарвлення (наприклад, позитивне (тобто слова чи речення мають позитивне значення), негативне (тобто слова чи речення мають негативне значення), нейтральне (тобто слова чи речення мають нейтральне значення, не виражаючи позитивну чи негативну оцінку)).

Іноді науковці в атрибуті «емоційне забарвлення» виділяють і такі лексичні одиниці як:

- слова-підсилювачі (слова, які не мають самостійного емоційного забарвлення, але можуть підсилювати емоційне

забарвлення пов'язаного із ними компонента. Наприклад, дуже, вкрай, тощо);

- слова-інвертори (слова, які не мають самостійного емоційного забарвлення, але можуть змінювати емоційне забарвлення пов'язаного із ними компонента. Наприклад, не, нема, нереально, тощо).

1.3 Графічні засоби передачі емоцій та їх роль у комунікації Instagram

Стрімка популяризація соціальних мереж сприяє динамічному використанню мови користувачами. Особливістю комунікації в соціальних мережах, зокрема в Instagram, є не лише стислість висловлювання, але й явні риси усного мовлення. Користувачі активно використовують графічні засоби передачі емоцій, адже через екран стає недоступно передати емоцію, жести чи тембр свого голосу. Таким чином з'явилася потреба у приближенні друкованого тексту до реального мовлення. Задля найбільш точної передачі настрою, користувачі почали використовувати емотикони.

Емотикони (від англ. „emoticon = emotion icon” — емоційна іконка або, як їх часто називають у просторіччі, — „смали”, „смайлики” (від англ. „Smile” — посмішка :-)) — кластери типографічних символів, які наявні на клавіатурі комп'ютера та мобільного телефону і які компенсують нестачу паралінгвістичних засобів спілкування в мережі та виражають емоційне ставлення комуніканта до сказаного [48].

Від традиційних прикладів графічної інформації — емотикони відрізняються, бо вони мають обмежену кількість та їхні значення є чітко прописаними, тобто можуть конвертуватися у текстову інформацію. Зазвичай дослідники укладають довідники-списки позитивних чи негативних емотиконів, аби потім ідентифікувати ці символи у наборі текстової інформації. Тобто, будь-яка емоція-емотикон може бути виявлена: як негативна, якщо вона є знайденою у довіднику-списку із негативними

емотиконами та як позитивна, якщо вона є знайденою у довіднику-списку із позитивними емотиконами.

Насправді смайлики відіграють важливу роль у визначенні тональності тексту, адже вони несуть у собі значний емоційний сенс. Емотикони набули широкого вжитку серед користувачів мережі Інтернет, тому не є доречним вилучати їх із аналізу тексту, як «шум».

За допомогою смайлика автор констатує ту емоцію, яку хотів передати реченням, наприклад: :-)) — означає посмішку, жартівливе ставлення до сказаного; :-(— сумне повідомлення; :)) — позначає почуття щастя тощо [48].

Багато лінгвістів стверджують, що емодзі — це шлях до створення нової мови, що означає, що люди знову переходять від слів до символів і зображень. Спосіб нашого спілкування в соціальних мережах змінюється. Нова мова зображень стає такою популярною останнім часом завдяки використанню емодзі та піктограм замість довгих пояснень шляхом написання довгих текстів. Люди використовують смайлики з тієї причини, що їм потрібно менше пояснень, менше часу та місця, а почуття передаються значною мірою. Люди реагують на смайлики так само, як на реальне людське обличчя [34].

Використання смайлів стало частиною культури спілкування в соціальних мережах. Смайлики подолали мовні бар'єри, оскільки вони мають форму малюнків або зображень, які є універсальними, і, майже, всюди можна зрозуміти повідомлення. Ніл Кон стверджує, що коли емодзі з'являються разом із текстом, то вони часто доповнюють або покращують текст. Це схоже на жести, які з'являються разом із мовою. Протягом останніх трьох десятиліть дослідження показали, що наші руки надають важливу інформацію жестами. Отже, емотикони мають потенціал стати вагомим носієм інтернаціональної інформації [34].

Розділ 2. ОГЛЯД ЗАДАЧІ СЕНТИМЕНТ-АНАЛІЗУ

2.1. Поняття, сутність та рівні аналізу тональності текстів

Сентимент-аналіз тексту (тональний аналіз тексту, англ. Sentiment analysis, англ. Opinion mining) – це область досліджень, яка аналізує думки людей, настрої, оцінки, ставлення та емоції щодо таких об'єктів, як продукти, послуги, організації, особи, проблеми, події, теми та їх атрибути. Із 2000 року ця область досліджень розвивається та передбачає залучення низки методів аналізу природної мови в комп'ютерній лінгвістиці, задля виявлення у контенті емоційно забарвленої лексики, або ж оціночного маркера щодо ставлення автора тексту до об'єктів про які йде мова [25].

Тональність – це особисте ставлення автора щодо певних об'єктів, про які йде мова в його висловлюванні. Лексична тональність – це емоційна складова тексту, яка є вираженою на рівні лексем або ж комунікативного фрагмента. Для визначення тональності всього тексту, тобто чи має заданий матеріал позитивне чи негативне забарвлення, створюють функцію (або ж суму) із лексичних тональностей, тобто речень та правил їх поєднання.

Наразі, сентимент-аналіз тексту – це величезне поле для обробки природної мови (Natural language processing, NLP), включаючи аналіз мікротексту, виявлення сарказму, виявлення суб'єктивності та виділення аспектів; це системи, що мають бути спроможними ідентифікувати та «витягти» думки із тексту [27].

Із теоретичної точки зору, аналіз тональності висловлюваного є цікавим для вивчення напрямком у прикладній лінгвістиці, але, з іншого боку, має багато способів практичного застосування. Оскільки публічна інформація в мережі Інтернет постійно зростає, з'являється багато емоційно забарвлених текстів, відгуків, висловлювань, рецензій, тощо. За допомогою систем сентимент-аналізу текстів, ці масиви незгрупованої інформації можна автоматично групувати задля створення повноцінної картини даних про

громадську думку щодо певного об'єкта: продукції, послуги, події, тощо. Така відсортована інформація може стати у нагоді для маркетингового аналізу, комерційних досліджень та огляду настроїв громадськості.

При сентимент-аналізі тексту зазначають, що інформацію поділяють на два основні типи: факти і думки [25]. Загалом, фактами є певні об'єктивні вирази автора про щось, а от думка виражається, як правило, суб'єктивними виразами, які описують оцінку, почуття чи ставлення мовця до теми.

Більше того, тональний аналіз тексту можна здійснювати на різних рівнях. Загалом можна виділити аналіз на таких рівнях [26]:

- на рівні речення;
- на рівні документу;
- на рівні об'єкту та аспекту.

Тональний аналіз на рівні речення дозволяє визначити настрій кожного речення в документі окремо. Емоційна оцінка дається як така: позитивна, негативна, нейтральна. Сентимент-аналіз на рівні документу здійснюється, щоб визначити загальний настрій всього документу. Наразі, такий рівень аналізу є ефективним при оцінці настрою щодо одного суб'єкта (продукту, послуги, тощо). Аналіз на рівні документу не є ефективним при оцінці декількох суб'єктів.

Сентимент-аналіз на рівні документу та на рівні речення не дає змогу визначити що саме подобається автору, а от аналіз на рівні об'єкту та аспекту дає можливість хоча б визначити сентимент, що стосується безпосередньо об'єкту висловлення. Наприклад речення «Мені подобається ця пісня». Загальний настрій речення позитивний, адже спостерігається позитивний сентимент стосовно пісні. Тобто, метою тонального аналізу на рівні об'єкту та аспекту є виокремлення сентименту (позитивного чи негативного) від об'єкту висловлення або його аспекту.

2.2 Класифікація систем сентимент-аналізу тексту

Основні розробки в області тонального аналізу можна класифікувати, як:

- системи на основі правил, що виконують тональний аналіз тексту за допомогою набору правил, які були вручну розроблені лінгвістами;
- автоматичні системи, які базуються на методах машинного навчання;
- гібридні системи, що поєднують у собі обидва наведені вище підходи, тобто машинне навчання із набором ручних правил.

Загалом, системи на основі правил, мають певний набір правил, який ідентифікує суб'єктність, полярність або предмет думки. Правила можуть використовувати різні входи, такі як: класичні методи NLP (маркування, розбір мовлення та ін.) та лексикони (списки слів та виразів).

Для таких систем варто розробляти набір правил із врахуванням усіх особливостей майбутнього тонального аналізу матеріалів, адже додавання нових правил може спровокувати небажані наслідки, через невдалу взаємодію нових та раніше створених правил. Системи на основі правил потребують постійного моніторингу людиною та не здатні самонавчатися.

Автоматичні системи не мають тенденції покладатися на правила, які створюють вручну. Машинне навчання є основою для таких систем. Тональний аналіз на основі автоматичних систем здійснюється за методом класифікації, тобто коли класифікатор подається текстом і повертає відповідну категорію, наприклад у випадку проведення аналізу полярності тексту, позитивний, негативний чи нейтральний. Автоматичні системи мають здатність самонавчатися.

Гібридні системи є поєднанням машинного навчання із набором ручних правил. Аналіз на базі таких систем, на мою думку, є найбільш ефективним. Гібридні системи не просто здатні самонавчатися, але й мають можливість удосконалення результатів цього навчання за допомогою кваліфікованого

фахівця. У такому випадку, людина безпосередньо бере участь в ефективному навчанні машини, що не тільки пришвидшує результативність, але й правильність аналізу.

1.3 Існуючі системи сентимент-аналізу

Застосування низки різноманітних ІТ-інструментів допомагає вченим не тільки зрозуміти колективну поведінку та оцінити настрої громадської реакції на маркетингові заходи, але й розвивати значний потенціал практичного застосування створення і розвитку технологій обробки тексту. Незважаючи на численні способи використання аналізу тональності, проєктів, які б зосередили свою увагу на емоційному відклику для україномовних текстів, є надзвичайно мало [12].

Світ ніколи не стоїть на місці, тому українські лінгвісти мають собі як приклад системи тонального аналізу від іноземних розробників. Наприклад, система SentiStrength, розроблена M. Thelwall, K. Buckley, G. Paltoglou та D. Cai. Початковою ціллю цієї розробки був сентимент-аналіз коротких структурованих неформальних текстів англійською мовою. Наразі, система навчається налаштовуватися так, аби могли працювати з рядом інших мов. Результат аналізу тексту видається у вигляді двох оцінок: негативної чи позитивної за п'ятибальною шкалою [45].

Програмна система ВААЛ розроблена з метою автоматизованого прогнозування ефекту неусвідомлюваного впливу текстів на масову аудиторію. Для кожного виду психоемоційного впливу в системі передбачена своя категорія. Ступінь належності тексту до певної категорії подано шкалою оцінки даного виду впливу. Кожна шкала задається двома протилежними за змістом поняттями (наприклад, хороший – поганий, гарний – відштовхуючий, легкий – важкий і тд.). Такий функціонал не повністю відповідає вирішенню задачі визначення емоційного забарвлення текстових даних, які створені природною мовою, але є достатньо близьким до неї. В основі системи

оцінювання емоційного впливу алгоритми, які базуються на аналізі емоційного впливу звуків в окремих словах. Тобто основною одиницею, що формує вплив тексту на підсвідомість людини в даному випадку є звук. Для розв'язання задачі автоматизованого визначення емоційного забарвлення такий підхід не є коректним. Емоційне забарвлення тексту формується за допомогою емоційно забарвлених слів, а також зв'язків між ними та нейтрально забарвленими словами. Тобто, елементарною одиницею, що робить внесок у ту чи іншу емоційну забарвленість, повинно бути слово [2].

YouScan – це система для аналітики соціальних медіа на базі штучного інтелекту. Вона, ніби виконує роль спостерігача та займається моніторингом відгуків користувачів, які публікуються у соціальних мережах. Така система може не тільки дати змогу бренду бути в курсі будь-якого його згадування у мережі, але й визначає тональність відгуку, тобто позитивний чи негативний настрій. YouScan спрямована не тільки на визначення оцінки тексту, але й на визначення того, що саме спровокувало такі настрої у споживача. Користувач системи має змогу дізнатися, як клієнти сприймають різні аспекти його продукту чи сервісу, навіть сприйняття від якості та ціни до смаку, клієнт-сервісу та упаковки. Сервіс YouScan позиціонує себе як високоточна система визначення тональності тексту та оцінюють точність оцінки тексту на 95% [50].

The Cloud Natural Language API система, що має кілька методів для виконання аналізу та анотації вхідного тексту. Кожен рівень аналізу надає цінну інформацію для розуміння мови. Сентимент-аналіз перевіряє даний текст і визначає емоційну думку, що переважає в тексті. Загалом застосовується для визначення ставлення письменника як позитивного, негативного чи нейтрального. Система надає оцінку настрою документа, яка вказує на загальну емоцію документа. Важливо зазначити, що The Cloud Natural Language вказує на відмінності між позитивними та негативними емоціями в документі, але не визначає конкретні позитивні та негативні емоції. Наприклад, «злий» і «сумний» обидва вважаються негативними емоціями.

Однак, коли система аналізує текст, який вважається «злим», або текст, який вважається «сумним», відповідь лише вказує, що настрої в тексті є негативними. Документ з нейтральною оцінкою (близько 0,0) може вказувати на документ з низьким рівнем емоцій або може вказувати на змішані емоції з високими позитивними і негативними значеннями, які скасовують кожне. У системі є можливість порівнювати різні документи за тональністю [44].

Системи сентимент-аналізу тексту є важливим інструментом для автоматизованого розпізнавання емоційного забарвлення текстів, що знаходить застосування у багатьох сферах, зокрема аналізі соціальних мереж та дослідженнях громадської думки. Серед основних систем, які використовуються для аналізу тональності тексту, можна виокремити такі, як TextBlob та VADER (Valence Aware Dictionary and sEntiment Reasoner).

TextBlob забезпечує зручність у використанні й дозволяє проводити аналіз тексту з точки зору полярності (позитивна, негативна, нейтральна) та суб'єктивності. Однак його слабкою стороною є недостатня адаптованість до специфіки коротких текстів та розмовного стилю, характерних для соціальних мереж.

VADER, у свою чергу, розроблений спеціально для аналізу коротких текстів, включно з твітами, коментарями та іншими цифровими форматами. Його перевагою є врахування сленгу, емодзі та великих літер, що дозволяє більш точно оцінювати емоційне забарвлення текстів у контексті цифрової комунікації.

Отже, розвиток систем сентимент-аналізу продовжує прогресувати. Завдяки інтеграції машинного навчання, нейронних мереж і лексикографічних підходів, розширено можливості аналізу емоційного контексту текстів у різноманітних форматах.

Розділ 3. РЕАЛІЗАЦІЯ ТА ПРАКТИЧНЕ ЗАСТОСУВАННЯ ПРОГРАМНОГО КОДУ СЕТИМЕНТ-АНАЛІЗАТОРА ТЕКСТІВ

3.1. Опис ідеї розробки проєкту сентимент-аналізу коментарів у Instagram

Сентимент-аналіз розпочинав свій розвиток в галузі комп'ютерних наук, але пізніше зацікавив дослідників напрямку гуманітарних наук. Зазвичай емоції, пізнання та поведінка є взаємопов'язаними, тому сентимент-аналіз стає у нагоді для того, щоб краще зрозуміти та в майбутньому передбачити поведінку людини, щоб змінити чи скорегувати реакцію індивіда на певні речі.

Тональність тексту може бути прийнята для позначення почуттів, що лежать в основі головної думки людини. Сентимент може бути позитивним, негативним чи нейтральним, та підсумовувати почуття, які хотів виразити автор думки. Наразі, сентимент-аналіз може застосовуватися не лише в академічних колах. Зрозуміло, що, через збільшення популярності соціальних мереж у повсякденному житті, перед науковцями постає завдання дослідити масиви мовних матеріалів, які залишають користувачі кожного дня, щоб у режимі реального часу спостерігати за динамікою зміни мови та передбачати майбутні виклики [35].

