Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

Використання методів opinion mining для визначення суспільних настроїв

(тема)

Виконав:
здобувач четвертого року навчання,
групи ІТШІ-21-3

Артур Лавров
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)
Тип програми освітньо-професійна
Освітня програма Штучний інтелект

(повна назва освітньої програми)

Керівник доц. Марина Кудрявцева
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ _____ Олег ЗОЛОТУХІН
(підпис) (власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Штучного інтелекту

Рівень вищої освіти перший (бакалаврський)

Спеціальність 122 Комп'ютерні науки
(код і повна назва)

Тип програми освітньо-професійна

Освітня програма Штучний інтелект
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«_____» _____ 20 ___ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Лаврову Артуру Олександровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Використання методів opinion mining для визначення суспільних настроїв

затверджена наказом університету від 19 травня 20 25 р. № 378Ст

2. Термін подання студентом роботи до екзаменаційної комісії 19 червня 20 25 р.

3. Вихідні дані до роботи Науково-технічні публікації з обробки природної мови та машинного навчання, дослідження класичних та вдосконалених архітектур нейронних мереж (RNN, CNN, Transformers), дослідження з аналізу настроїв, видобутку думок, аспектного аналізу настроїв та виявлення емоцій, література про різні алгоритми обробки та аналізу тексту, документація та дослідження щодо конкретних моделей, таких як BERT, RoBERTa та інших, відкриті набори даних для загального аналізу настроїв, аспектів та емоцій (наприклад, Financial PhraseBank, завдання SemEval, набори даних соціальних мереж).

4. Перелік питань, що потрібно опрацювати в роботі

1) Аналіз предметної галузі та постановка задачі

2) Теоретичне підґрунтя та обрання референтних моделей для видобутку думок

3) Розробка емпіричних моделей, оцінювання та виведення архітектур

4) Практична реалізація

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	19.05.2025	виконано
2	Аналіз предметної галузі та постановка задачі	26.05.2025	виконано
3	Огляд методологій для аналізу настроїв	02.06.2025	виконано
4	Моделювання проектних архітектур	05.06.2025	виконано
5	Експериментальне дослідження моделей	07.06.2025	виконано
6	Аналіз отриманих результатів	09.06.2025	виконано
7	Написання пояснювальної записки	10.06.2025	виконано
8	Перевірка на академічний плагіат	10.06.2025	виконано
9	Нормоконтроль	12.06.2025	виконано
10	Підготовка презентації та доповіді	16.06.2025	виконано
11	Рецензуваня	17.06.2025	виконано
12	Захист перед ЕК	19.06.2025	

Дата видачі завдання 19 травня 2025 р.

Здобувач _____
(підпис)

Керівник роботи _____ доц. Марина Кудрявцева
(підпис) (посада, власне ім’я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 85 с., 19 рис., 3 табл., 4 дод., 30 джерел.

АНАЛІЗ НАСТРОЇВ, АНАЛІЗ ТОНАЛЬНОСТІ, АСПЕКТНИЙ АНАЛІЗ НАСТРОЇВ, ГРОМАДСЬКА ДУМКА, ОБРОБКА ПРИРОДНОЇ МОВИ, РОЗПІЗНАВАННЯ ЕМОЦІЙ, РОЗПІЗНАВАННЯ ІМЕНОВАНИХ СУТНОСТЕЙ.

Об’єкт дослідження – динамічна взаємодія між суспільними настроями з новин і соціальних медіа, з подальшим акцентом на її вплив на ключові галузі, об’єкти або глобальні події.

Предмет дослідження – підходи для визначення та розуміння сукупності індивідуальних поглядів або переконань, яких дотримується населення та нюансів, на кшталт емоційного тону, усередині фрагмента тексту, щоб вловити його повне значення та оцінити можливий вплив.

Мета роботи – дослідження методів, що дасть змогу оцінити вплив громадської думки, щоб надати зацікавленим сторонам цінні відомості про взаємозв’язок між інформацією про аспекти і громадською думкою.

Методи дослідження – вивчення і реалізація систем для забезпечення розробки як і класичних, так і передових архітектур нейронних мереж для налагодження вилучення контексту, розпізнавання настроїв та емоцій.

У даній роботі проводиться всебічне дослідження комп’ютерного аналізу громадської думки, вираженої в цифровому тексті. Для досягнення мети був реалізований, оцінений і продемонстрований набір еталонних моделей, призначених для вирішення різних аналітичних завдань. Завдання варіюються від класифікації полярності офіційних новин і неформальних коментарів у соціальних мережах до тонкого розпізнавання емоцій і, в кінцевому підсумку, до детального аспектного аналізу настроїв (ABSA), підтримуваного розпізнаванням іменованих сутностей (NER).

ABSTRACT

Bachelor’s thesis contains: 85 pp., 19 fig., 3 tabl., 4 ann., 30 references.

ASPECTUAL SENTIMENT ANALYSIS, EMOTION RECOGNITION, NAMED ENTITY RECOGNITION, NATURAL LANGUAGE PROCESSING, PUBLIC OPINION, SENTIMENT ANALYSIS, TONALITY ANALYSIS.

The object of study is the dynamic interaction between public sentiment from news and social media, with a subsequent focus on its impact on key industries, objects, or global events.

The subject of the study is approaches to identifying and understanding the set of individual views or beliefs held by the population and nuances, such as emotional tone, within a text fragment in order to grasp its full meaning and assess its potential impact.

The goal of the work is to investigate methods that will allow the impact of public opinion to be assessed in order to provide stakeholders with valuable information about the relationship between information about aspects and public opinion.

Research methods include the study and implementation of systems to ensure the development of classical and advanced neural network architectures for context extraction, sentiment recognition, and emotion recognition.

This work conducts a comprehensive study of computer analysis of public opinion expressed in digital text. To achieve this goal, a set of reference models designed to solve various analytical tasks was implemented, evaluated and demonstrated. The tasks range from classifying the polarity of official news and informal comments on social networks to subtle emotion recognition and, ultimately, to detailed aspect-based sentiment analysis (ABSA) supported by named entity recognition (NER).

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ШІ – штучний інтелект;

ABSA – Aspect-Based Sentiment Analysis – аспектно-орієнтований аналіз настроїв;

AI – Artificial Intelligence – штучний інтелект;

CNN – Convolutional Neural Networks – конволюційні нейронні мережі;

CRF – Conditional Random Fields – умовні випадкові поля;

GPT – Generative Pre-trained Transformer – генеративний попередньо навчений трансформатор;

GRU – Gated Recurrent Unit – закритий рекурентний блок;

LSTM – Long Short-Term Memory – довга короткочасна пам’ять;

NER – Named Entity Recognition – розпізнавання іменованих сутностей;

NLP – Natural Language Processing – обробка природної мови;

PMI – Pointwise Mutual Information – поточкова взаємна інформація;

POS – Part-of-Speech – мітки частини мови;

RNN – Recurrent Neural Networks – рекурентні нейронні мережі;

SVM – Support Vector Machines – метод опорних векторів.

ВСТУП

Сучасна епоха характеризується безпрецедентним поширенням цифрового опосередкованого спілкування, що призводить до експоненціального зростання обсягу текстового контенту. Це цифрове середовище, що охоплює соціальні мережі, онлайн-агрегатори відгуків, дискусійні форуми й особисті блоги, являє собою велике сховище людських пізнавальних і афективних станів. Так званий «цифровий архів» містить значний обсяг суб'єктивної інформації – неструктурованих висловлювань, що відображають спектр людських переживань, включно зі схваленням, невдоволенням і нюансами емоційних реакцій. Тому й здатність систематично опрацьовувати ці колективні думки, виявляти переважаючі настрої, світоглядні орієнтації та основні проблеми перетворилася із суто академічного завдання на стратегічний імператив для ефективної роботи в сучасній соціально-економічній та інформаційній екосистемі.

Фундаментально наукові дослідження, спрямовані на інтерпретацію цифрових настроїв, являють собою масштабну характеристику громадської та індивідуальної думки. Традиційні методики оцінювання громадської думки, як-от опитування і фокус-групи, хоча й мають корисність, але за своєю суттю обмежені часовим дозволом і обмеженнями вибірки. Сучасне цифрове середовище пропонує безпрецедентний безперервний потік висловлених думок, являючи собою як значну можливість для комплексного аналізу, так і серйозну методологічну проблему. Можливість полягає в доступі до динамічного, ґлобального дискурсу; проблема потребує розроблення надійних обчислювальних методів, здатних ефективно прослуховувати, семантично розчленовувати й витягувати корисні відомості з терабайтів текстових даних. Це і представляє сфера діяльності Opinion Mining, орієнтована на обчислювальну ідентифікацію, категоризацію та кількісне оцінювання думок, емоцій і суб'єктивних станів, що формулюються в безлічі текстових артефактів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Передумови концепцій та важливість розуміння

Аналіз тональності тексту, який також часто називають «Аналіз настроїв», є важливою галуззю в області обробки природної мови (NLP), комп’ютерної лінгвістики, інтелектуального аналізу текстів та пошуку інформації (IR). Основна його спрямованість – автоматизований процес ідентифікації, вилучення, аналізу та класифікації суб’єктивної інформації, такої як думки, настрої, судження, ставлення та емоції, що виражені в письмовій або усній мові. Основна мета полягає у визначенні загального ставлення або емоційного тону, переданого в тексті, зазвичай класифікуючи його як позитивний, негативний або нейтральний.

Хоча ці терміни часто використовуються як взаємозамінні, особливо в промисловості, іноді між ними проводяться певні академічні відмінності.

«Opinion Mining» можна розглядати як дещо ширшу концепцію, що зосереджується на отриманні та аналізі думок про конкретні об’єкти та їхні атрибути чи особливості, включаючи переконання та судження, які не обов’язково ґрунтуються на фактах. Це часто передбачає визначення того, про що говорять і яку думку висловлюють.

«Sentiment Analysis», навпаки, іноді розглядають як підзадачу, що фокусується на емоційному забарвленні використаної мови та передусім стосується визначення та класифікації загальної полярності настроїв (позитивних/негативних/нейтральних) у тексті.

Однак для більшості практичних цілей і в більшій частині дослідницької літератури вони вважаються синонімічними парасольковими термінами, що охоплюють цілу низку пов’язаних завдань.