Аналіз тональності коментарів у соціальній мережі може стати ключовим інструментом для того, аби можна було зрозуміти основний настрій та емоційний стан аудиторії. Соціальні мережі стали головними майданчиками для обміну інформацією. Саме там, часто можна спостерігати формування громадської думки та комунікації між найрізноманітнішими прошарками суспільства. Instagram, безперечно, є однією із найпопулярніших соціальних мереж у світі. У цьому середовищі щодня генеруються мільйони постів, коментарів та реакцій. Свобода слова в Instagram спонукає користувачів активно висловлювати свої думки, надавати оцінки та виражати емоції. Для бізнес-користувачів (брендів, блогерів, підприємств, компаній і т.д.) це джерело інформації є надзвичайно важливим, щоб мати можливість зрозуміти

та проаналізувати те, як саме їхній контент або продукт впливає на аудиторію. З одного боку, негативні відгуки можуть вказувати бізнесу на проблеми, які потрібно якнайшвидше вирішити у продукті, а, з іншого боку, позитивні коментарі допоможуть зрозуміти та підкреслити сильні сторони продукту, щоб потім можна було акцентувати увагу користувачів на переваги продукту в майбутніх рекламних кампаніях.

Сентимент-аналіз, тобто визначення полярності тексту (позитивної, негативної чи нейтральної), дає змогу не лише ідентифікувати настрої користувачів, а й дозволяє лінгвістам провести дослідження того, як саме люди висловлюють свою думку та виражають емоції в цифровому середовищі. Більше того, опрацювання великого масиву коментарів мануальним методом є складним і часозатратним завданням. Отже, автоматизований сентимент-аналіз стає оптимальним рішенням для того, щоб науковці могли досліджувати велику кількість матеріалів, залучаючи менше ресурсів. Програми, що працюють на основі штучного інтелекту, спрямовані на те, щоб швидко та точно обробляти величезні обсяги даних, надаючи лінгвістам більше можливостей та опрацьованої інформації для дослідження певних мовних явищ.

Розробка програмного коду, що створений для того аби аналізувати настроїв висловлюваного, має оптимізувати проведення ряду лінгвістичних досліджень, та може мати такі сильні сторони як:

- швидкість опрацювання (автоматизована система за кілька секунд може не лише прочитати великий за обсягом текст, а й виділити слова-маркери настрою);
- об'єктивність (алгоритм аналізує текст без упереджень, враховуючи лише його зміст, тональність та ключові теми).

Дослідження конфліктної комунікації в Instagram потребує обробки великої кількості реальних коментарів. Завдання постає в тому, щоб визначити емоційне забарвлення висловлюваного користувачами в мережі. Важливим є й те, що на меті є проаналізувати та визначити мовні маркери

негативної комунікації, дослідити вибірку даних за різними критеріями, щоб сформулювати висновки, щодо «природи» негативної комунікації в цифровому середовищі.

Зрозуміло, що сентимент-аналіз не є простою задачею, якщо нею займатися вручну. По-перше, потрібно уважно вчитати велику кількість коментарів, по-друге, визначити мовні маркери, які передають емоційне чи оцінне забарвлення негативного висловлювання, по-третє, підрахувати самостійно тональну оцінку словам-маркерам у співвідношенні до кількості слів у тексті.

Було прийнято рішення розробити програмний код, який автоматично буде визначати тональність тексту. Основну ідею проєкту описано у таблиці 3.1.

Таблиця 3.1 – Опис ідеї проєкту

Зміст ідеї	Застосування	Зручності для користувача
Створення програмного коду, який автоматично аналізує тональність тексту та виділяє позитивні/негативні мовні маркери у коментарях	Тональний аналіз текстових коментарів в Instagram для визначення характеру оцінки користувача, щодо певного контенту	Можливість швидко визначити емоційне забарвлення вибірки коментарів, без мануального підрахунку полярності тексту

3.2 Проєктування та створення програмного коду для аналізу тональності тексту

У сучасному ІТ-світі розвивається напрям створення сентимент-аналізатора за допомогою API (Application Programming Interface), тобто спеціального протоколу для взаємодії комп'ютерних програм, який дозволяє використовувати функції однієї програми всередині іншої. У контексті аналізу та звітності даних API може застосовуватися для збору даних із різних джерел, щоб потім аналізувати ці дані таким чином, щоб надати користувачу точний результат. Загалом існує декілька варіантів проєктування систем аналізу настроїв через API. Їх можна поділити на:

- бібліотеки з відкритим кодом;
- комерційні бібліотеки.

Для обробки таких бібліотек застосовують мову програмування Python. Вона є однією із провідних мов для ефективної обробки та аналізу даних. Python об'єднує сильну спільноту програмістів, яка вільно ділиться навчальними посібниками та різноманітною документацією, тому мова програмування має великий набір інструментів для реалізації різнозадачних моделей NLP (Natural Language Processing). Python пропонує численні можливості для сентимент-аналізу, завдяки своїй великій екосистемі бібліотек та інструментів, призначених для обробки природної мови (NLP). Відомі бібліотеки, такі як NLTK (Natural Language Toolkit) та TextBlob надають зразки того, як побудувати функції та шаблони для сентимент-аналізу та обробки тексту, суттєво економлячи час розробника. З огляду на вище сказане, варто обрати Python для проєкту розробки автоматизованого аналізу текстів, оскільки ця мова надає пріоритет простоті та легкості розуміння алгоритму дій написання коду, стаючи доступним інструментом для лінгвістів із початковим рівнем програмування [37].

Отже, саме цю мову програмування буде застосовано для проєктування та створення коду-аналізатора тональності відгуків.

Для написання коду буде застосовуватися інтегроване середовище розробки PyCharm. Це середовище є досить потужним, адже безперебійно забезпечує аналіз коду, має графічний налагоджувач, який виділяє помилки в коді. PyCharm постачає постійну підтримку системи контролю версій для операційних систем macOS, Windows та Linux. У самому середовищі можна легко завантажувати різні модулі та бібліотеки для створення майбутнього коду. Середовище PyCharm є зручним для користувача, тому що воно оснащено зрозумілим інтерфейсом [36].

Для реалізації програмного коду сентимент-аналізу коментарів обрано дві з найбільш ємних за обсягом бібліотек, а саме TextBlob та VADER. Вони одні з найбільш ефективних бібліотек для Python, щоб опрацьовувати природну мову. Обидві бібліотеки є безкоштовними та у відкритому доступі.

TextBlob — це бібліотека Python для опрацювання текстових даних. Вона надає простий API для занурення в поширені завдання обробки природної мови (NLP), такі як додавання тегів до частин мови, виділення іменників, аналіз настроїв, класифікація тощо. Бібліотека активно застосовує Natural Language ToolKit (NLTK) для вирішення своїх задач. NLTK є потужним інструментом, який забезпечує зручний доступ до різних лексичних ресурсів і дає змогу виконувати завдання, пов'язані з категоризацією, класифікацією та іншими аспектами обробки тексту. TextBlob, у свою чергу, є простою у використанні бібліотекою, що дозволяє проводити глибокий аналіз і виконувати різноманітні операції з текстовою інформацією [42].

Бібліотеку TextBlob створив Стівен Лорія, який є розробником та ентузіастом Python. Він опублікував першу версію TextBlob в 2013 році, покращуючи її для того, щоб домогтися простоти у використанні користувачами бібліотеки для обробки природної мови. Метою Стівена було надати легкий у сприйнятті інтерфейс, щоб можна було виконувати базові

NLP-завдання, такі як аналіз тональності тексту, токенизація, розмітка частин мови тощо [40].

Стівен Лорія створив TextBlob на основі вже існуючих на той час бібліотек, таких як NLTK та Pattern, поєднавши їхні найкращі можливості. Розробник спроектував бібліотеку так що вона є інтуїтивно зрозумілою та придатною навіть для тих, хто є новачком у NLP. До того ж, автор створив детальну документацію, що допомагає початківцям швидко освоїти основи роботи з TextBlob.

Особливістю роботи з TextBlob є те, що вона повертає полярність та суб'єктивність речення. Полярність лежить між $[-1,1]$, де -1 позначає негативний настрій тексту, а 1 визначає позитивний настрій. TextBlob має семантичні мітки, які допомагають у тональному аналізі. Наприклад — смайлики, знак оклику, емодзі тощо. Суб'єктивність бібліотека визначає між $[0,1]$. Суб'єктивність визначає співвідношення між особистими судженнями та фактами, представленими в тексті. Якщо суб'єктивність висока, це свідчить про переважання особистої точки зору автора над об'єктивною інформацією. У простих випадках бібліотека TextBlob демонструє відмінні результати у визначенні таких характеристик тексту.

Для того, щоб почати працювати із бібліотекою TextBlob її потрібно інсталиувати. Процес встановлення бібліотеки відбувається через командний рядок із введенням prompt (Рис. 3.2.1) [29]:

```
pip install textblob
```

Рис.3.2.1. Інсталяція бібліотеки TextBlob

Наступним кроком є застосування бібліотеки TextBlob за допомогою мови програмування Python. Створення коду відбувається у кілька етапів.

Залучення аналізатора відбувається через імпорт класу TextBlob із бібліотеки textblob (Рис. 3.2.2).

```
from textblob import TextBlob
```

Рис 3.2.2. Імпорт бібліотеки

Введення тексту для аналізу від користувача забезпечує функція `input()`. Вона дозволяє отримати введення тексту через консоль. Користувач вводить текст (наприклад, коментар з Instagram), надалі цей текст зберігається у змінній `sentence`. Рядок («Коментар:») слугує підказкою для користувача, яка пояснює, що потрібно ввести текст коментаря (Рис 3.2.3.).

```
sentence = input("Коментар: ")
```

Рис 3.2.3. Стрічка вводу тексту від користувача

Далі прописується блок створення екземплярів `TextBlob` для аналізу тексту. Вводиться змінна `analysisPol`, яка зберігатиме результат функції, яка обчислює полярність тексту `TextBlob(sentence).polarity`. Наступна змінна `analysisSub`, яка зберігатиме результат функції, яка обчислює суб'єктивність тексту `TextBlob(sentence).subjectivity`. Функція `TextBlob(sentence).words` розбиває текст на окремі слова та повертає їх у вигляді списку (Рис 3.2.4);

```
analysisPol = TextBlob(sentence).polarity  
analysisSub = TextBlob(sentence).subjectivity  
words = TextBlob(sentence).words
```

Рис 3.2.4. Екземпляри TextBlob об'єктів

Наступним кроком створюємо ініціалізацію списків збереження слів. Порожній список, у який зберігатимуться слова з позитивною полярністю разом із їх значеннями полярності, введемо як `positive_words`, а інший порожній список, у який навпаки зберігатимуться слова з негативною полярністю разом із їх значеннями полярності, введемо як `negative_words` (Рис. 3.2.5).

```
positive_words = []  
negative_words = []
```

Рис. 3.2.5. Введення списків для збереження слів

Блок циклу для кожного слова призначений для аналізу полярності слів, тобто визначення слів позитивними чи негативними, або ж нейтральними. Надалі ці слова зберігатимуться у списках описаних вище. Цикл `for` проходить через кожне слово в списку і виконує певні дії для кожного з них. Для кожного

слова, яке ми отримуємо в циклі, TextBlob(word) створює об'єкт для аналізу конкретного слова, у свою чергу word_analysis.polarity дає значення полярності слова (Рис. 3.2.6).

```
for word in words:  
    word_analysis = TextBlob(word)  
    word_polarity = word_analysis.polarity
```

Рис 3.2.6. Цикл опрацювання для кожного слова

Для перевірки, аналізу полярності та додавання результатів до списків, потрібно встановити параметри позитивної та негативної полярності слів. Якщо полярність слова більше нуля, то це означає, що слово має позитивну полярність, і додаємо його до списку positive_words. Якщо полярність слова менше нуля, то це означає, що слово має негативну полярність, і його додаємо до списку negative_words. Для того, щоб вивести кортеж слів прописуємо positive_words.append((word, word_polarity)) та за подібним принципом прописуємо параметри кортежу для негативних слів (Рис. 3.2.7).

```
if word_polarity > 0:  
    positive_words.append((word, word_polarity))  
elif word_polarity < 0:  
    negative_words.append((word, word_polarity))
```

Рис 3.2.7. Аналіз полярності слів та додавання до списків

Функція print() виводить результат на екран (в консоль), у нашому випадку ця функція буде виводити результат-оцінку сентимент-аналізу тексту та об'єктивність/суб'єктивність тексту. Якщо полярність проаналізованого тексту більше нуля, то текст має позитивну тональність, а якщо полярність проаналізованого тексту менше нуля, то текст має негативну тональність. У випадку, коли аналіз тексту дорівнює нулю, то цей текст можна вважати нейтральним (Рис. 3.2.8).

```

print(f"Полярність тексту: {analysisPol}")
if analysisPol > 0:
    print("Позитивна")
elif analysisPol < 0:
    print("Негативна")
else:
    print("Нейтральна")

print(f"Суб'єктивність/об'єктивність тексту: {analysisSub}")
if analysisSub >= 0.1:
    print("Суб'єктивний")
else:
    print("Об'єктивний")

```

Рис 3.2.8. Функція, яка виводить оцінку-результат

Заключним етапом є виведення позитивно чи негативно забарвлених слів у вигляді списку. Умовою форматування списку є те, що після слова мають бути дві крапки «:» та оцінка до кожного слова. Якщо слова з негативною полярністю є у коментарі, то вони будуть збережені у списку `negative_words`, але якщо слів не знайдено, то функція виведе у консоль слово «немає». Аналогічно й для позитивних слів (Рис. 3.2.9).

```

print("\nСлова з позитивним забарвленням:")
if positive_words:
    for word, polarity in positive_words:
        print(f"{word}: {polarity}")
else:
    print("Немає")

print("\nСлова з негативним забарвленням:")
if negative_words:
    for word, polarity in negative_words:
        print(f"{word}: {polarity}")
else:
    print("Немає")

```

Рис 3.2.9. Функція, яка виводить слова з позитивним/негативним забарвленням та їх оцінку полярності

Кінцевим кроком є тестування коду. Перевірка коду відбувалася довільними реченнями з емоційно-забарвленою лексикою. Наприклад

речення, «Фільм був приголомшливим і захоплюючим» з англійським еквівалентом «The movie was amazing and exciting» має такий результат (Рис. 3.2.10):

```
Коментар: The movie was amazing and exciting
Полярність тексту: 0.45000000000000007
Позитивна
Суб'єктивність/об'єктивність тексту: 0.8500000000000001
Суб'єктивний
```

Рис. 3.2.10. Результат сентимент-аналізу позитивного речення за допомогою бібліотеки TextBlob

Отже маємо результат, що речення «The movie was amazing and exciting» є позитивно забарвленим, адже його оцінка в межах проміжку від 0 до 1. Таке речення має суб'єктивний зміст, адже воно виражає особисте ставлення мовця до певного об'єкту, тобто оцінка знаходиться у проміжку від 0 до 1.

У консолі було виведено результат оцінки тональності слів у реченні. Слів із негативним забарвленням не було виявлено у бібліотеці TextBlob, тому функція вивела «Немає». Виявлено два слова із позитивним забарвленням, такі як: amazing та exciting результат (Рис. 3.2.11).

```
Слова з позитивним забарвленням:
amazing: 0.60000000000000001
exciting: 0.3

Слова з негативним забарвленням:
Немає
```

Рис. 3.2.11. Результат сентимент-аналізу позитивних слів за допомогою бібліотеки TextBlob

Позитивну полярність програмний код може виявляти, тому варто протестувати речення із негативним настроєм. Наприклад речення, «Фільм був жахливим і нудним» з англійським еквівалентом «The movie was terrible and boring» має такий результат (Рис. 3.2.12):

```
Коментар: The movie was terrible and boring
Полярність тексту: -1.0
Негативна
Суб'єктивність/об'єктивність тексту: 1.0
Суб'єктивний
```

Рис. 3.2.12. Результат сентимент-аналізу негативного речення за допомогою бібліотеки TextBlob

Отже маємо результат, що речення «The movie was terrible and boring» є негативно забарвленим, адже його оцінка - 1. Таке речення має суб'єктивний зміст, адже воно так само виражає особисте враження щодо певного об'єкту, тобто оцінка знаходиться в межах від 0 до 1.

Далі функція вивела у консолі результат оцінки тональності слів у реченні. Слів із позитивним забарвленням не було виявлено у бібліотеці TextBlob, тому було виведено «Немає». Більше того, було виявлено два слова із негативним забарвленням, такі як: terrible та boring результат (Рис. 3.2.13).

```
Слова з позитивним забарвленням:
Немає

Слова з негативним забарвленням:
terrible: -1.0
boring: -1.0
```

Рис. 3.2.13. Результат сентимент-аналізу негативних слів за допомогою бібліотеки TextBlob

Отже, код створений та протестований. Функції повертають очікувані результати, але надалі потрібно провести дослідження, щоб зрозуміти те, якою є похибка правильного визначення сентимент-аналізу тексту з бібліотекою TextBlob. Надалі можна переходити до написання подібного за структурою програмного коду, але для бібліотеки VADER, з якою будемо проводити порівняння, щодо певних параметрів та точності виведених результатів.