Процес сам по собі багатогранний і виходить за рамки простої класифікації полярності. Основні завдання, які часто включають в себе:

а) виявлення/класифікація суб’єктивності: розрізнення між фактичними твердженнями (об’єктивними) і текстом, що містить думки або почуття (суб’єктивними). Об’єктивний текст часто менш релевантний для прямого аналізу настроїв. Суб’єктивність відноситься до мовного вираження особистих станів, таких як думки, оцінки та переконання [1];

б) класифікація настроїв: присвоєння полярності (позитивна, негативна, нейтральна) суб’єктивному тексту. Це можна зробити на різних рівнях деталізації [2]:

– рівень документа: класифікація загального настрою всього документа (наприклад, огляду продукту);

– рівень речення: визначення настрою, вираженого в окремих реченнях;

– рівень аспекту/особливості (ABSA): визначення настрою, пов’язаного з конкретними аспектами або особливостями об’єкта (наприклад, «Їжа була чудовою [позитивна оцінка], але обслуговування було повільним [негативна оцінка]»). Це має вирішальне значення для отримання детальних інсайтів;

в) виділення носіїв думки: визначення суб’єкта (особи або організації), який висловлює думку;

г) виділення аспекту/суб’єкта: визначення об’єкта думки – конкретного продукту, послуги, теми або функції, що обговорюється;

д) виявлення емоцій: вихід за межі полярності для виявлення конкретних емоцій, таких як радість, гнів, смуток, страх тощо;

е) шкала інтенсивності: визначення сили почуття (наприклад, злегка позитивне проти дуже позитивного).

Застосовуються різні методи, які загалом поділяються на лексикографічні (з використанням словників слів, позначених оцінками настроїв), машинне навчання (контрольоване, неконтрольоване, напівконтрольоване) та гібридні підходи. Моделі глибокого навчання,

зокрема трансформатори, такі як BERT, продемонстрували значний прогрес у визначенні контексту та нюансів у мові для завдань аналізу настроїв.

Однак проблеми залишаються, особливо щодо контекстно-залежних значень, сарказму, іронії, обробки заперечень, імпліцитних думок і багатомовного аналізу.

1.2 Прихильність зацікавлених сторін

Важливість аналізу настроїв і вивчення думок зросла в геометричній прогресії, чому в основній мірі сприяло вибухове зростання користувацького контенту в Інтернеті. Зручність і доступність онлайн-додатків значно змінилася, збільшилася кількість соціальних платформ для обміну думками та поглядами, сайтів онлайн-оглядів і особистих блогів, які привернули увагу зацікавлених сторін, не просто таких як клієнти, а й організації, і уряди. Можливість автоматично обробляти та розуміти подібні дані забезпечить безцінно вигідну інформацію для її власника.

Основною і найпомітнішою сферою застосування, ймовірно, буде Бізнес і Маркетинг. Компанії отримують думки з відгуків клієнтів, оглядів, коментарів у соціальних мережах і різних опитуваннях. Аналіз думок дає змогу визначити напрямок розвитку продукту, з’ясувавши, що подобається або не подобається покупцям у поточних пропозиціях. Підприємства можуть впроваджувати інновації та адаптувати продукти для задоволення потреб ринку, визначаючи пріоритетність функцій на основі переваг клієнтів, перебуваючи попереду конкурентів.

Аналіз думок допомагає керувати репутацією бренду, відстежуючи настрої на платформах соціальних мереж, реагуючи на коментарі та негативні відгуки. Захист бренду може допомогти запобігти потенційній кризі компанії, що перебуває в занепаді або на піку популярності.

В основі будь-якого бізнесу лежатиме задоволеність клієнтів, і аналіз думок має в цьому завданні найважливішу роль. Підприємства можуть

поліпшити, наприклад, свої служби підтримки, виявляючи і вирішуючи проблеми, зазначені в думках, забезпечуючи позитивний клієнтський досвід. Адже одна з вічних проблем будь-якої компанії це розробка продуктів, які знаходять відгук у споживачів. Виробник може ухвалювати зважені рішення і дозволити змінити пріоритетність функцій і адаптувати продукт на основі переваг отриманих майнінгом думок.

Підприємства щодня піддаються бомбардуванню даними, але аналіз думок дає змогу відфільтрувати шум і витягти значущі відомості. Аналізуючи тенденції різних компаній будь-якому бізнесу буде простіше триматися на плаву і бути готовими до різних результатів. Адже аналіз настроїв дає змогу прогнозувати ринок і оцінки продажів.

Вид обробки природної мови швидко набув неабиякої популярності в різних галузях, зокрема й у політиці. Аналіз настроїв, збираючи коментарі на форумах і в соціальних мережах, дає змогу відстежувати посил суспільства щодо певного закону, політики, тощо.

Несподівано нещодавно новою критично важливою сферою застосування аналізу настроїв стала Громадська безпека. Аналіз настроїв і пошук думок використовуються з метою оповіщення населення про надзвичайні події, стихійні лиха або катастрофи. Під час криз виходячи з оцінки настроїв у суспільствах для грамотного управління зусиллями з надання допомоги. Адже людське джерело для з’ясування місця розташування різних точок виявлення загроз є одним із найкращих за своєю натурою.

Здатність автоматично обробляти і розуміти величезні обсяги текстового контенту, висловлюваного в Інтернеті, робить аналіз настроїв та розбір думок потужним інтрументом і в соціальних науках, охороні здоров’я, розробці програмного забезпечення, розвагах, і в багатьох інших областях.

1.3 Переслідувані цілі

Головна амбіція цього дослідження – концептуалізація, розробка і створення надійних аналітичних інструментів. Передбачається, що в разі об’єднання система буде здатна оцінювати багатогранний вплив суспільних настроїв, виражених на різних онлайн-платформах, на певні ключові галузі та глобальні події, котрі відбуваються. Це першорядне завдання підкріплюється кількома найважливішими підзавданнями, які визначатимуть процес дослідження.

Найважливіша вторинна частина дослідження передбачає глибоке і ретельне вивчення передових методологій для точної ідентифікації і тонкого розуміння колективних точок зору, переконань і тонких емоційних відтінків, закладених у величезних потоках текстових даних. Необхідно вийти за рамки спрощених класифікацій «позитивний-негативний-нейтральний» і прагне охопити повніший спектр настроїв, їхню інтенсивність і глибинні семантичні конотації, що роблять внесок у їхнє загальне значення. Це зумовлює необхідність вивчення методів, що дають змогу ефективно визначати не тільки те, що говориться, а й те, як це говориться, з яким ступенем переконаності або емоційного напруження. Надаючи інтерпретовану й актуальну інформацію, можна буде подолати розрив між аналітичними даними та прийняттям складних стратегічних рішень.

Додатковим, але важливим завданням є вивчення і потенційна потенційна інтеграція методів оцінки упередженості джерела ЗМІ. Увага приділяється тому, що вплив громадської думки може суттєво змінюватися залежно від сприйняття упередженості або схильності джерела інформації або самого контенту. Тому включення рівня аналізу, що дозволяє ідентифікувати або оцінити ймовірність певної упередженості або намірів, вважається важливим кроком до забезпечення більш цілісної та надійної оцінки суспільних настроїв.

1.4 Охоплення та обмеження

Для досягнення поставлених цілей дослідження проводитиметься в певних рамках, фокусуючи свій аналітичний погляд на конкретних типах даних і певному наборі обчислювальних методів. Основна увага буде приділена текстовій інформації, що може бути зібрана з різноманітних цифрових джерел. У їх числі маються на увазі загальнодоступні новинні статті з відомих і доступних ЗМІ, а також користувацький контент із впливових платформ соціальних мереж. Критерії відбору джерел даних можуть визначатися їхньою значущістю для різних ключових галузей, як-от фінанси або політика, які, як відомо, чутливі до суспільних дискусій.

Методологічним підґрунтям цього дослідження стане комплексне вивчення, впровадження та порівняльна оцінка різних методів оброблення природної мови та машинного навчання. Починаючи з вивчення усталених, або «класичних», підходів до аналізу настроїв і пошуку думок. Ці методи, що часто спираються на лексичні ресурси і традиційні алгоритми класифікації, слугуватимуть базою і дадуть основоположні знання про побудову ознак і статистичний аналіз тексту.

Методологічним підґрунтям цього дослідження стане комплексне вивчення, впровадження та порівняльна оцінка різних методів оброблення природної мови та машинного навчання. Починаючи з вивчення усталених, або «класичних», підходів до аналізу настроїв і пошуку думок. Ці методи, що часто спираються на лексичні ресурси і традиційні алгоритми класифікації, слугуватимуть базою і дадуть основоположні знання про побудову ознак і статистичний аналіз тексту.

Однак значна увага буде приділена вивченню і застосуванню передових нейромережевих архітектур. Сюди входить ціла низка моделей, включно з RNN та їхніми складнішими різновидами (LSTM, GRU), що добре підходять для опрацювання послідовних текстових даних або дають змогу визначати важливі локальні патерни та n-грами в тексті, як CNN, які

вказують на настрій. Важливе значення має вивчення моделей на основі трансформерів, таких як BERT і його похідні з потужними механізмами привернення уваги та здатності вловлювати глибоке контекстуальне розуміння.

Прагнучи високої точності та тонкого розуміння, проєкт не ставить перед собою завдання досягти ідеальної інтерпретації на людському рівні в усіх можливих контекстах. Початковий фокус буде переважно на англомовних текстах через ширшу доступність ресурсів і фундаментальних досліджень, хоча потенціал майбутнього багатомовного розширення розглядатиметься як перспектива. Етичні наслідки збору та аналізу даних, особливо щодо конфіденційності та потенційної упередженості алгоритмів, також будуть постійно розглядатися в процесі дослідження, хоча повна етична експертиза виходить за рамки безпосередніх технічних рамок, описаних у дослідженні.

1.5 Історичний контекст та ключові терміни

Хоча обчислювальний аналіз текстів сягає корінням у далеке минуле (наприклад, контент-аналіз у середині XX століття для аналізу пропаганди), конкретна галузь аналізу настроїв і дослідження поглядів почала формуватися й набирати значних обертів приблизно на рубежі тисячоліть. Перші роботи в царині комп’ютерної лінгвістики та NLP у 1990-х роках були присвячені вивченню аспектів з інтерпретації метафор, почуттів, класифікації суб’єктивності, виокремлення точок зору та афектів [3].

Сплеск інтересу з початку 2000-х років тісно пов’язаний з розвитком Інтернету, зокрема технологій Веб 2.0, таких як блоги, форуми та соціальні мережі. Це створило безпрецедентний обсяг загальнодоступних текстових даних, багатих на думки, що забезпечило як мотивацію (розуміння цих

даних), так і ресурс (дані для навчання і тестування моделей) для швидкого розвитку галузі.