Бібліотека VADER (Valence Aware Dictionary and sEntiment Reasoner) була розроблена у 2014 році Еріком Гілбертом, американським фахівцем із

інформатики та доцентом Школи інформації Мічиганського університету імені Джона Дербі Еванса, та Клейтоном Хатто, головним науковим співробітником комп'ютерної інженерії за Інституту технічних досліджень Штату Джорджії. VADER є ефективним інструментом для аналізу тональності. Метою розробників бібліотеки було забезпечити точне визначення настроїв у текстах соціальних медіа, які часто є короткими за обсягом, неформальними та насиченими сленгом, аббревіатурами та емотиконами [33].

Особливістю бібліотеки VADER є те, що вона спеціалізується на аналізі контенту, який публікують користувачі соціальних мереж, але цей інструмент можна застосовувати й в інших сферах досліджень, таких як аналіз відгуків клієнтів або заголовків новин. Завдяки відкритому вихідному коду (MIT License), вона стала поширеною серед дослідників та розробників. MIT License є однією із найбільш популярних ліцензій для програмного забезпечення з відкритим кодом. Завдяки цій ліцензії розробникам та користувачам надаються широкі права на використання, зміну та розповсюдження програмного забезпечення. До того ж, MIT License має ряд й інших переваг, як-от: простота, адже ліцензія має короткий за обсягом текст-пояснення, тому є зрозумілою для користувачів; сприятливість до модифікацій, тобто користувачі мають право змінювати код та поширювати його в оригінальному або зміненому вигляді [43].

Сентимент-аналіз із VADER проводиться за допомогою нормалізованої оцінки, яка показує загальний настрій тексту, а саме позитивний, нейтральний або негативний. Основними перевагами бібліотеки можна зазначити швидкість, простоту інтеграції та здатність працювати з різними нюансами в тексті. Бібліотека має функцію враховування впливу пунктуації на тональність тексту. Наприклад, до таких пунктуаційних маркерів належать – знаки оклику, які є підсилювачами емоцій. Наприклад, слово «чудово!!!» буде оцінено значно позитивнішим, ніж просто «чудово».

Ефективність VADER була доведена в численних дослідженнях. Бібліотеку застосовували для аналізу громадських настроїв під час політичних виборів, моніторингу сторінок брендів у соціальних мережах та оцінки тональності у відгуках на платформах, таких як Amazon. Бібліотека VADER здатна аналізувати як формальні, так і неформальні тексти, тому можна вважати її універсальним інструментом для лінгвістичних досліджень. Оцінки, які виводить VADER є не тільки обнадійливими – вони є справді на рівні високої точності та правильності результату. В експериментальних дослідженнях науковців, VADER показує себе так само добре (або й в більшості випадків краще ніж) інші високо оцінені інструменти для тонального аналізу тексту [33].

Для початку роботи з VADER необхідно виконати інсталяцію бібліотеки. Цей процес здійснюється за допомогою командного рядка шляхом введення відповідної команди (Рис. 3.2.14).

```
pip install vaderSentiment
```

Рис.3.2.14. Інсталяція бібліотеки VADER

Наступний етап включає використання бібліотеки vaderSentiment через мову програмування Python. Реалізація коду здійснюється наступним чином: для активації аналізатора необхідно імпортувати клас SentimentIntensityAnalyzer, щоб залучити інструмент для обробки текстових даних (Рис. 3.2.15).

```
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
```

Рис.3.2.15. Імпорт бібліотеки VADER

Потрібно створити об'єкт, який буде використовуватись для тонального аналізу введеного тексту або окремих слів (Рис. 3.2.16).

```
analyzer = SentimentIntensityAnalyzer()
```

Рис.3.2.16. Створення об'єкту analyzer

Для введення тексту для аналізу будемо використовувати функцію `input()`, яка дозволяє отримувати дані безпосередньо із консолі. Текст, введений користувачем, стане входними даними для аналізу настрою та зберігатиметься у змінній `sentence`. У цьому ж рядку варто ввести підказку як («Коментар:»), щоб користувачу пояснювалося, який саме текст потрібно для аналізу (Рис. 3.2.17).

```
sentence = input("Коментар: ")
```

Рис 3.2.17. Стрічка вводу текстового коментаря від користувача

Наступним кроком потрібно прописати блок коду із екземплярами для аналізатора VADER (Рис. 3.2.18).

```
sentiment = analyzer.polarity_scores(sentence)
vader_polarity = sentiment['compound']
```

Рис 3.2.18. Екземпляри VADER об'єктів

Програмуючи наступну стрічку коду, створюємо ініціалізацію двох порожніх списків для збереження слів із різними полярностями. Перший список, у який будуть додаватися слова із позитивною полярністю та їх значення, позначимо як `positive_words`. Інший список, який міститиме слова із негативною полярністю разом з їх оцінками, позначимо як `negative_words`. Для розбивки слів у реченні варто використати метод `split` (Рис. 3.2.19).

```
words = sentence.split()
positive_words = []
negative_words = []
```

Рис. 3.2.19. Введення списків для збереження емоційно-забарвлених слів

Цикл, який обробляє кожне слово, призначений для оцінки полярності слів, тобто для визначення, чи є слово позитивним, негативним чи нейтральним. Після аналізу, ці слова з відповідними значеннями полярності зберігаються в раніше створених списках. Цикл `for` перебирає всі слова в списку `words`. Для кожного слова, яке обробляється, за допомогою

`analyzer.polarity_scores(word)` створюється змінна, що дозволяє здійснити аналіз цього слова `word_sentiment` (Рис. 3.2.20).

```
for word in words:
    word_sentiment = analyzer.polarity_scores(word)
```

Рис. 3.2.20. Цикл опрацювання для кожного слова у VADER

Надалі, потрібно встановити параметри оцінки полярності слів. Якщо загальна оцінка настрою тексту чи слова «`compound > 0`», то слово вважається позитивним. Слово разом із його позитивною оцінкою додається до списку `positive_words`, якщо ж оцінка настрою «`compound < 0`», то слово вважається негативним та прямує до списку `negative_words`. Слово буде вважається нейтральним, якщо «`compound < 0`», за таких умов такі слова ігноруються, адже вони не є релевантними для аналізу, аби їх виводити у окремий список (Рис. 3.2.21).

```
if word_sentiment['compound'] > 0:
    positive_words.append((word, word_sentiment['compound']))
elif word_sentiment['compound'] < 0:
    negative_words.append((word, word_sentiment['compound']))
```

Рис. 3.2.21. Введення списків для збереження емоційно-забарвлених слів та їх аналіз

Функція `print()` у Python використовується для виведення результатів на екран, зокрема в консоль. У контексті сентимент-аналізу, ця функція дозволяє показати на екрані оцінку тональності та об'єктивності або суб'єктивності аналізованого тексту. У нашому випадку, функція буде виводити полярність тексту, визначену на основі аналізу тексту кодом. Якщо значення полярності більше за нуль, це свідчить про те, що текст має позитивну тональність, тобто автор коментаря висловлює позитивні емоції або думки. Якщо полярність менше за нуль, отже, текст містить негативні емоції або думки. У випадку, коли значення полярності дорівнює нулю, можна зробити висновок, що текст є нейтральним, тобто він не містить яскраво виражених позитивних чи негативних емоцій. Варто ввести підказку-уточнення таку як «Полярність тексту за VADER», щоб серед виведених даних аналізу коментарів із різних

бібліотек-аналізаторів, можна було швидко ідентифікувати бібліотеку, за допомогою якої виконували аналіз (Рис. 3.2.22).

```
print(f"Полярність тексту за VADER: {vader_polarity}")
if vader_polarity > 0:
    print("Позитивна")
elif vader_polarity < 0:
    print("Негативна")
else:
    print("Нейтральна")
```

Рис 3.2.22. Функція виведення оцінки-результату за VADER

Аналіз тональності всього тексту не є достатнім джерелом для досліджень, тому постала потреба виокремити емоційно-забарвлені слова та вивести їх списки в консолі. Отже, потрібно створити блок коду, де має бути функція, яка спрямована на виведення списку слів із позитивною та негативною полярністю разом із їхніми оцінками. Якщо ж у тексті немає емоційно-забарвлених слів, то до списку не потрапляють слова без тональності, таким чином список вважається порожнім. Користувачу виводиться повідомлення «Немає» (Рис. 3.2.23).

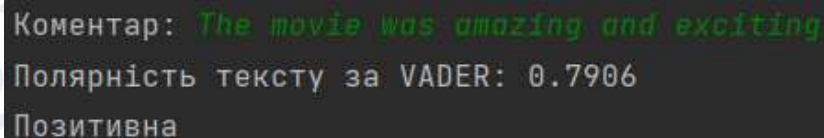
```
print("\nСлова з позитивним забарвленням (VADER):")
if positive_words:
    for word, score in positive_words:
        print(f"{word}: {score}")
else:
    print("Немає")

print("\nСлова з негативним забарвленням (VADER):")
if negative_words:
    for word, score in negative_words:
        print(f"{word}: {score}")
else:
    print("Немає")
```

Рис 3.2.23. Функція виведення слів з позитивним/негативним забарвленням та їх оцінку полярності за VADER

Заключним етапом у процесі роботи з кодом-аналізатором є перевірка виводу результату. Для тестування використовувалися різні речення, що

містять емоційно забарвлену лексику. Наприклад, речення «Фільм був приголомшливим і захоплюючим», англійською мовою перекладається як «The movie was amazing and exciting», було використане для оцінки результатів аналізу. Отримані результати були правильними (Рис. 3.2.24).

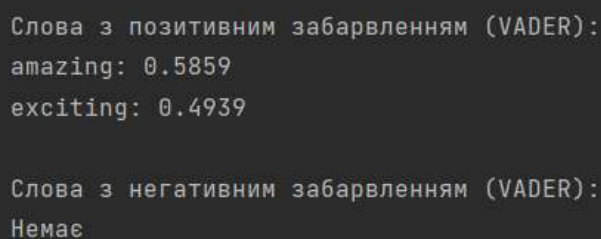


```
Коментар: The movie was amazing and exciting
Полярність тексту за VADER: 0.7906
Позитивна
```

Рис. 3.2.24. Результат сентимент-аналізу позитивного речення за допомогою бібліотеки VADER

Аналізатор опрацював речення із позитивним забарвленням та вивів полярність 0.7906, яка є більшою нуля, отже, вважається позитивною.

У консоль було виведено результати аналізу тональності слів у реченні. Жодного слова з негативним забарвленням із бібліотеки VADER не було ідентифіковано в реченні, тому функція повернула значення «Немає». У той же час було визначено два слова з позитивною полярністю: amazing та exciting, які підтвердили емоційно-позитивний характер аналізованого тексту (Рис. 3.2.25).



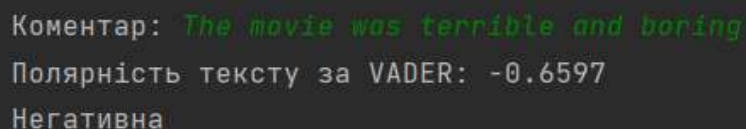
```
Слова з позитивним забарвленням (VADER):
amazing: 0.5859
exciting: 0.4939

Слова з негативним забарвленням (VADER):
Немає
```

Рис. 3.2.25. Результат сентимент-аналізу позитивних слів за допомогою бібліотеки VADER

Програмний код здатний розпізнавати позитивну полярність у текстах, однак для повноцінного тестування слід перевірити його здатність працювати з реченнями, що містять негативний емоційний контекст. Наприклад, речення «Фільм був жахливим і нудним», англійський еквівалент якого — «The movie was terrible and boring», є вдалим прикладом для такої перевірки. Аналіз цього

речення дозволяє оцінити, чи правильно алгоритм визначає негативну тональність, відображаючи результат відповідно до очікувань (Рис. 3.2.26).

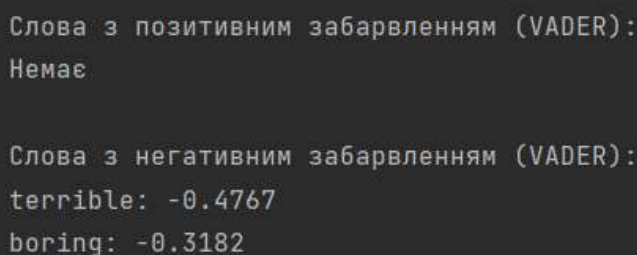


```
Коментар: The movie was terrible and boring
Полярність тексту за VADER: -0.6597
Негативна
```

Рис. 3.2.26. Результат сентимент-аналізу негативного речення за допомогою бібліотеки VADER

У результаті аналізу було встановлено, що речення «The movie was terrible and boring» має яскраво виражену негативну тональність, оскільки його оцінка становить -0.6597. Це свідчить про те, що текст відображає негативне емоційне забарвлення, спрямоване на конкретний об'єкт.

Функція обробки тексту вивела у консоль детальний результат аналізу тональності слів. У цьому випадку слова з позитивною полярністю не були знайдені, тому програма повернула значення «Немає». Водночас у реченні були ідентифіковані два слова із негативним емоційним забарвленням — "terrible" та "boring". Їхні показники полярності відображають суттєвий негативний вплив на загальний настрій тексту, що підтверджується результатами аналізу (Рис. 3.2.27).



```
Слова з позитивним забарвленням (VADER):
Немає

Слова з негативним забарвленням (VADER):
terrible: -0.4767
boring: -0.3182
```

Рис. 3.2.25. Результат сентимент-аналізу негативних слів за допомогою бібліотеки VADER

Таким чином, програмний код успішно створений та пройшов початкове тестування. Його функції демонструють очікувані результати, проте варто зосередитися на оцінці рівня похибки при визначенні тональності тексту за допомогою бібліотеки VADER.

3.3 Збір та підготовка даних для вибірки текстових коментарів

Досягнення основних цілей дослідження потребувало збору та підготовки текстових коментарів негативного характеру, які залишили користувачі соціальної мережі Instagram. До уваги були взяті коментарі англійською мовою для аналізу за допомогою TextBlob та VADER, оскільки обидві ці бібліотеки оптимізовані саме для роботи з текстами англійською мовою. Їхні словники та правила для визначення тональності базуються на особливостях англійської мови, таких як граматична структура, лексичний запас, фразеологізми та інші мовні нюанси. Використання текстів українською мовою з перекладом на англійську могло б викликати спотворення вихідного сенсу через неточність перекладу. Це, у свою чергу, вплине на точність визначення полярності текстів, що зробить результати аналізу ненадійними. Було вирішено, що, для отримання достовірних і точних результатів, варто зосередитися на роботі саме з коментарями, написаними англійською мовою, щоб максимально використати та дослідити можливості бібліотек TextBlob та VADER.

На меті є не лише дослідити аспект використання мовних маркерів негативного забарвлення в конфліктній комунікації, але й перевірити точність роботи бібліотек TextBlob та VADER у розпізнаванні емоційного забарвлення текстів. Негативні коментарі стали основним матеріалом для текстової вибірки, адже подібна комунікація є справжнім викликом для автоматичного аналізатора настроїв.

Зібрані дані було отримано з постів у Instagram, які мали високу активність користувачів у коментарях. Вибірка базувалася на контенті, що викликав резонанс серед аудиторії, включаючи новинні публікації, дописи про особисте життя знаменитостей та спортивні репортажі.

Збір даних проводився з платформи Instagram, яка є однією з найпопулярніших соціальних мереж для висловлення думок та коментарів у

світі. Джерелами даних було обрано три різні категорії постів. До них належали пости з офіційних сторінок:

- новинна публікація авторитетного видання CNN (Cable News Network) на тему надання дозволу Україні президентом США Джо Байденом використовувати потужну американську далекобійну зброю [28];
- пост південно-кореїської супер-зірки Хьюни Кім, яка оголосила про своє весілля та отримала багату хейту від фанатів [15];
- спортивні новини одного з найбільших каналів ESPN (Entertainment and Sports Programming Network) про те, що 16-ти річний золотий медаліст Квінсі Вілсон придбав собі авто марки Rolls Royce [30] та про те, що менеджери професійної хокейної команди «Boston Bruins» звільнили головного тернера Джима Монтгомері [31].