Термін «Sentiment Analysis», вперше з'явився в роботі Насукави та Йі [4], а термін «Opinion Mining» – в роботі Дейва, Лоуренса та Пеннока [5]. Дослідницькі зусилля швидко вийшли за межі простої класифікації полярності і почали вирішувати більш складні завдання. Еволюція бачила перехід від початкових лексичних та евристичних методів до підходів машинного навчання (таких як наївний Байєс, машини опорних векторів), підживлюваних еталонними наборами даних, такими як рецензії на фільми, а нещодавно – до складних моделей глибокого навчання (CNN, LSTM, Transformers), здатних вловлювати складні лінгвістичні патерни та контекст.

Якщо заглиблюватися в термінологію, у нашому випадку ми обговорюємо обчислювальне вивчення думок, настроїв, оцінок, поглядів та емоцій людей стосовно об'єктів та їхніх атрибутів.

Об'єкт може мати на увазі продукт, тему, іншу людину або навіть послугу, те, що було прокоментовано. Об'єкт сам по собі може мати набір компонентів і атрибутів, які, своєю чергою, розбиваються на підкомпоненти і так далі.

Думка – це судження, точка зору, що охоплює оцінки, ставлення тощо, що складатиметься з виразів, які мають на увазі певні почуття. Думка також вважається твердженням, яке не є остаточним, на відміну від фактів, які є істинними твердженнями.

Настрій вирізняється тим, що більше відноситься до почуття, вираженого емоційним тоном. Часто класифікується за полярністю (позитивна, негативна, нейтральна).

Полярність слова або думки – це спрямованість висловленої думки, яку також називають семантичною орієнтацією, вказує на те, в якому напрямку слово або думка відхиляється від норми для своєї семантичної

групи. Іншими словами, полярність слів або фраз безпосередньо показує думку автора тексту [6].

Носій думки – це особа або джерело, яке висловлює думку, в деяких випадках завдання вилучення носія думки корисне для того, щоб знати, хто має таку саму позицію, які питання хвилюють конкретну особу, і чи існують різні думки від деяких конкретних осіб [7].

У термінах можна так само зустріти виявлення суб'єктивності та орієнтації тексту [8], обидва мають вирішальне значення для точного вилучення думок. Основна проблема полягає у фундаментальній різниці між суб'єктивною та об'єктивною мовою. Виявлення суб'єктивності може заощадити аналіз маси нерелевантних даних, які містять факти і є лише об'єктивними. Такі тексти не можуть мати орієнтацію або полярність і спроба застосувати аналіз настроїв до об'єктивного тексту буде схожа на пошук сигналу там, де його немає, що призведе до неточних або оманливих висновків.

1.6 Лексикони та раннє машинне навчання

Напередодні широкого впровадження складних нейронних мереж у сфері аналізу настроїв були розроблені базові методи, які залишаються актуальними та інформативними. Ці «традиційні» підходи в основному поділяються на дві основні категорії: методи, що спираються на заздалегідь визначені лексикони настроїв, і методи, що використовують класичні алгоритми машинного навчання з ретельно розробленими функціями.

1.6.1 Підходи на основі використання словників

Маючи підготовлений набір, найінтуїтивніший спосіб визначити настрій тексту – це вивчити емоційну полярність слів, що містяться в ньому. Підходи, засновані на лексиці, формалізують цю інтуїцію, використовуючи

словники, або лексикони, в яких слова (а іноді й фрази) попередньо маркуються оцінками або категоріями настрою (позитивний, негативний, нейтральний). Загальну оцінку настрою документа або речення зазвичай визначають шляхом підсумовування оцінок настрою слів, що його складають.

Основний процес включає в себе:

– токенізація, що являє собою розбиття вхідного тексту на окремі слова або значущі одиниці (лексеми);

– для кожної лексеми перевіряється, чи існує вона в лексиконі настроїв, і витягується пов’язаний із нею бал(и) полярності;

– об’єднання оцінок окремих слів для отримання загальної оцінки настрою тексту. Проста агрегація може включати підсумовування позитивних і негативних оцінок окремо та їх порівняння або обчислення середнього бала.

Складніші системи, засновані на лексиконі, часто включають обробку заперечення, інтенсифікацію або модифікацію, і перемикачі валентності для обробки нюансів [9].

Визначення слів із запереченням (наприклад, «не», «ніколи», «ні») і зміна полярності прилеглих слів настрою (наприклад, «не дуже добре» стає негативним, незважаючи на позитивну оцінку «добре»), інтенсифікаторів (наприклад, «дуже», «надзвичайно», «неймовірно»), що підсилюють силу настрою, та тих, що її зменшують (наприклад, «трохи», «дещо», «ледве»), що призводить до відповідного коригування оцінок.

Крім заперечення, інші слова або фрази можуть змінювати настрій, що вимагає дотримання особливих правил. Точне і більш специфічне опрацювання прилеглих слів може виявитися дуже складним завданням.

Ефективність зазначених методів значною мірою залежить від якості та охоплення базового лексикону. Для створення ефективного словника використовують дві основні стратегії.