Було обрано наведені вище пости, через передбачену здатність провокувати активне обговорення та хейтерські коментарі черед користувачів, що підходить для тестування алгоритмів аналізу настроїв.

Методи збору даних включали ручне копіювання коментарів із Instagram постів. Під час збору дані були фільтрувалися, щоб залишити лише ті коментарі, які містили виражену конфліктну та негативну комунікацію. Цей крок був необхідний для зосередження уваги на аналізі негативних емоцій. Крім того, дані були попередньо очищені від емотиконів та спеціальних символів, які не відображаються в консолі PyCharm під час аналізу та не можуть опрацьовуватися. Після очищення коментарі були структуровані в таблицю, яка включала такі параметри, як джерело (офіційний Instagram акаунт), текст коментаря, загальна кількість слів, кількість слів із негативним забарвленням, нікнейм автора, статя, тема публікації та результат аналізу настроїв.

До вибірки було відібрано 10 коментарів під новинним постом офіційної Instagram сторінки CNN, які мають негативне забарвлення. Ці коментарі

характеризуються критикою та засудженням теми обговорення, чиновників, які приймали рішення. Новина про надання дозволу Україні президентом США застосовувати потужну американську зброю великої дальності сколихнула емоції підписників спільноти CNN. У коментарях всі охочі відреагували на публікацію. Стаття коментаторів – чоловіча, адже при мануальному огляді всіх коментарів під постом не було виявлено негативних коментарів від жінок. Зібрані дані систематизовано у Таблиці 3.2, що наведена нижче.

Таблиця 3.2 – Вибірка негативних коментарів (джерело новинний канал CNN)

№	Текст коментаря	Нікнейм автора	Стать
1	RIP Ukraine!!!! Russia is about to let loose a gangster arsenal!!	spreah	Ч
2	Shame on the democrats!!!	hackenbusch_official	Ч
3	Absolutely idiotic decision it will not change anything but it might create the world War three	temo_matcharashvili	Ч
4	Disastrous	patrickalmeida_s	Ч
5	Trying to sabotage the country because they lost	mikemoneyrun	Ч
6	Stupid decision	javmolina1	Ч
7	USA president is a sick man	mufeed0100	Ч
8	Dangerous senile old shit.. couldn't care less for peoples lives	fionnbarr_86	Ч
9	This idiot gonna have us in a WAR!	nemesis711_	Ч
10	Fuck it start a war before leaving	richieforever	Ч

Наступною резонансною темою було обрано нішу шоу-бізнесу. Одним із найскандальніших постів стала публікація весільних фото південно-корейської супер-зірки Хьюни Кім. Переважна кількість коментарів була залишена користувачами жіночої статі, а негативних коментарів від чоловіків не було під постом.

Отже, до вибірки було обрано 10 хейтерських коментарів з вираженою конфліктною комунікацією. Зібрані дані були систематизовані для подальшого аналізу. Усі коментарі структуровані до таблиці, що дозволяє чітко та прозоро представити отриману інформацію. Таблиця зібраних даних представлена нижче як Таблиця 3.3.

Таблиця 3.3 – Вибірка негативних коментарів (джерело особиста Instagram сторінка Хьюни Кім)

№	Текст коментаря	Нікнейм автора	Стать
1	I hate you	meryamibrahimi_	Ж
2	Disappointed in you	ellensmith2002_	Ж
3	What the fuck, man	hellh69ss	Ж
4	The worst wedding dress I've ever seen	sound_of_sumin.04	Ж
5	I hate her, but that flowers are SO PRETTY	your.luzz	Ж
6	Both of you can go to hell and burn forever	morningstarlight_	Ж
7	You deserve nothing but the worst.	that_gvrlie	Ж
8	Spreading hate to you both	_shar.miiiiiii_	Ж
9	That's what I call ruining your life.	kayn_azul	Ж
10	Shame on you	keylacutie	Ж

Для тематики наступної вибірки було надано перевагу ніші спортивних новин. Стало відкриттям те, що позитивних коментарів та слів підтримки переважна кількість у коментарях під спортивними новинами, тому було обрано не один, а два пости одного із найавторитетніших та найпопулярніших каналів спортивних новин – ESPN. Під публікаціями вдалося ідентифікувати 10 коментарів негативного забарвлення, які транслювали негативні думки та емоції коментаторів. Особливістю зібраних даних є те, що всі негативні коментарі були залишені чоловіками. Жінки, які коментували обрані пости, уникали негативних висловлювань та залишали позитивні коментарі, щоб підтримати улюблену команду чи спортсмена. Зібрані коментарі систематизовано у Таблиці 3.4, яку наведено нижче.

Таблиця 3.4 – Вибірка негативних коментарів (джерело новинний канал ESPN)

№	Текст коментаря	Нікнейм автора	Стать
1	His Barber Should Be Ashamed At That Hairline	pluto.skiiii	Ч
2	He looks like hes legally blind	champagnepop22	Ч
3	Already horrible with his money	c_abelendee24	Ч
4	and will be dead broke in less than 10 years	joemacnaughton	Ч
5	That girls skin complexion is disgusting	return_of_jaafar	Ч
6	We still suck regardless	christianpx648	Ч
7	What an absolutely terrible decision. Not his fault our players and goaltending suck	silaspalker	Ч
8	OK, that was another bad idea in a long list of bad ideas. Montgomery was not the problem.	goaliedad.13	Ч
9	Bad move	itsjohnawad	Ч
10	i hate the bruins for this	baby_jordo	Ч

Підготовлена вибірка в сумі налічує 30 коментарів із негативним емоційним забарвленням. Дані готові до опрацювання бібліотеками TextBlob та VADER. Надані аналізаторами результати сприятимуть створенню візуалізацій різних за параметрами експериментів, таких як діаграми та графіки. Вибірка коментарів, які залишені реальними користувачами, має забезпечити виявлення оцінки ефективності алгоритмів бібліотек TextBlob та VADER у роботі з текстами природної мови.

У підсумку процес підготовки даних дозволив створити вибірку, яка відповідає поставленим завданням майбутнього експериментального дослідження. Дана вибірка стане у нагоді, щоб стало можливим оцінити ефективність обраних бібліотек як інструментів для автоматичного сентимент-аналізу.

3.4 Експериментальне дослідження вибірки та оцінка результатів

Сучасний цифровий світ посприяв стрімкому розвитку соціальних мереж, які на сьогодні стали невід'ємною частиною життя мільйонів людей. Подібні платформи для обміну думками та емоціями несуть і позитивний, і негативний характер, адже часто слугують середовищем для конфліктної комунікації, яка проявляється у вигляді агресивних, образливих та токсичних коментарів. Зокрема, у соціальній мережі Instagram, де комунікація відбувається у вільній формі, негативні висловлювання, або хейт стають поширеним явищем, що значно впливають на атмосферу обговорень під публікаціями та психологічне благополуччя інших користувачів.

Конфліктна комунікація в мережі має свої відмінності від авторської інтелектуальної прози-публікації, адже негативні коментарі часто виникають спонтанно, підкріплюються анонімністю користувачів та можуть набувати масової підтримки серед токсичних користувачів. Дослідження цих аспектів є важливими для розуміння природи негативної взаємодії в соціальних медіа.

Особливу увагу в зібраній вибірці коментарів буде покладено на тональний аналіз висловлювань, які містять мовні маркери конфліктної комунікації, використовуючи автоматизовані інструменти TextBlob і VADER.

Важливим є проведення експериментального дослідження задля визначення характерних рис конфліктних висловлювань, таких як частота використання негативних слів, їхнє смислове навантаження, а також вплив гендеру на рівень негативу в конфліктній комунікації. Додатково, зробити аналіз ефективності двох найефективніших бібліотек сентимент-аналізу тексту та визначити ключові слова-маркери, які найчастіше з'являються у конфліктних висловлюваннях.

Проведене дослідження матиме на меті не лише глибше зрозуміти «природу» негативної комунікації в соціальних мережах, але й сприятиме розробці більш ефективних підходів до створення сентимент-аналізаторів.

Соціальні мережі надають користувачам свободу висловлення, але, в той же час, встановлюють параметри, які обмежують коментар користувача під постом до 2000 тисяч друкованих символів. Подібні ліміти сприяють культурі коротких та спонтанних текстових висловлювань. Таким чином, можна висунути гіпотезу, що коментарі з негативним забарвленням не будуть містити багато слів. Отже, будуть короткими за обсягом.

Зібрана вибірка негативних коментарів налічує 30 коментарів з 3-ох різних груп: «Новини» (Додаток А, Б), «Спортивні новини» (Додаток В, Г) та «Шоу-бізнес» (Додаток І, Д). Рубрика «Новини» об'єднує 10 коментарів, у яких всього 68 слів, «Спортивні новини» на 10 коментарів мають обсяг у 78 слів, а от група «Шоу-бізнес» містить найменше слів на 10 коментарів – 57. Для того, аби зрозуміти та передбачити обсяг слів негативних коментарів потрібно вирахувати середнє значення кількості слів (Рис. 3.4.1).

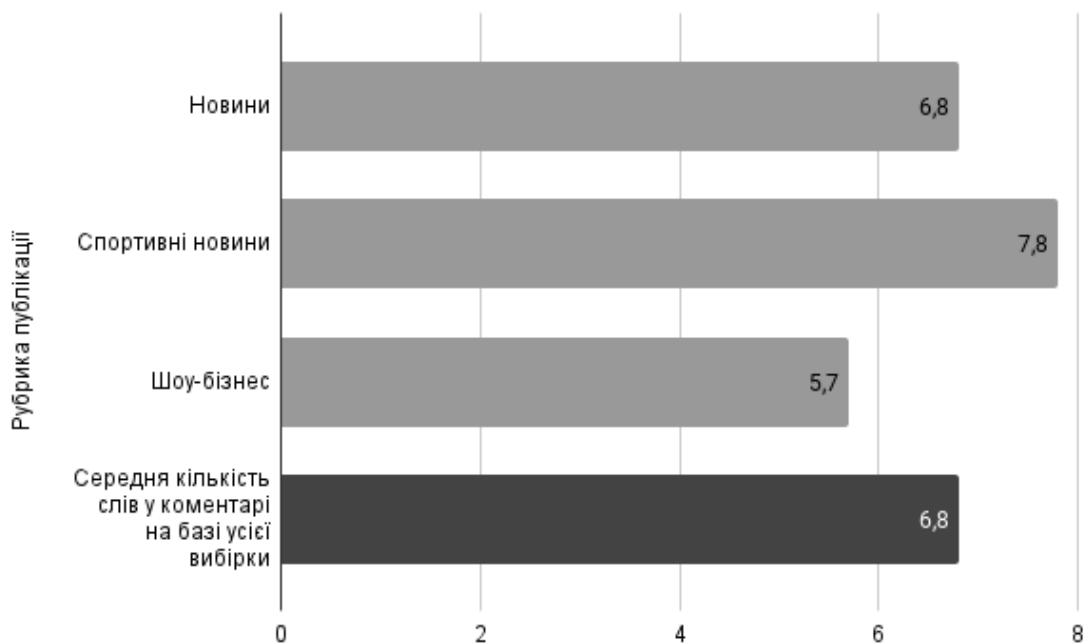


Рисунок 3.4.1. Середня кількість слів у коментарі з емоційним негативним забарвленням

Отже, аналізуючи діаграму наведену вище, спостерігаємо підтвердження гіпотези, що висловлювання у соціальних мережах зазвичай не

містять багато слів, адже середня кількість слів у негативному коментарі – 6,8 слів.

Найменше слів у коментарях було залишено в рубриці «Шоу-бізнес» – 5,7 слів, а в той же час «Спортивні новини» лідирують за середньою кількістю слів у коментарі, адже налічують 7,8 слів. Маємо припущення, що «Спортивні новини» зазвичай викликають високий резонанс у користувачів, зокрема, через динамічність тем, адже спортивні події часто супроводжуються миттєвими емоційними реакціями на поточні події, результати матчів, які стимулюють дискусії про майбутні перемоги чи невдачі спортсменів.

Отримана інформація про середній словесний обсяг негативного висловлювання в Instagram надає можливість проаналізувати кількість слів на коментар. Дослідження кількісного використання слів дозволяє спрогнозувати мовну модель, яку автори коментарів використовують для висловлення невдоволення чи критики в цифровому середовищі.

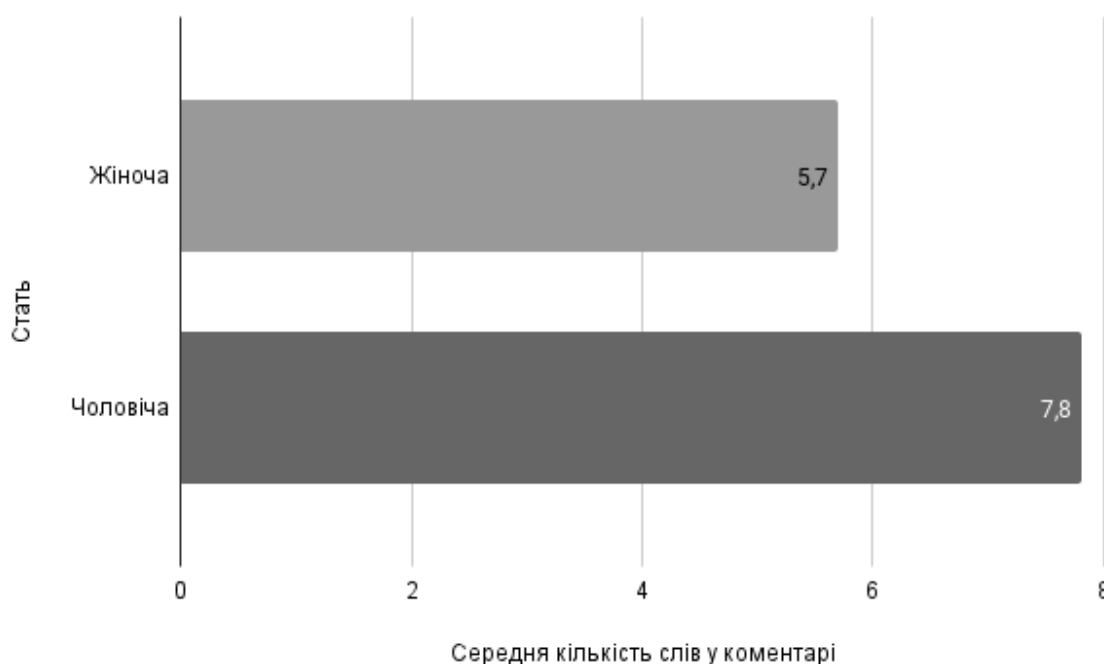


Рис.3.4.2. Середня кількість слів у коментарі з емоційним негативним забарвленням за гендером коментатора

Наприклад, врахування гендерних особливостей, існує думка, що жінки більш емоційні та більш схильні до довгих бесід, аніж чоловіки. З огляду на дану думку, можна висунути гіпотезу, що жінки використовуватимуть більше слів для вираження негативних емоцій під публікацією, ніж чоловіки.

Даним порівняльним аналізом негативних коментарів чоловіків на тему «Спортивні новини» та коментарів жінок з рубрики «Шоубізнес» (Рис.3.4.2) доведено, що подібне припущення може бути лише аргументом побудованим на стереотипних уявленнях про гендерні відмінності.

Аналіз коментарів бібліотеками TextBlob і VADER надає можливість дослідникам отримувати дані про позитивні та негативні слова в тексті. Негативні коментарі в соціальних мережах, можуть містити як позитивну, так і негативну лексику, адже саме це відображає складність природного мовлення. Навіть у негативному контексті людина може використовувати позитивні слова, але негативних слів буде переважна кількість, що підтверджує сентимент-аналіз негативних коментарів з Instagram (Рис. 3.4.3 та Рис. 3.4.4).