Методи, засновані на словниках, зазвичай починають із невеликого набору визначених вручну «затравочних» слів із відомою позитивною та негативною полярністю. Потім цей набір автоматично розширюється за рахунок використання структури наявних словників або лексичних ресурсів. Процес часто включає ітеративне додавання синонімів (припускається, що вони мають однакову полярність) і антонімів (мають протилежну полярність), які на даний час знаходяться в лексиконі. Прикладом такого словника послужить WordNet. Він має велику лексичну базу даних англійської мови, в якій іменники, дієслова, прикметники та прислівники згруповані в набори когнітивних синонімів, які називаються «синсетами». Методи, засновані на словниках, використовують ці відносини синонімії та антонімії. Ще один видатний приклад, побудований на основі WordNet. Де замість бінарної мітки «позитивний/негативний» SentiWordNet присвоює три числові оцінки кожному синсету WordNet: оцінку позитивності, оцінку негативності та оцінку об’єктивності (нейтральності). Ці три бали для будь-якого даного синсету дорівнюють 1,0. Наприклад, синсет може бути 0,75 позитивним, 0,0 негативним і 0,25 об’єктивним. Це дає змогу проводити більш тонкий аналіз, визнаючи, що слова можуть мати змішані або слабкі почуття. SentiWordNet 3.0, остання версія на момент проведення дослідження, використовує WordNet 3.0 і застосовує такі методи, як випадкове переміщення по графу, щоб переконатися, що семантична спорідненість переходить у спорідненість за настроєм.

Методи, що ґрунтуються на корпусах, також часто починають з вихідних слів, але визначають орієнтацію настрою, аналізуючи шаблони використання слів у великих текстових корпораціях. Вони спрямовані на виявлення специфічних для даної області слів настрою або уточнення оцінок полярності на основі фактичного використання.

Одна з методик шукає в корпусі текстів синтаксичні патерни у вигляді прикметників, які пов’язані сполучниками на кшталт «і» (які передбачають

схожу полярність) або «але» (протилежну полярність). Потім можна використовувати алгоритми кластеризації графів, щоб згрупувати слова в позитивні та негативні набори.

Інший підхід розраховує статистичну асоціацію (PMI) між словом-кандидатом і відомими позитивними («чудовими») і негативними («поганими») вихідними словами. Вищий зв'язок із «чудовим» словом передбачає позитивну полярність, і навпаки. Різниця в оцінках PMI щодо позитивних і негативних семплів дає оцінку семантичної орієнтації слова-кандидата. Для оцінки ймовірності збігу часто доводиться запитувати великі веб-корпорації через пошукові системи.

Примітні лексичні ресурси аналогічні SentiWordNet:

– LIWC (Linguistic Inquiry and Word Count): Програма для аналізу тексту, що постачається з ретельно перевіреними словниками, що класифікують слова за психологічними параметрами, включно з позитивними та негативними емоціями. Не будучи виключно лексиконом почуттів, його категорії емоцій часто використовуються в дослідженнях з аналізу почуттів і послужили основою для інших інструментів;

– VADER (Valence Aware Dictionary and sEntiment Reasoner): Лексикон та інструмент, заснований на правилах, спеціально оптимізований для почуттів, виражених у мікроблогах і соціальних мережах. Він включає поширений інтернет-сленг, смайлики і враховує модифікатори інтенсивності та заперечення;

– лексикон суб'єктивності MPQA: Створений вручну ресурс, що містить тисячі слів, анотованих за суб'єктивністю (сильний/слабкий) і попередньою полярністю (позитивний, негативний, нейтральний, обидва).

Саме простота і дешевизна є сильними сторонами підходів заснованих на використанні лексикону. Відсутність необхідності в навчальних даних для моделі машинного навчання є важливою перевагою. З огляду на легку інтерпретацію, що дає змогу простежити оцінку до конкретних слів або їхніх записів у лексиконі, і швидкість пророблених комп'ютером обчислень,

робить ці підходи дещо ефективнішими порівняно зі складними моделями машинного навчання.

Однак, їхній головний недолік – контекстна сліпота; їм важко впоратися з контекстно-залежними значеннями слів (наприклад, «sick» як «хворий» порівняно з «sick» як «крутий»), сарказмом, іронією та нюансами виразів, у яких настрій не виражений ключовими словами. Ще одну проблему становить специфіка домену, оскільки лексика, створена для загального використання, може погано працювати в спеціалізованих галузях, як-от фінанси або медицина, де існують унікальні жаргони та смислові конвенції. Розробка лексики, специфічної для конкретної галузі, розвиток мови, сленгів і термінів вимагає дедалі більше й більше часу роботи для опрацювання та підтримання актуальності підходу, що ґрунтується на лексиці. Нарешті, такі підходи часто не справляються зі складними структурами речень, де настрій виражається або змінюється дуже тонко, що робить просте об’єднання слів недостатнім методом для точного аналізу настрою.

1.6.2 Традиційні підходи з машинним навчанням

Усвідомивши обмеженість фіксованих лексиконів, дослідники звернулися до машинного навчання з супервізією. Цей підхід розглядає аналіз настрою як завдання класифікації тексту: задавши текст, передбачити його клас настрою. Замість того щоб покладатися на заздалегідь визначені оцінки, моделі ML вивчають шаблони, пов’язані з різними класами настрою, на основі позначених прикладів.