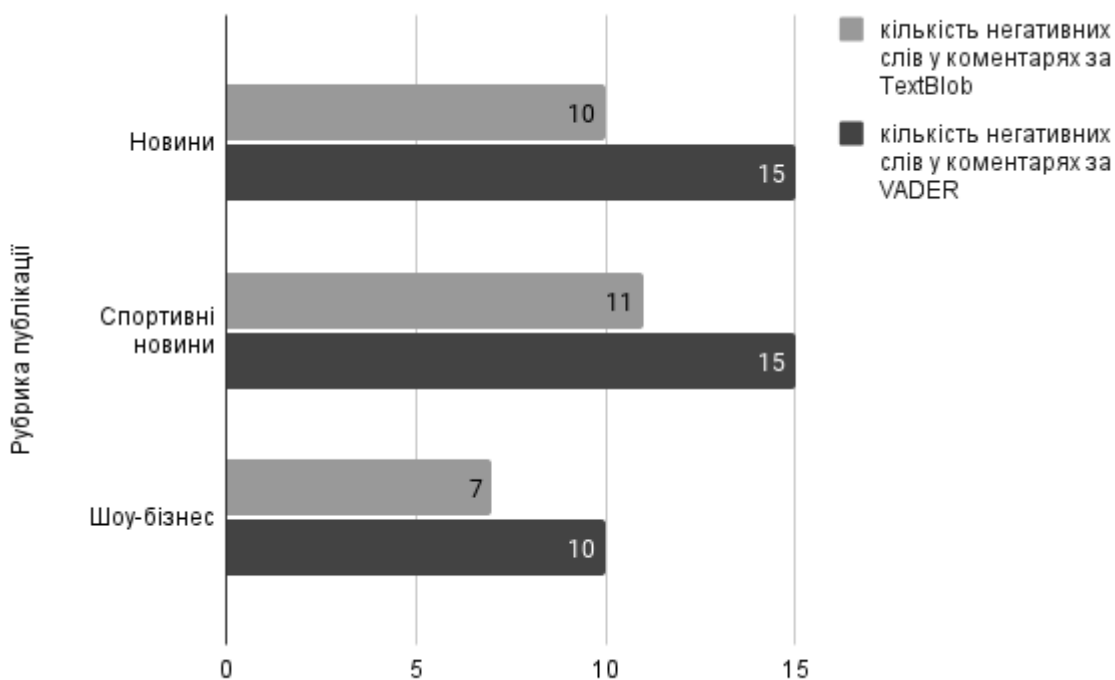


Рис. 3.4. 3. Кількість негативних слів у кожній рубриці публікацій

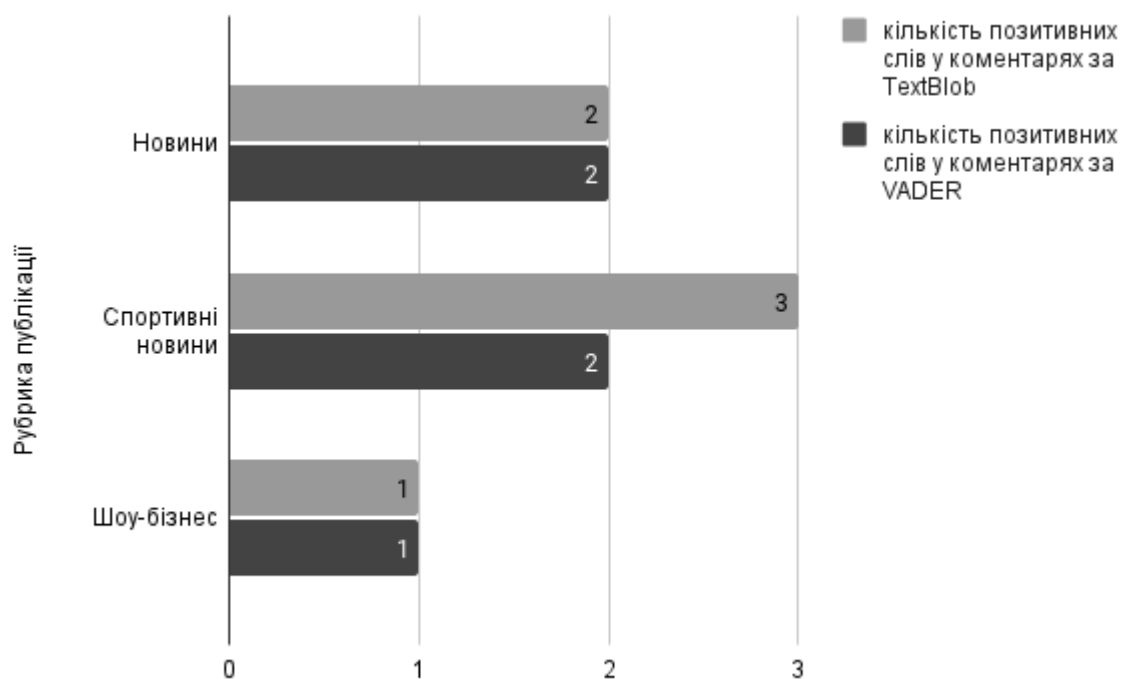


Рис. 3.4. 4. Кількість позитивних слів у кожній рубриці публікацій

Варто зазначити, що позитивні слова становлять приблизно 4,44% від усіх 135 досліджених слів. Негативні слова досягають приблизно 20,74% за TextBlob та 29,63% за VADER серед обсягу всієї вибірки конфліктної комунікації. Виявлена диспропорція між частками позитивних й негативних слів свідчить про перевагу негативного тону в коментарях.

Значна різниця у аналізі негативних слів між TextBlob (20,74%) і VADER (29,63%) підкреслює важливість використання різних підходів та інструментів для тонального аналізу. Виокремлюємо те, що VADER краще адаптований до сленгу та неформальних висловлювань, тому, ймовірно, може точніше розпізнавати негативні емоції.

Зважаючи на те, що негативних слів переважна кількість у конфліктній комунікації, можемо припустити, що не менш, ніж одне слово з негативною оцінкою, має бути присутнє у негативному коментарі. Дослідження цього аспекту дозволяє зрозуміти, як люди висловлюють негативні почуття та скільки негативно забарвлених слів потрібно мовцю, щоб виразити негативну оцінку (Рис.3.4.5).

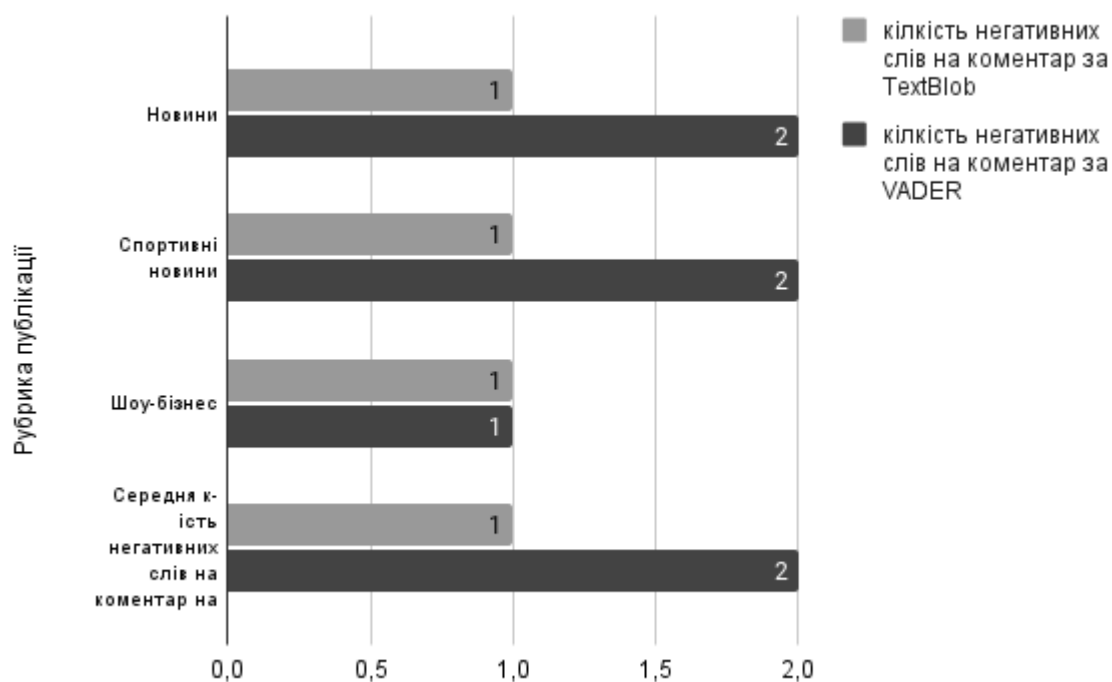


Рис. 3. 4. 5. Кількість негативних слів на коментар у кожній рубриці публікацій

Знову спостерігається різниця результатів, що надають бібліотеки TextBlob і VADER, але наявна діаграма свідчить про те, що мінімум одне слово негативного забарвлення має бути присутнє у коментарі негативного характеру.

Негативні слова є ключовими маркерами ідентифікації негативних коментарів, в контексті соціальних мереж, для інструментів автоматизованого сентимент-аналізу. Ці слова часто виражають емоції, такі як розчарування, злість, невдоволення чи образу. Вони включають прямі прояви агресії (лайки, зневажливі фрази) або непрямі форми критики, наприклад, сарказм чи іронію.

Аналіз слів-маркерів допомагає автоматизованим системам навчатися розпізнавати негатив у тексті, саме тому, стає можливим не лише класифікувати коментарі, але й передбачати загальний тон дискусій чи характер взаємодії у спільноті.

Сентимент-аналіз коментарів з допомогою бібліотеки TextBlob виділив 28 слів негативного емоційного забарвлення серед всієї вибірки коментарів. У той же час, VADER розпізнав 40 слів-маркерів негативної комунікації (Таблиці 3.4.1, 3.4.2, 3.4.3).

Таблиця 3.4.1. – Слова-маркери негативної комунікації в рубриці «Новини», які були розпізнані бібліотеками TextBlob та VADER

Рубрика публікації	Слово-маркер	Полярність за TextBlob	Слово-маркер	Полярність за VADER
Новини	Dangerous	-0.6	Dangerous	-0.4767
	Disastrous	-0.7	Disastrous	-0.5994
	Fuck	-0.4	Fuck	-0.5423
	idiot	-0.8	idiot	-0.5106
	idiotic	-0.6666666666666666	idiotic	-0.5574
	less	-0.16666666666666666	loose	-0.3182
	loose	-0.07692307692307693	lost	-0.3182
	shit	-0.2	sabotage	-0.5267
	Sick	-0.7142857142857143	Shame	-0.4767
	Stupid	-0.7999999999999999	shit	-0.5574
			Sick	-0.5106
			Stupid	-0.5267
			War	-0.5994
			war	-0.5994
			WAR!	-0.636

Аналізуючи таблицю наведену вище, можна винести висновок, що найбільш негативно забарвленими словами в коментарях під публікацією рубрики «Новини» є: «idiot» (-0.8) та «WAR!» (-0.636). Бібліотека VADER призначає слову-маркеру «war» полярність у -0.5994, але якщо додано до слова пунктуаційне підсилення у вигляді знаку оклику, то оцінка негативної полярності стає вищою.

У групі вибірки «Спортивні новини» було виявлено не менше слів-маркерів, аніж у групі «Новини». Спостерігається тенденція до вживання повторів негативних слів в залишених коментарях.

Таблиця 3.4.2. – Слова-маркери негативної комунікації в рубриці «Спортивні новини», які були розпізнані бібліотеками TextBlob та VADER

Рубрика публікації	Слово-маркер	Полярність за TextBlob	Слово-маркер	Полярність за VADER
Спортивні новини	bad	-0.6999999999999998	Ashamed	-0.4767
	bad	-0.6999999999999998	bad	-0.5423
	Bad	-0.6999999999999998	bad	-0.5423
	blind	-0.5	bad	-0.5423
	dead	-0.2	broke	-0.4215
	disgusting	-1.0	blind	-0.4019
	hate	-0.8	dead	-0.6486
	horrible	-1.0	disgusting	-0.5267
	less	-0.16666666666666666	fault	-0.4019
	long	-0.05	hate	-0.5719
	terrible	-1.0	horrible	-0.5423
			problem	-0.4019
			suck	-0.4404
			suck	-0.4404
			terrible	-0.4767

Проаналізувавши таблицю, наведені вище дані дозволяють зробити висновок, що найбільш негативно забарвленими словами за TextBlob в коментарях до публікації рубрики «Спортивні новини» є: «disgusting» (-1.0), «horrible» (-1.0), «terrible» (-1.0), які оцінені максимально негативно за параметрами. Бібліотека VADER не оцінила наведені маркери максимально негативно, найбільшу негативну оцінку (-0.6486) присвоєно слову-маркеру «dead».

За попереднім аналізом всієї вибірки негативних коментарів стало відомо, що найменше негативних слів було визначено бібліотеками TextBlob та VADER в рубриці «Шоу-бізнес».

Таблиця 3.4.3. – Слова-маркери негативної комунікації в рубриці «Шоу-бізнес», які були розпізнані бібліотеками TextBlob та VADER

Рубрика публікації	Слово-маркер	Полярність за TextBlob	Слово-маркер	Полярність за VADER
Шоу-бізнес	Disappointed	-0.75	Disappointed	-0.4767
	fuck	-0.4	fuck	-0.5423
	hate	-0.8	hate	-0.5719
	hate	-0.8	hate	-0.5719
	hate	-0.8	hate	-0.5719
	worst	-1.0	hell	-0.6808
	worst	-1.0	ruining	-0.25
			Shame	-0.4767
			worst	-0.6249
			worst	-0.6249

Опрацювання даних, що наведені у таблиці вище, сприяло виведенню висновків про те, що найбільш негативними за полярністю є слова «worst» (-1.0) та «hell» (-0.6808). З огляду на унікальність негативних слів у коментарях – дана вибірка «Шоу-бізнес» є лідером серед частих повторів слів-маркерів.

Отже, сентимент-аналіз всієї вибірки конфліктної комунікації в Instagram спонукає висунути гіпотезу, що слова-маркери негативної лексики можуть повторюватися у коментарях різних осіб, тому вважаємо доцільним визначити слова, які повторюються користувачами найчастіше для вираження негативних емоцій (Рис. 3.4.6 та Рис. 3.4.7).

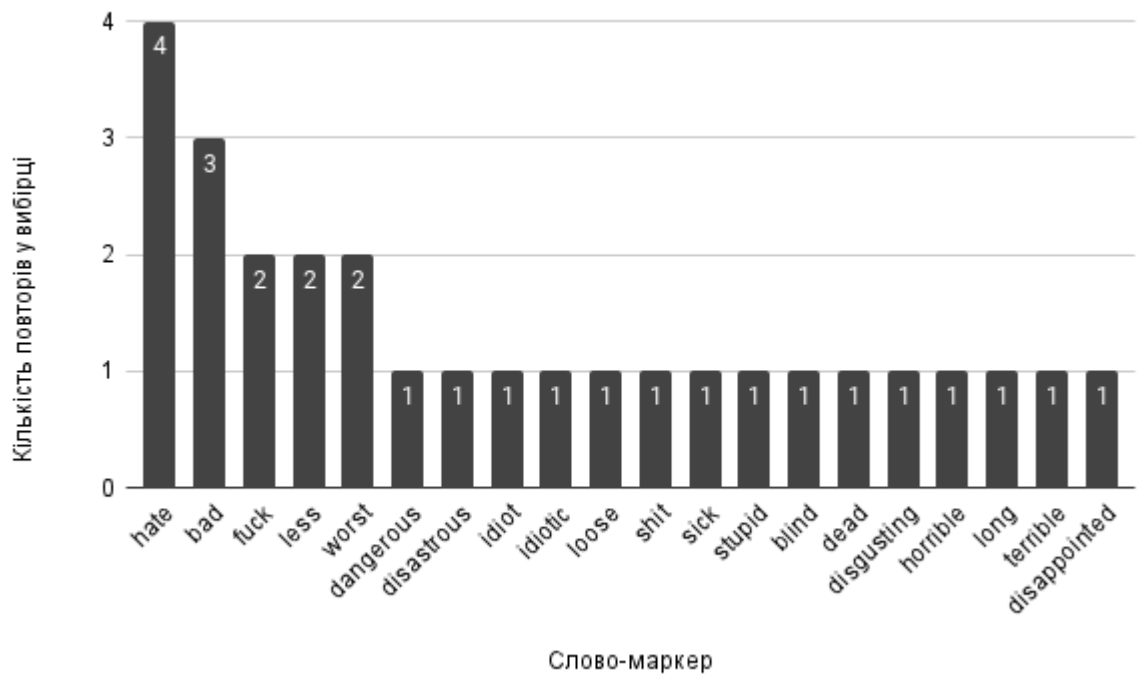


Рис. 3.4.6. – Частота повторів слів-маркерів негативної комунікації за TextBlob

Дана діаграма ілюструє те, що коментатори найчастіше вживали слово «hate» для того, аби передати почуття ненависті, які викликає у них публікація. Слова «bad», «fuck», «less» та «worst» займають позицію як топ-5 найуживаніших мовних маркерів негативної комунікації. Із п'яти найчастіше використовуваних у коментарях негативних слів є одне слово з прошарку нецензурної лексики, що вказує на те, що люди здатні вживати подібні слова, задля вираження сильного емоційного потрясіння, але, на жаль, часте вживання лайливих слів можемо сприймати як ознаку низького рівня культури спілкування чи відсутності розмірковування над кращими способами самовираження.

Одним із правил соцмережі Instagram є те, що лайка заборонена, якщо спрямована на образи або мову ворожнечі, але аналіз вибірки коментарів підтверджує, що алгоритми соціальної мережі не є досконалими та не можуть розпізнавати та блокувати ненормативну лексику.

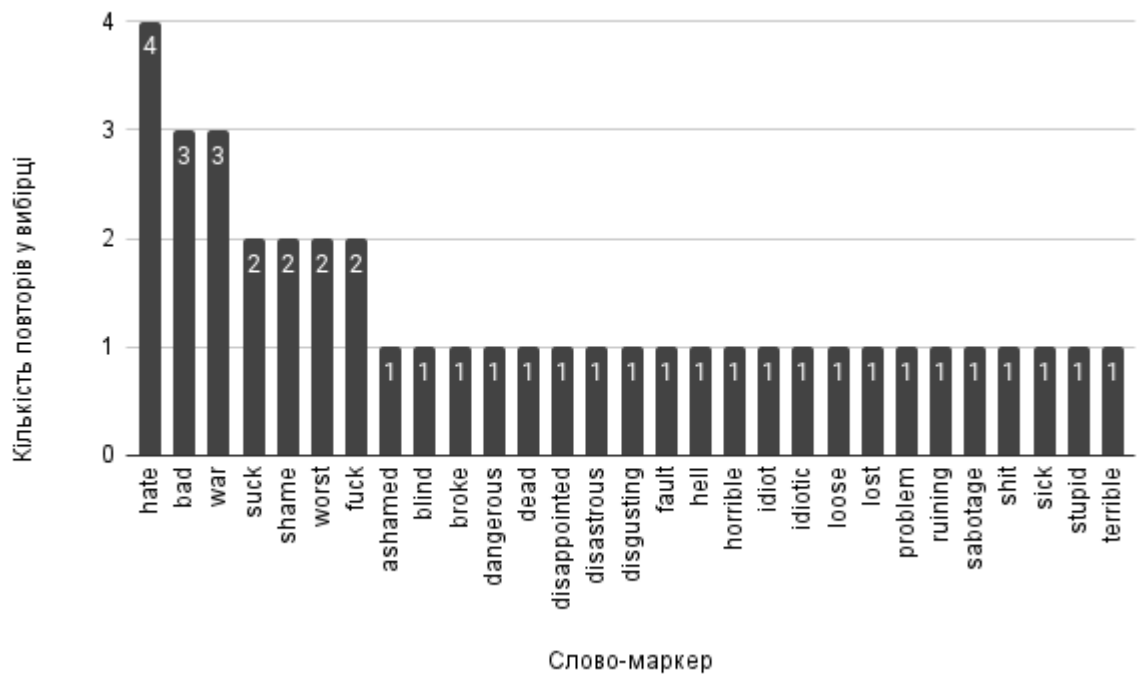


Рис. 3.4.7. – Частота повторів слів-маркерів негативної комунікації за VADER

За результатами сентимент-аналізу за допомогою бібліотеки VADER виявлено, що слово «hate» найчастіше повторно вживається як маркер негативу у вибірках, вказуючи на сильну неприязнь чи ненависть автора. Слова «bad» та «war» мають частоту використання у 3 рази. Загалом, «bad» є універсальним словом для опису негативу, а слово «war» було неочікувано виявити в аналізі, так як бібліотека TextBlob проігнорувала цю лексему. Такий мовний маркер може вказувати на тематику конфліктів або напруження, що несуть негативний характер.

Зі списку наданих слів знову спостерігаємо нецензурну лексику. Наприклад, слово «fuck», яке є загальновизнаним лайливим словом в англійській мові, часто вживається для вираження негативних емоцій. Слово «suck» у певних контекстах також може мати лайливий відтінок, хоча воно вважається менш грубим, ніж попереднє.

Отже, ненормативна лексика є однією із складових конфліктної комунікації.

Сентимент-аналіз текстів можна застосовувати в соціолінгвістичних дослідженнях. Наприклад, питання гендерних відмінностей та розвінчування стереотипів є досі актуальною проблемою. Існує думка, що жінки гіперболізованіше виражають свої емоції, тобто їх негативні емоції можуть бути більш яскраво вираженими, ніж у чоловіків.

Аналіз вибірки коментарів рубрик «Спортивні новини» та «Шоу-бізнес» продемонстрував, що за обома бібліотеками TextBlob і VADER – чоловіки, дещо негативніше висловлюються у соцмережах. Середня полярність негативного коментаря за TextBlob від жінок становить -0.5025, а від чоловіків -0.6421. Із бібліотекою VADER схожа ситуація: середня полярність коментаря від коментаторів жіночої статі -0.4550, а чоловічої -0.4607 (Рис. 3.4.8).

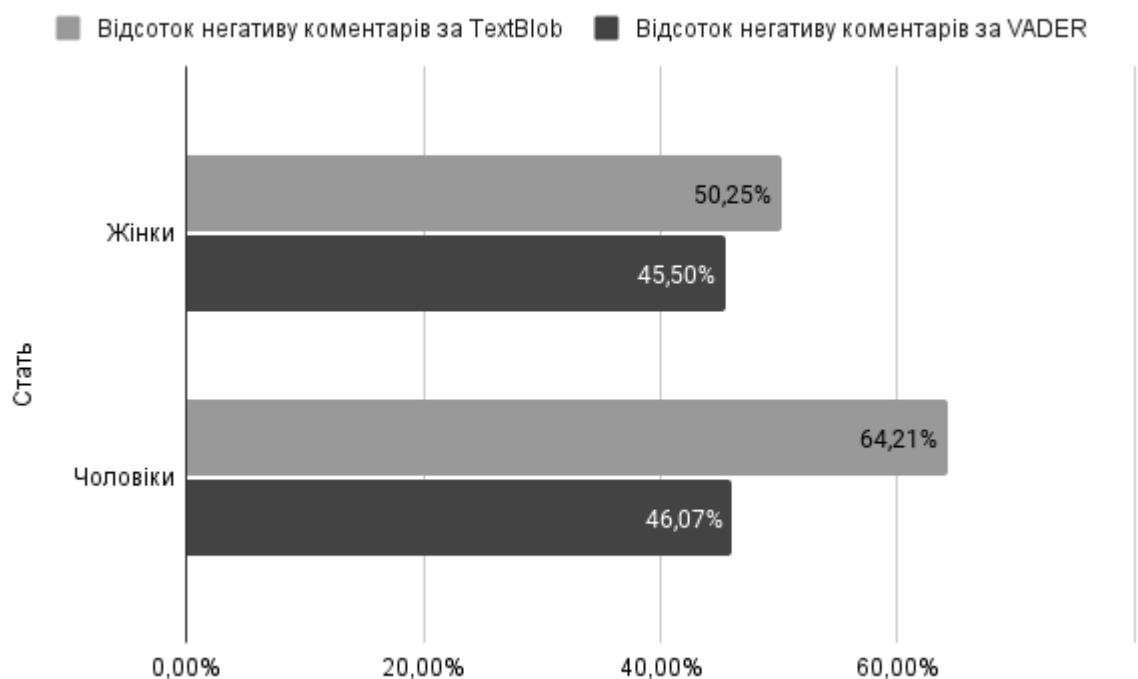


Рис. 3.4.8. Відсоток негативу в коментарях за гендером

Отже, за даними вибірки, спостерігається тенденція того, що коментатори чоловічої статі виражають свої негативні емоції більш інтенсивно, ніж жінки.

Передбачається, що користувачі будуть по-різному реагувати на тематики : «Новини», «Спортивні новини» та «Шоу-бізнес». Новини часто асоціюються з серйозними подіями, такими як політика, катастрофи чи соціальні проблеми, тому це може провокувати додаткову тривогу, обурення або розчарування. Спортивні новини, навпаки, можуть викликати широкий спектр емоцій: від радості та гордості за перемоги улюбленої команди до гніву або розчарування через поразки. У шоу-бізнесі спостерігається, що емоції зазвичай менш інтенсивні, часто спрямовані на захоплення або цікавість, але іноді можуть виникати й негативні реакції на скандали чи суперечки. Тематика рубрики визначає, як коментатори реагують на публікацію, тому це створює різноманітність емоційного тону.

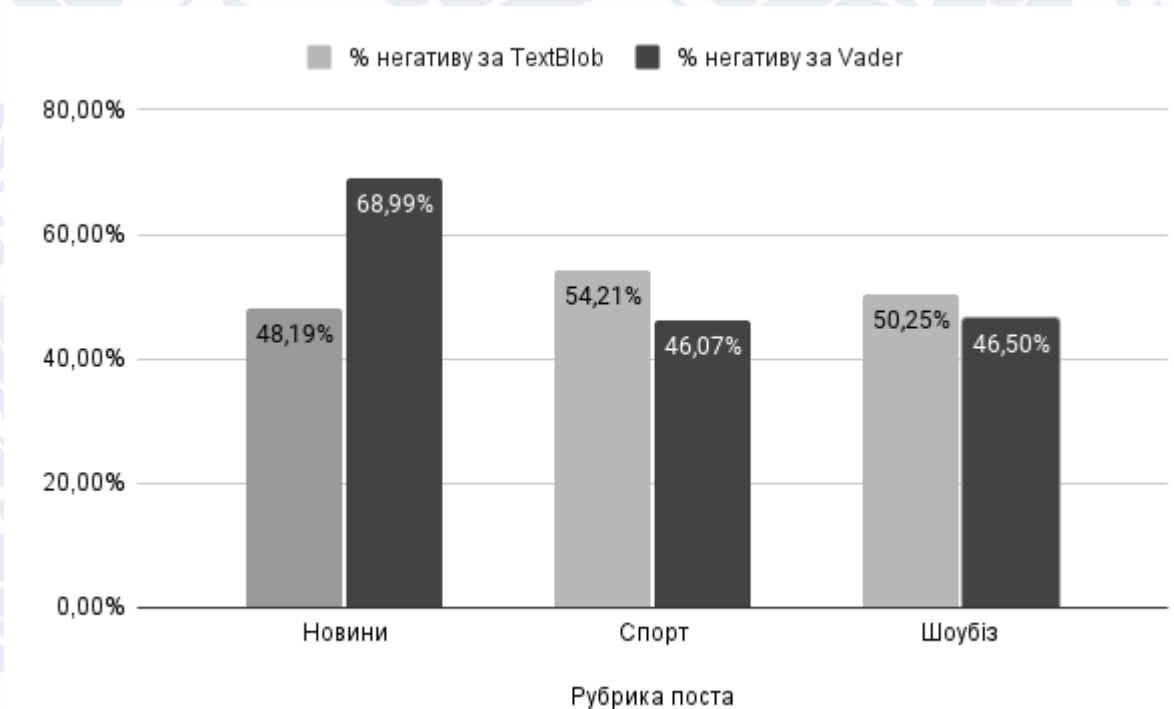


Рис. 3.4.9. Відсоток негативу в коментарях за рубрикою публікації

Дослідження відсотка негативу, висловленого в різних рубриках, таких як новини, спортивні новини та шоу-бізнес, є важливим для розуміння емоційної реакції користувачів на різний контент (Рис. 3. 4. 9). Це дозволяє визначити, в яких категоріях читачі інтенсивніше висловлюють негативні емоції.

За кожною рубрикою підраховано середню полярність відібраних коментарів. До рубрики «Новини» за бібліотекою TextBlob оцінка становила -0.4819, а за VADER -0.6899. Останній результат є найбільш приближеним до максимальної негативної оцінки -1. Група «Спортивні новини» отримала середню оцінку полярності -0.5421 за TextBlob та -0.4607 за VADER. Публікація категорії «Шоу-бізнес» викликала менший резонанс серед користувачів та отримала оцінки -0.5025 за першою бібліотекою, та -0.4650 за другою відповідно.

Діаграма наведена вище ілюструє те, що за аналізом даних з TextBlob, найбільш негативними коментарями були висловлювання під публікаціями спортивних новин. Тональний аналіз із VADER показує, що найбільший відсоток негативності в коментарях, що були залишені під новинним постом CNN.

Аудиторія соціальних мереж активно висловлює свої думки на різні теми. З одного боку, існує стереотипне твердження, що чоловіки більш аргументовано можуть доводити свою думку, тому мають використовувати більшу кількість слів. З іншого боку, жінки більш емоційні, тому можна припустити, що теж використовують багато слів у своїх коментарях, щоб якнайточніше виразити свої почуття.

Дослідження кількості використання негативних слів у коментарях серед жінок і чоловіків є важливим завданням, оскільки дозволяє краще зрозуміти особливості гендерних відмінностей у комунікації. Аналіз таких даних дає можливість виявити, як формуються конфліктні комунікації та які групи суспільства є більш схильними до використання негативної лексики. Діаграма наведена нижче (Рис. 3.4.10) ілюструє кількість використання слів негативної емоції у вибірці відповідно до гендеру.

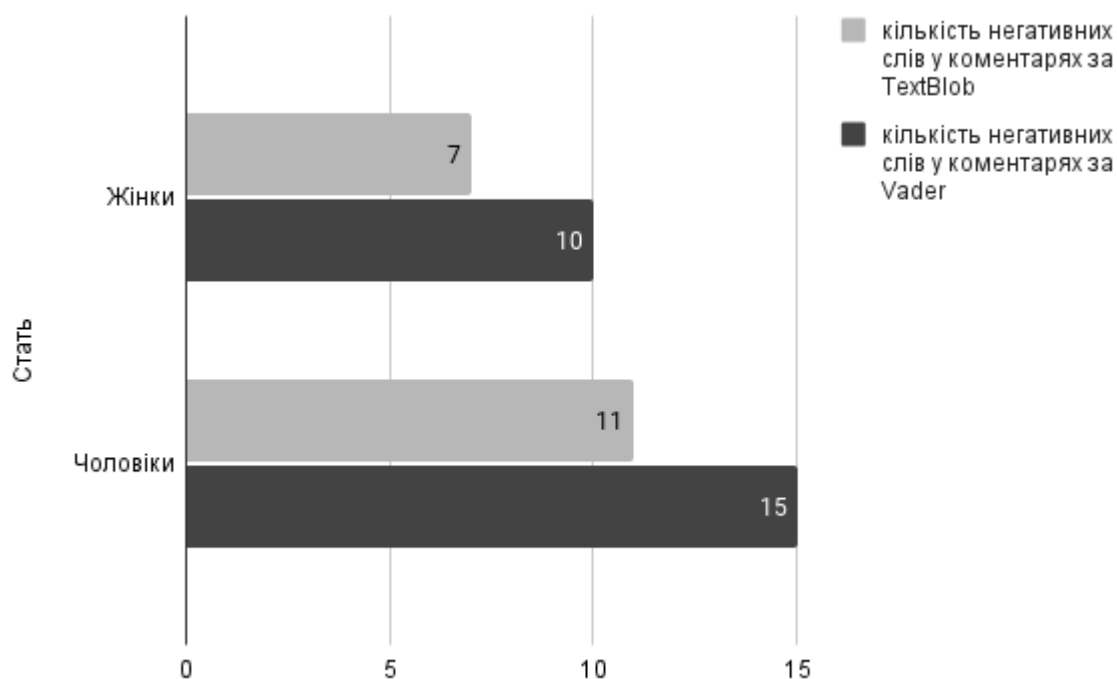


Рис. 3.4.10. Кількість вживаних негативних слів у вибірках відповідно до гендеру

Для аналізу були взяті вибірки груп «Спортивні новини» та «Шоу-бізнес». До першої групи коментарі залишали лише чоловіки, а до другої – жінки. Групи публікацій є спорідненими, тому проведення аналізу є доцільним. За обома інструментами автоматичного сентимент-аналізу більше негативних слів використовують чоловіки, аби виражати свої негативні емоції та оцінки. Отже, судження про те, що чоловіки висловлюють більшою кількістю слів, аніж жінки, має право на буття.