Технологія контрольованого навчання насамперед має на увазі збір набору даних, у якому кожен екземпляр тексту (наприклад, відгук) зіставлено з відомою міткою настрою. Поширеним джерелом є онлайн-огляди з оцінками зірок (1–5 зірок), де оцінки зіставляються з мітками (наприклад, 4–5 зірок → позитивні, 1–2 зірки → негативні,

3 зірки → нейтральні або відкинуті). Обов'язковим є також очищення зібраних текстових даних. Загальні кроки включають перетворення на малі літери, видалення розділових знаків і спеціальних символів, обробку HTML-тегів і потенційне видалення поширених «стоп-слів» (таких як «a», «the», «is»), які можуть не нести значної інформації про настрої. Також може бути застосовано стеммінг (скорочення слів до кореневої форми, наприклад, «running» → «run») або лематизацію (скорочення слів до їхньої словникової форми, наприклад, «ran» → «run»).

Критичний крок – розробка/видобування ознак. Сирий текст необхідно перетворити в числовий формат (вектори ознак), який можуть обробляти алгоритми. Найпоширеніший підхід (BoW), представляє текст на основі наявності або частоти слів, ігноруючи граматику та порядок слів. Для цього часто використовують уніграми (окремі слова) або n-грами (послідовності з n слів, наприклад, біграми, триграми). N-грами відображають деяку локальну інформацію про порядок слів (наприклад, біграми типу «не дуже добре», триграми типу «дуже поганий фільм»). Частота термінів (TF) або зважування TF-IDF (Term Frequency-Inverse Document Frequency) часто застосовують для надання більшого значення відмітним словам. Бінарна присутність (наявність/відсутність слова/n-грами) або частотна (підрахунок кількості входжень). Дослідження Pang та інші. (2002) показало, що використання присутності слів часто працює так само добре, як і частота, або навіть краще, ніж частота, для класифікації настроїв.

Прикметники та прислівники часто особливо показові для визначення настрою. Використовується (POS), тобто граматична роль слів (прикметник, прислівник, іменник, дієслово) як ознаки. Включається й інформація з наявних лексиконів, така як кількість позитивних/негативних слів, що трапляються в тексті, або сума/середнє значення їхніх оцінок настрою, що може бути використане як додаткові ознаки для моделі ML.

Додатково процес виявлення може включати в себе створення специфічних ознак для представлення наявності та обсягу заперечення, ідентифікацію відомих слів або фраз, що виражають думку, та обробку синтаксичних ознак, отриманих із граматичної структури речень, такі як шляхи залежності між словами, виявлені в результаті синтаксичного аналізу.

Після виявлення ознак відбувається навчання моделі. Подаються витягнуті вектори ознак і відповідних їм міток в обраний алгоритм класифікації. Алгоритм навчається функції відображення або межі прийняття рішень, щоб розрізняти класи настроїв на основі вхідних ознак. Останній етап – оцінювання ефективності навченої моделі на тестовому наборі даних (дані, які не використовували під час навчання), проробляють за допомогою стандартних показників класифікації, як-от точність, достовірність, відкликання і F1-score.

Трьома основними поширеними традиційними алгоритмами ML є Naive Bayes, Support Vector Machines і Maximum Entropy (MaxEnt) або Logistic Regression. Також але рідше застосовуються дерева рішень, k-Nearest Neighbors (k-NN) і ансамблеві методи, такі як Random Forests.

Сімейство простих імовірнісних класифікаторів, заснованих на теоремі Байєса, припускають умовну незалежність ознак від мітки класу. Незважаючи на це часто нереалістичне припущення, NB-класифікатори ефективні з погляду обчислень, добре працюють із високорозмірними даними, як-от текст, і можуть забезпечити високу базову продуктивність, особливо за обмеженої кількості навчальних даних.

SVM під час класифікації націлені на пошук оптимальної гіперплощини, яка найкращим чином розділяє точки даних, що належать до різних класів у високорозмірному просторі ознак. Вони особливо ефективні у високорозмірних просторах (що характерно для текстових ознак) і можуть моделювати нелінійні залежності за допомогою функцій ядра. SVM часто досягали найвищих результатів у класифікації текстів до епохи глибокого

навчання і часто перевершували Naive Bayes за наявності достатньої кількості навчальних даних.

ML підходи часто досягають більш високої точності і є більш адаптивними, ніж методи, засновані винятково на лексиконі, особливо під час навчання на достатній кількості позначених даних, які стосуються конкретної галузі. Вони можуть вивчати складні патерни і взаємодії ознак безпосередньо з даних, виходячи за рамки простих списків слів.

Як і в методах зі словниками, успіх моделей значною мірою залежить від якості створених вручну ознак. Розроблення ефективних функцій вимагає значних зусиль, досвіду та експериментів. І незважаючи на успіх в уловлюванні деякого локального контексту, ці моделі не справляються з глибоким семантичним розумінням, далекими залежностями між реченнями, сарказмом і неявними настроями, що передаються через складні мовні структури.

Усе одно лексичні методи і традиційні методи машинного навчання є основою для аналізу настроїв. Методи на основі лексики вирізняються простотою і зручністю інтерпретації, але їм не вистачає контекстного розуміння. ML методи підвищують точність завдяки навчанню на основі різноманітних даних, але також спираються на марковані дані та ресурсовитратне ручне розроблення ознак, усе ще не справляючись із глибшими складнощами мови. Ці підходи підкреслюють важливість лексичних підказок і закономірностей, обумовлених даними. Але властиві їм обмеження в кінцевому рахунку були стимулом для переходу до здатних створювати ширші представлення з тексту – методів глибокого навчання.