Вибірка негативних коментарів із 3-ох груп надала можливість не лише спостерігати мовні аспекти, а й виявити соціальні закономірності серед проаналізованих даних.

Отримано велику кількість даних для досліджень завдяки використанню інструментів TextBlob та VADER, які стали у нагоді для проведення детального аналізу текстів автоматично. Ці алгоритми дозволили швидко оцінити емоційне забарвлення коментарів і виділити негативно-забарвлену лексику у великій вибірці даних.

Однак результати аналізу, проведеного за допомогою TextBlob і VADER, дещо різняться. Подібні виводи даних спонукають дослідити, який із цих інструментів є ефективнішим і точнішим для sentiment-аналізу (Рис. 3.4.10).

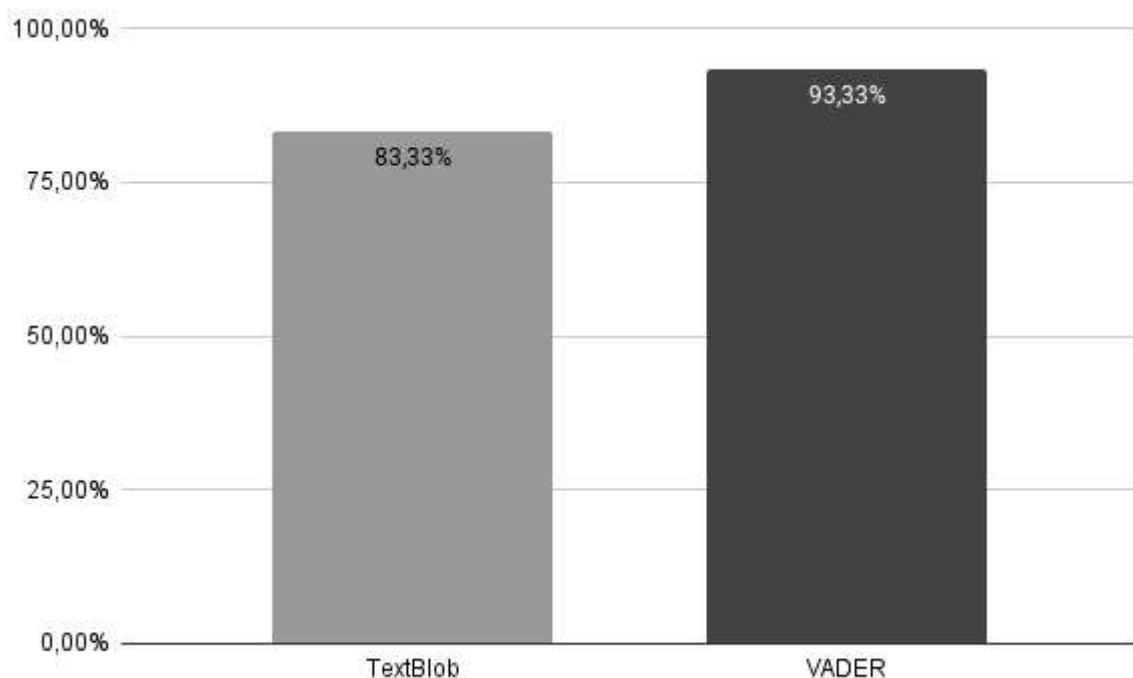


Рис. 3.4.10. Порівняння точності результатів sentiment аналізу за використання бібліотек TextBlob і VADER

Кожна бібліотека проаналізувала по 30 негативних коментарів. Вибірка була створена мануально, а не автоматично, тому всі висловлювання мали негативний характер. Із всієї вибірки бібліотека TextBlob неправильно вивела результат sentiment-аналізу 5-ти коментарям, а VADER – 2-ом.

Вибірка коментарів з Instagram дозволила оцінити ефективність двох популярних інструментів для проведення тонального аналізу. Результати демонструють, що VADER є точнішим для аналізу коментарів з негативним забарвленням, тоді як TextBlob виявився менш надійним у таких умовах. Це підкреслює важливість тестування та порівняння інструментів для вибору оптимального рішення для дослідження коментарів природною мовою із соціальних мереж.

ВИСНОВКИ

Сучасна комунікація в мережі Інтернет часто постає джерелом для вивчення конфліктів. Подібні ситуації виникають як у приватному спілкуванні, так і на публічних обговореннях публікацій в соціальних мережах. Мовні засоби вираження, як-от слова чи візуальні елементи, служать інструментами для вираження емоцій та настроїв, які намагаються передати коментатори. Аналіз впливу цих засобів на взаємодію комунікантів є важливим аспектом цифрової культури, необхідним для попередження конфліктів та їхнього загострення.

У процесі дослідження було зосереджено увагу на автоматизації аналізу тонального забарвлення текстів, на вивченні мовних та графічних засобів, які передають емоції та оцінки, що підкреслюють важливість емоційного забарвлення у розумінні настроїв текстового контенту.

У результаті проведеної роботи вдалося досягти поставленої мети, а саме проаналізувати мовні та графічні засоби, що застосовуються під час конфліктної комунікації, а також оцінити ефективність автоматизованих систем сентимент-аналізу тексту. Дослідження було зосереджене на опрацюванні негативних коментарів у соціальній мережі Instagram, саме вони стали об'єктом детального аналізу.

Виявлено, що мовні засоби негативної комунікації характеризуються активним використанням емоційно забарвленої лексики, скорочень, розмовного сленгу та нецензурної лексики. Дослідження графічних елементів, таких як емотикони, сприяло виведенню гіпотези, що емодзі не лише доповнюють текст, але й виступають самостійними засобами вираження емоцій.

Було здійснено аналіз негативного дискурсу в різних типах публікацій, які охоплюють політичні, спортивні та особисті акаунти. Такий підхід дозволив глибше зрозуміти специфіку цих публікацій та їхній вплив на форму та вираження негативних настроїв.

Використання автоматизованих інструментів TextBlob та VADER для аналізу тональності коментарів дозволило оцінити їхню ефективність. З'ясовано, що TextBlob краще працює з більш формальними текстами, тоді як VADER точніше аналізує короткі та неформальні повідомлення, характерні для соціальних мереж.

Отримані результати підкреслюють важливість урахування специфіки цифрової комунікації при створенні алгоритмів автоматичного аналізу текстів. Практичне значення дослідження полягає у можливості інтеграції виділених мовних маркерів негативу до алгоритмів модерації контенту. Це сприятиме ефективнішому виявленню та нейтралізації конфліктних чи агресивних висловлювань у соціальних мережах.

Таким чином, проведена робота здійснила певний внесок у розвиток прикладної лінгвістики, продемонструвавши наочно взаємодію мовних засобів та сучасних технологій у дослідженні комунікаційних конфліктів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Бацевич Ф. С. Основи комунікативної лінгвістики: підручник. Київ: ВЦ «Академія», 2009. 376 с.
2. Бегунов, А.Б. Алгоритм автоматизованого визначення емоційного забарвлення текстових даних/ Ю.Г. Разуванова ; наук. кер. Заболотня Т.М// Матеріали міжнародної науково-практичної конференції аспірантів і студентів «Інженерія програмного забезпечення 2011», Київ, 2011. — с. 93-97. URL: <https://jrnl.nau.edu.ua/index.php/IPZ/article/view/6631/7406>
3. Білоконенко Л.А. морфологічні маркери мовного конфлікту. URL: <https://elibrary.kdpu.edu.ua/bitstream/0564/1826/1/%D0%9C%D0%BE%D1%80%D1%84%D0%BE%D0%BB%D0%BE%D0%B3%D1%96%D1%87%D0%BD%D1%96%20%D0%BC%D0%B0%D1%80%D0%BA%D0%B5%D1%80%D0%B8%20%D0%BC%D0%BE%D0%B2%D0%BD%D0%BE%D0%B3%D0%BE%20%D0%BA%D0%BE%D0%BD%D1%84%D0%B%D1%96%D0%BA%D1%82%D1%83.pdf>
4. Вахула Б. Я. Соціальні інтернет-мережі, їхні функції та роль у формуванні громадянського суспільства. *Вісник Львівського університету*. Львів, 2012. Вип. 6. с. 312-313.
5. Дикан О. В. Емоційно-оцінна лексика як лінгвістична проблема / Дикан О. В. // Івано-Франківський національний медичний університет, Україна. – 2013. URL: http://www.confcontact.com/2013-alyans-nauk/fl4_dikan.htm
6. Ідзьо М. В. Засоби експресивізації текстів сучасного мас-медійного дискурсу / Мар'яна Ідзьо // Польські студії. – 2014. – № 7. – С. 172-180.
7. Ідзьо М. В. Стилiстичні засоби експресивізації (на матеріалі текстів сучасних польських ЗМІ) / М. В. Ідзьо // Наукові записки Національного університету "Острозька академія". Серія : Філологічна. - 2014. - Вип. 44. - с. 107-109. - URL: http://nbuv.gov.ua/UJRN/Nznuoaf_2014_44_34.

8. Калужинська Ю. Негативнооцінна лексика в мові сучасної української газетної періодики // *Лексика на перетині наукових парадигм: монографія / за ред. Л. Струганець. Тернопіль: Осадца Ю. В., 2018. с. 166–185.* URL: http://dspace.tnpu.edu.ua/bitstream/123456789/17365/1/14_Kaluzhynska.pdf
9. Кузьменко К. О. Емоційно-оцінна лексика в мові сучасних періодичних видань. *Матеріали XLVI наук.-техн. конф. підрозділів ВНТУ (м. Вінниця, 22–24 березня 2017 р.). 2017.* URL: <https://conferences.vntu.edu.ua/index.php/all-hum/all-hum-2017/paper/view/2311>
10. Кульчицький А. Б. Методи аналізу тональності тексту в системах аналізу опіній користувачів соціальних мереж [Текст] / Кульчицький А. Б., Турко Ю. В. // *Комп'ютерні інформаційні технології : матеріали школи-семінару молодих вчених і студентів CIT'2020 [м. Тернопіль, 30 листопада 2020 р.] / відп. за вип. М. П. Дивак. - Тернопіль : ЗУНУ, 2020. - с. 53-54.* URL: <http://dspace.wunu.edu.ua/bitstream/316497/39163/1/53.pdf>
- ЛНУ імені Івана Франка, 2009. – 296 с.
11. Мартин Н. В. Місце оцінності у семантичній структурі слова. URL: <https://naub.ua.edu.ua/mistse-otsinnosti-u-semantychnij-strukturi-slova/>
12. Оленич І., Притула М., Сінкевич О., Хамар О. Система автоматичного визначення тональності тексту// *Електроніка та інформаційні технології. 2021. Випуск 15. с. 16–23* URL: <http://publications.lnu.edu.ua/collections/index.php/electronics/article/download/3467/3770>
13. Панченко К.І. Лінгвістичні особливості текстів малої форми. *Матеріали II Міжнар. наук.-практ. заоч. конф. Київ, 2013. с. 213–214.*
14. Приходько Г.І. Способи вираження оцінки в сучасній англійській мові : монографія. Запоріжжя : Вид-во Запорізького державного ун-ту, 2001. 362 с.

- 15.Публікація південно-корейської зірки в Instagram. URL:
https://www.instagram.com/p/C_adnCovorZ/?img_index=1
- 16.Разуванова, Ю.Г. Мовні засоби вираження емоційного стану радості (на матеріалі твору Сомерсета Моєма "Театр") [Текст] / Ю.Г. Разуванова ; наук. кер. Н.А. Приходько // *Перекладацькі інновації : матеріали III Всеукраїнської студентської науково-практичної конференції, м. Суми, 15-16 березня 2013 р. / Ред.кол.: С.О. Швачко, І.К. Кобякова, О.В. Ємельянова та ін. — Суми : СумДУ, 2013. — с. 119-123.* URL:
<https://essuir.sumdu.edu.ua/bitstream-download/123456789/30575/1/Razuvanova%20Yu.H.%20Movni%20zasoby.pdf> .
- 17.Романюк А., Романишин М. Тональний словник української мови на основі сентимент-анотованого корпусу. *Міжвідомчий науковий збірник «Українське мовознавство»*. — К: Київський національний університет імені Тараса Шевченка, 2013. — Випуск 43. с. 63-74. URL:
http://ukr.movoznavstvo.knu.ua/wp-content/uploads/2019/11/UM_43.pdf
- 18.Сагач Г. М. Слова з емоційними відтінками значень / Г. М. Сагач // *Рідне слово*. — К.: Наук. думка, 1974. — № 9. — с. 36 — 44. URL:
<http://kulturamovy.univ.kiev.ua/KM/pdfs/Magazine9-12.pdf>
- 19.Селіванова О.О. Сучасна лінгвістика: напрями та проблеми : підручник / О. О. Селіванова. — Полтава: Довкілля-К, 2008. — 711 с.
- 20.Федоренко О. І. Основи лінгвістичних досліджень : Підручник / О. І. Федоренко, С. М. Сухорольська О. В. Руда. — Львів : Видавничий центр
- 21.Форманова С. Коментар у соціальних мережах: інвективні форми. URL:
<http://publications.lnu.edu.ua/bulletins/index.php/philology/article/view/10930/11152>
- 22.Шаховська Н. Б. Шкалювання емоційно забарвлених слів для використання у методах класифікації тональності / Н. Б. Шаховська, Х. Ю. Гірак // *Вісник Національного університету «Львівська політехніка»*. Серія: Інформаційні системи та мережі. — Львів :

- Видавництво Львівської політехніки, 2017. — № 872. — с. 195–203.
URL: <http://science.lpnu.ua/sites/default/files/journal-paper/2018/jun/13016/ilovepdfcom-195-203.pdf> .
23. Шингалов Д. В. Методи автоматичного аналізу тональності контенту у соціальних мережах для виявлення інформаційно-психологічних впливів / Д. В. Шингалов, Є. В. Мелешко, Р. М. Минайленко, В. А. Резніченко // *Збірник наукових праць Кіровоградського національного технічного університету. Техніка в сільськогосподарському виробництві, галузеве машинобудування, автоматизація*. - 2017. - Вип. 30. - с. 196-202. - URL: http://nbuv.gov.ua/UJRN/znpkntu_2017_30_29.
24. Andang K., Barli B. Swear words and their implications for English language learning-teaching. URL: https://www.researchgate.net/profile/Barli-Bram/publication/326152797_SWEAR_WORDS_AND_THEIR_IMPLICATIONS_FOR_ENGLISH_LANGUAGE_LEARNING-TEACHING/links/5b7d048a92851c1e1226992f/SWEAR-WORDS-AND-THEIR-IMPLICATIONS-FOR-ENGLISH-LANGUAGE-LEARNING-TEACHING.pdf
25. Bing Liu. Sentiment Analysis and Opinion Mining / Bing Liu. — Morgan & Claypool Publishers, May 2012. URL: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>
26. Bo Pang, Lillian Lee. Opinion Mining and Sentiment Analysis / Bo Pang, Lillian Lee. — 2008. URL: <https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>
27. Cambria E., Das D., Bandyopadhyay A. Practical Guide to Sentiment Analysis, Springer – 2017, 199 с. URL: <https://sentic.net/practical-guide-to-sentiment-analysis.pdf>
28. CNN публікація в Instagram. URL: <https://www.instagram.com/p/DCfXoJmtTJr/>

29. Ditiman Hazarika¹, Gopal Konwar¹, Shuvam Deb¹, Dr. Dibya Jyoti Bora. Sentiment Analysis on Twitter by Using TextBlob for Natural Language Processing. URL: https://www.researchgate.net/publication/348637239_Sentiment_Analysis_on_Twitter_by_Using_TextBlob_for_Natural_Language_Processing
30. ESPN публікація в Instagram (1). URL: https://www.instagram.com/p/DCCgAUfpQNK/?img_index=1
31. ESPN публікація в Instagram (2). URL: <https://www.instagram.com/p/DCkW2PRJutV/>
32. G. Katz. Context-based sentiment analysis / G. Katz, N. Ofek, and B. Shapira // Knowledge-Based Systems. ConSent. — 2015. Vol. 84, No. 1. с. 162—178.
33. Hutto C., Gilbert E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), с. 216-225.
34. Iqbal, L., Safi, F., & Ullah, I. (2020). The Use of Symbols (Emoticons) in Social Media: A Shift of Language from Words to Symbols. Global Mass Communication Review, V(III), с. 124-135. URL: [https://doi.org/10.31703/gmcr.2020\(V-III\).10](https://doi.org/10.31703/gmcr.2020(V-III).10)
https://www.researchgate.net/publication/352189355_The_Use_of_Symbols_Emoticons_in_Social_Media_A_Shift_of_Language_from_Words_to_Symbols
35. Nip J.Y.M., Berthelie B. Social Media Sentiment Analysis. *Encyclopedia* 2024, 4, 1590–1598. URL: https://www.researchgate.net/publication/385165077_Social_Media_Sentiment_Analysis
36. PyCharm [Електронний ресурс]. URL: <https://www.jetbrains.com/pycharm/features/>
37. Reis João. Enhancing Python-Based Sentiment Analysis: Empowering Industrial Engineers and Managers in the Service Industry. URL: https://www.researchgate.net/publication/383157980_Enhancing_Python-

- [based Sentiment Analysis Empowering Industrial Engineers and Managers in the Service Industry](#)
38. Sentiment Analysis in Python With TextBlob [Електронний ресурс]. URL: <https://stackabuse.com/sentiment-analysis-in-python-with-textblob/> .
39. SentiStrength [Електронний ресурс]. URL: <https://mi-linix.wlv.ac.uk/~cm1993/sentistrength/>
40. Steven Loria Website [Електронний ресурс]. URL: <https://stevenloria.com/>
41. Taboada M., Brooke J., Tofiloski M., Kimberly V., Manfred S. Lexicon-Based Methods for Sentiment Analysis. URL: <https://direct.mit.edu/coli/article/37/2/267/2105/Lexicon-Based-Methods-for-Sentiment-Analysis>
42. TextBlob: Simplified Text Processing [Електронний ресурс]. URL: <https://textblob.readthedocs.io/en/dev/#>
43. The MIT License (MIT) [Електронний ресурс]. URL: <https://mit-license.org/>
44. The Natural Language API [Електронний ресурс]. URL: <https://cloud.google.com/natural-language/docs/basics>
45. Thelwall M., Buckley K., Paltoglou G. Sentiment strength detection in short informal text // *Journal of the American Society for Information Science and Technology*. 2010. No61. с. 2544–2558 URL: <http://www.scit.wlv.ac.uk/~cm1993/papers/SentiStrengthPreprint.pdf>
46. Tripathil V., Joshil A., Pushpak B. Emotion Analysis from Text: A Survey. URL: <https://www.cfilt.iitb.ac.in/resources/surveys/emotion-analysis-survey-2016-vaibhav.pdf>
47. VADER Sentiment-Analysis Introduction. [Електронний ресурс]. URL: <https://vadersentiment.readthedocs.io/en/latest/index.html>
48. Volovenko, I., Grigorenko, I. (2018). Metamovna refleksiya merczhevykh tekstiv. [Metalanguage Network Display]. *Intercultural Communication*, ISSN 2451-0998, vol. 1 (4) 2018 pp. 57-70. (In Ukrainian) URL:

<https://bibliotekanauki.pl/api/full-texts/2022/2/7/ff98227a-3274-46f5-a29e-ebeb8089c8ad.pdf>

49. Wan Nur Syahirah Wan Min, Nur Zareen Zulkarnain. Comparative Evaluation of Lexicons in Performing Sentiment Analysis. URL: <https://jacta.utm.edu.my/jacta/article/view/5207>
50. YouScan [Електронний ресурс]. URL: <https://youscan.io/ua/>

Додатки
ДОДАТОК А

Вибірка негативних коментарів до посту новинного акаунту CNN в Instagram та результати сентимент аналізу за допомогою бібліотеки TextBlob

№	Текст коментаря	Кількість слів	Позитивні/негативні мовні маркери в коментарі	Результат сентимент-аналізу за TextBlob	Стать коментатора
1	RIP Ukraine!!!! Russia is about to let loose a gangster arsenal!!	10	Слова з негативним забарвленням: loose: -0.07692307692307693	Полярність тексту: -0.1201923076923077 Негативна	чоловіча
2	Shame on the democrats!!!	4		Полярність тексту: 0.0 Нейтральна	чоловіча
3	Absolutely idiotic decision it will not change anything but it might create the world War three	15	Слова з позитивним забарвленням: Absolutely: 0.2 Слова з негативним забарвленням: idiotic: -0.6666666666666666	Полярність тексту: -0.6666666666666666 Негативна	чоловіча
4	Disastrous	1	Слова з негативним забарвленням: Disastrous: -0.7	Полярність тексту: -0.7 Негативна	чоловіча

5	Trying to sabotage the country because they lost	7		Полярність тексту: 0.0 Нейтральна	чоловіча
6	Stupid decision	2	Слова з негативним забарвленням: Stupid: -0.7999999999999999	Полярність тексту: -0.7999999999999999 Негативна	чоловіча
7	USA president is a sick man	6	Слова з негативним забарвленням: Sick: -0.7142857142857143	Полярність тексту: -0.7142857142857143 Негативна	чоловіча
8	Dangerous senile old shit.. couldn't care less for peoples lives	9	Слова з позитивним забарвленням: old: 0.1 Слова з негативним забарвленням: Dangerous: -0.6 shit: -0.2 less: -0.16666666666666666	Полярність тексту: -0.21666666666666665 Негативна	чоловіча
9	This idiot gonna have us in a WAR!	7	Слова з негативним забарвленням: idiot: -0.8	Полярність тексту: -1.0 Негативна	чоловіча
10	Fuck it start a war before leaving	7	Слова з негативним забарвленням: Fuck: -0.4	Полярність тексту: -0.4 Негативна	чоловіча

ДОДАТОК Б

Вибірка негативних коментарів до посту новинного акаунту CNN в Instagram та результати sentiment аналізу за допомогою бібліотеки VADER

№	Текст коментаря	Кількість слів	Позитивні/негативні мовні маркери в коментарі	Результат sentiment-аналізу за VADER	Стать коментатора
1	RIP Ukraine!!!! Russia is about to let loose a gangster arsenal!!	10	Слова з негативним забарвленням (VADER): loose: -0.3182	Полярність тексту за VADER: -0.5374 Негативна	чоловіча
2	Shame on the democrats!!!	4	Слова з негативним забарвленням (VADER): Shame: -0.4767	Полярність тексту за VADER: -0.6093 Негативна	чоловіча
3	Absolutely idiotic decision it will not change anything but it might create the world War three	15	Слова з позитивним забарвленням (VADER): create: 0.2732 Слова з негативним забарвленням (VADER): idiotic: -0.5574 War: -0.5994	Полярність тексту за VADER: -0.7308 Негативна	чоловіча
4	Disastrous	1	Слова з негативним забарвленням (VADER): Disastrous: -0.5994	Полярність тексту за VADER: -0.5994 Негативна	чоловіча

5	Trying to sabotage the country because they lost	7	Слова з негативним забарвленням (VADER): sabotage: -0.5267 lost: -0.3182	Полярність тексту за VADER: -0.6908 Негативна	чоловіча
6	Stupid decision	2	Слова з негативним забарвленням (VADER): Stupid: -0.5267	Полярність тексту за VADER: -0.5267 Негативна	чоловіча
7	USA president is a sick man	6	Слова з негативним забарвленням (VADER): Sick: -0.5106	Полярність тексту за VADER: -0.5106 Негативна	чоловіча
8	Dangerous senile old shit.. couldn't care less for peoples lives	9	Слова з позитивним забарвленням (VADER): care: 0.4939 Слова з негативним забарвленням (VADER): Dangerous: -0.4767 shit..: -0.5574	Полярність тексту за VADER: -0.5423 Негативна	чоловіча
9	This idiot gonna have us in a WAR!	7	Слова з негативним забарвленням (VADER): Fuck: -0.5423 war: -0.5994	Полярність тексту за VADER: -0.8126 Негативна	чоловіча
10	Fuck it start a war before leaving	7	Слова з негативним забарвленням (VADER): idiot: -0.5106 WAR!:-0.636	Полярність тексту за VADER: -0.8491 Негативна	чоловіча

ДОДАТОК В

Вибірка негативних коментарів до посту акаунту південно-корейської зірки в Instagram та результати сентимент аналізу за допомогою бібліотеки TextBlob

№	Текст коментаря	Кількість слів	Позитивні/негативні мовні маркери в коментарі	Результат сентимент-аналізу за TextBlob	Стать коментатора
1	I hate you	3	Слова з негативним забарвленням: hate: -0.8	Полярність тексту: -0.8 Негативна	жіноча
2	Disappointed in you	3	Слова з негативним забарвленням: Disappointed: -0.75	Полярність тексту: -0.75 Негативна	жіноча
3	What the fuck, man	4	Слова з негативним забарвленням: fuck: -0.4	Полярність тексту: -0.4 Негативна	жіноча
4	The worst wedding dress I've ever seen	7	Слова з негативним забарвленням: worst: -1.0	Полярність тексту: -1.0 Негативна	жіноча
5	I hate her, but that flowers are SO PRETTY	9	Слова з позитивним забарвленням: PRETTY: 0.25 Слова з негативним забарвленням: hate: -0.8	Полярність тексту: -0.275 Негативна	жіноча

6	Both of you can go to hell and burn forever	10		Полярність тексту: 0.0 Нейтральна	жіноча
7	You deserve nothing but the worst.	6	Слова з негативним забарвленням: worst: -1.0	Полярність тексту: -1.0 Негативна	жіноча
8	Spreading hate to you both	5	Слова з негативним забарвленням: hate: -0.8	Полярність тексту: -0.8 Негативна	жіноча
9	That's what I call ruining your life.	7		Полярність тексту: 0.0 Нейтральна	жіноча
10	Shame on you	3		Полярність тексту: 0.0 Нейтральна	жіноча

ДОДАТОК Г

Вибірка негативних коментарів до посту акаунту південно-корейської зірки в Instagram та результати сентимент аналізу за допомогою бібліотеки VADER

№	Текст коментаря	Кількість слів	Позитивні/негативні мовні маркери в коментарі	Результат сентимент-аналізу за VADER	Стать коментатора
1	I hate you	3	Слова з негативним забарвленням (VADER): hate: -0.5719	Полярність тексту за VADER: -0.5719 Негативна	жіноча
2	Disappointed in you	3	Слова з негативним забарвленням (VADER): Disappointed: -0.4767	Полярність тексту за VADER: -0.4767 Негативна	жіноча
3	What the fuck, man	4	Слова з негативним забарвленням (VADER): fuck,: -0.5423	Полярність тексту за VADER: -0.5423 Негативна	жіноча
4	The worst wedding dress I've ever seen	7	Слова з негативним забарвленням (VADER): worst: -0.6249	Полярність тексту за VADER: -0.6249 Негативна	жіноча
5	I hate her, but that flowers are SO PRETTY	9	Слова з позитивним забарвленням (VADER): PRETTY: 0.4939 Слова з негативним	Полярність тексту за VADER: 0.8431 Позитивна	жіноча

			забарвленням (VADER): hate: -0.5719		
6	Both of you can go to hell and burn forever	10	Слова з негативним забарвленням (VADER): hell: -0.6808	Полярність тексту за VADER: -0.6808 Негативна	жіноча
7	You deserve nothing but the worst.	6	Слова з негативним забарвленням (VADER): worst.: -0.6249	Полярність тексту за VADER: -0.6249 Негативна	жіноча
8	Spreading hate to you both	5	Слова з негативним забарвленням (VADER): hate: -0.5719	Полярність тексту за VADER: -0.5719 Негативна	жіноча
9	That's what I call ruining your life.	7	Слова з негативним забарвленням (VADER): ruining: -0.25	Полярність тексту за VADER: -0.25 Негативна	жіноча
10	Shame on you	3	Слова з негативним забарвленням (VADER): Shame: -0.4767	Полярність тексту за VADER: -0.4767 Негативна	жіноча

ДОДАТОК Г

Вибірка негативних коментарів до посту спортивного новинного акаунту ESPN в Instagram та результати сентимент аналізу за допомогою бібліотеки TextBlob

№	Текст коментаря	Кількість слів	Позитивні/негативні мовні маркери в коментарі	Результат сентимент-аналізу за TextBlob	Стать коментатора
1	His Barber Should Be Ashamed At That Hairline	8		Полярність тексту: 0.0 Нейтральна	чоловіча
2	He looks like hes legally blind	6	Слова з позитивним забарвленням: legally: 0.2 Слова з негативним забарвленням: blind: -0.5	Полярність тексту: -0.5 Негативна	чоловіча
3	Already horrible with his money	5	Слова з негативним забарвленням: horrible: -1.0	Полярність тексту: -1.0 Негативна	чоловіча
4	and will be dead broke in less than 10 years	10	Слова з негативним забарвленням: dead: -0.2 less: -0.16666666666666666	Полярність тексту: - 0.18333333333333335 Негативна	чоловіча

5	That girls skin complexion is disgusting	6	Слова з негативним забарвленням: disgusting: -1.0	Полярність тексту: -1.0 Негативна	чоловіча
6	We still suck regardless	4		Полярність тексту: 0.0 Нейтральна	чоловіча
7	What an absolutely terrible decision. Not his fault our players and goaltending suck	13	Слова з позитивним забарвленням: absolutely: 0.2 Слова з негативним забарвленням: terrible: -1.0	Полярність тексту: -1.0 Негативна	чоловіча
8	OK, that was another bad idea in a long list of bad ideas. Montgomery was not the problem.	18	Слова з позитивним забарвленням: OK: 0.5 Слова з негативним забарвленням: bad: -0.6999999999999998 long: -0.05 bad: -0.6999999999999998	Полярність тексту: -0.23749999999999993 Негативна	чоловіча
9	Bad move	2	Слова з негативним забарвленням: Bad: -0.6999999999999998	Полярність тексту: -0.6999999999999998 Негативна	чоловіча
10	i hate the bruins for this	6	Слова з негативним забарвленням: hate: -0.8	Полярність тексту: -0.8 Негативна	чоловіча

ДОДАТОК Д

Вибірка негативних коментарів до посту спортивного новинного акаунту ESPN в Instagram та результати сентимент аналізу за допомогою бібліотеки VADER

№	Текст коментаря	Кількість слів	Позитивні/негативні мовні маркери в коментарі	Результат сентимент-аналізу за VADER	Стать коментатора
1	His Barber Should Be Ashamed At That Hairline	8	Слова з негативним забарвленням (VADER): Ashamed: -0.4767	Полярність тексту за VADER: -0.4767 Негативна	чоловіча
2	He looks like hes legally blind	6	Слова з позитивним забарвленням (VADER): like: 0.3612 legally: 0.1027 Слова з негативним забарвленням (VADER): blind: -0.4019	Полярність тексту за VADER: 0.0516 Позитивна	чоловіча
3	Already horrible with his money	5	Слова з негативним забарвленням (VADER): horrible: -0.5423	Полярність тексту за VADER: -0.5423 Негативна	чоловіча
4	and will be dead broke in less than 10 years	10	Слова з негативним забарвленням (VADER):	Полярність тексту за VADER: -0.7964 Негативна	чоловіча

			dead: -0.6486 broke: -0.4215		
5	That girls skin complexion is disgusting	6	Слова з негативним забарвленням (VADER): disgusting: -0.5267	Полярність тексту за VADER: -0.5267 Негативна	чоловіча
6	We still suck regardless	4	Слова з негативним забарвленням (VADER): suck: -0.4404	Полярність тексту за VADER: -0.4404 Негативна	чоловіча
7	What an absolutely terrible decision. Not his fault our players and goaltending suck	13	Слова з негативним забарвленням (VADER): terrible: -0.4767 fault: -0.4019 suck: -0.4404	Полярність тексту за VADER: -0.6168 Негативна	чоловіча
8	OK, that was another bad idea in a long list of bad ideas. Montgomery was not the problem.	18	Слова з негативним забарвленням (VADER): bad: -0.5423 bad: -0.5423 problem.: -0.4019	Полярність тексту за VADER: -0.6948 Негативна	чоловіча
9	Bad move	2	Слова з негативним забарвленням (VADER): bad: -0.5423	Полярність тексту за VADER: -0.5423 Негативна	чоловіча
10	i hate the bruins for this	6	Слова з негативним забарвленням (VADER): hate: -0.5719	Полярність тексту за VADER: -0.5719 Негативна	чоловіча