1.7 Методи глибокого навчання

Нейронні мережі з їхньою здатністю навчатися ієрархічних уявлень на основі даних здійснили революцію зі зміною парадигми в NLP і значно просунули вперед сучасні методи аналізу смислів. На відміну від

традиційного машинного навчання, яке вимагає від розробників явного визначення характеристик (наприклад, Bag-of-Words або TF-IDF), комплексні моделі націлені на автоматичне вивчення відповідних характеристик у процесі навчання. Починаючи, як правило, зі щільних векторних уявлень слів, званих вкрапленнями слів (наприклад, Word2Vec, GloVe, FastText або вкраплення, отримані в процесі навчання), моделі глибокого навчання пропускають вхідні дані через кілька рівнів нелінійних перетворень, що дає їм змогу поступово створювати більш абстрактні та складні уявлення смислу і настрою тексту.

1.7.1 Рекурентні нейронні мережі та їхні різновиди

Мережі RNN, зображені рисунку 1.1, ідеально підходять для роботи з послідовними даними, так і у випадку з текстом. Оскільки вони обробляють вхідні дані по одному елементу (наприклад, одному слову) за раз, зберігаючи внутрішню «пам’ять» або прихований стан, який підсумовує інформацію побачену раніше.

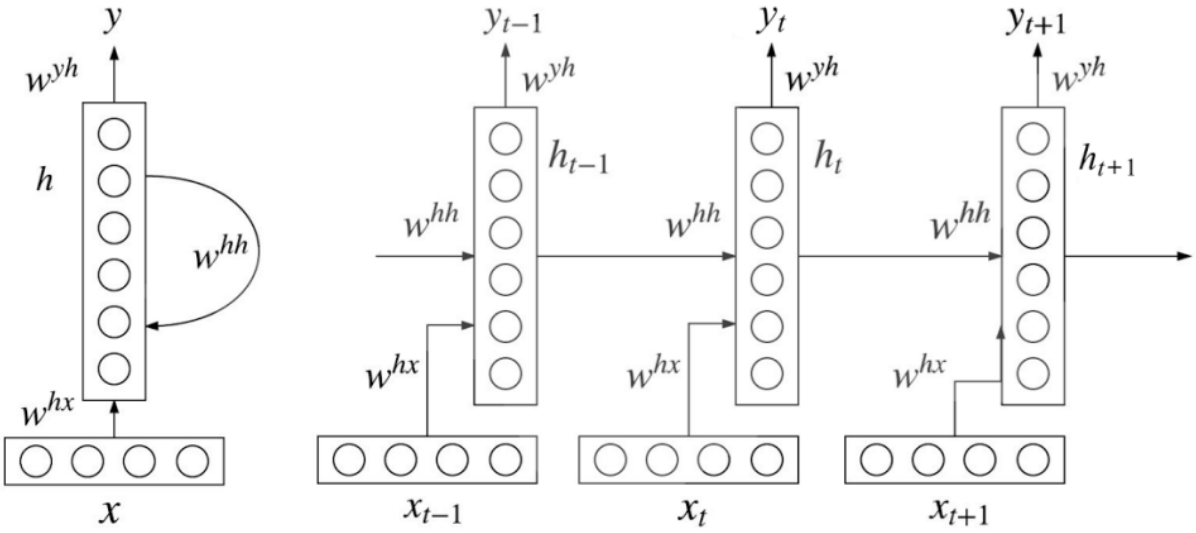

Рисунок 1.1 – Рекурентна нейронна мережа (RNN)

На кожному кроці RNN приймає на вхід подання поточного слова (зазвичай вектор вкладення) і прихований стан, отриманий на попередньому кроці. Він обчислює новий прихований стан і, потенційно, вихід (передбачення настрою на цей момент). Під час агрегації часто використовують усереднення або максимальний пул, а для шару класифікації – шар softmax. Остаточний прихований стан після опрацювання всієї послідовності може бути використано як подання всього тексту для класифікації.

Прості RNN страждають від проблеми «зникаючого градієнта», що ускладнює їхнє навчання залежностей між словами, які перебувають на великій відстані один від одного в довгій послідовності. Вплив перших слів має тенденцію зникати в міру того, як мережа обробляє більше слів.

Спеціально для вирішення цієї проблеми було розроблено LSTM [10]. Вони використовують систему «воріт», зображена на рисунку 1.2, для управління потоком інформації, що дає змогу вибірково запам’ятовувати релевантну інформацію впродовж тривалого часу і забувати нерелевантні деталі.

Gated Recurrent Unit є дещо простішим варіантом LSTM, який поєднує ворота забування (forget gate) і вхідні ворота (input gate) у «Update Gate» і об’єднує стан комірки і прихований стан. GRU часто досягають порівнянної продуктивності, але з меншою кількістю параметрів і потенційно швидшим навчанням.

Оптимальними для осягнення тексту будуть саме двонаправлені RNN/LSTMs/GRUs. Вони покращують розуміння контексту та нюансів, обробляючи вхідну послідовність в обох напрямках за допомогою двох окремих прихованих шарів. Виходи обох напрямків потім об’єднуються на кожному часовому кроці. Це дає змогу включити в представлення речі інформацію з попереднього і наступного контексту, і такий підхід часто призводить до поліпшення продуктивності.

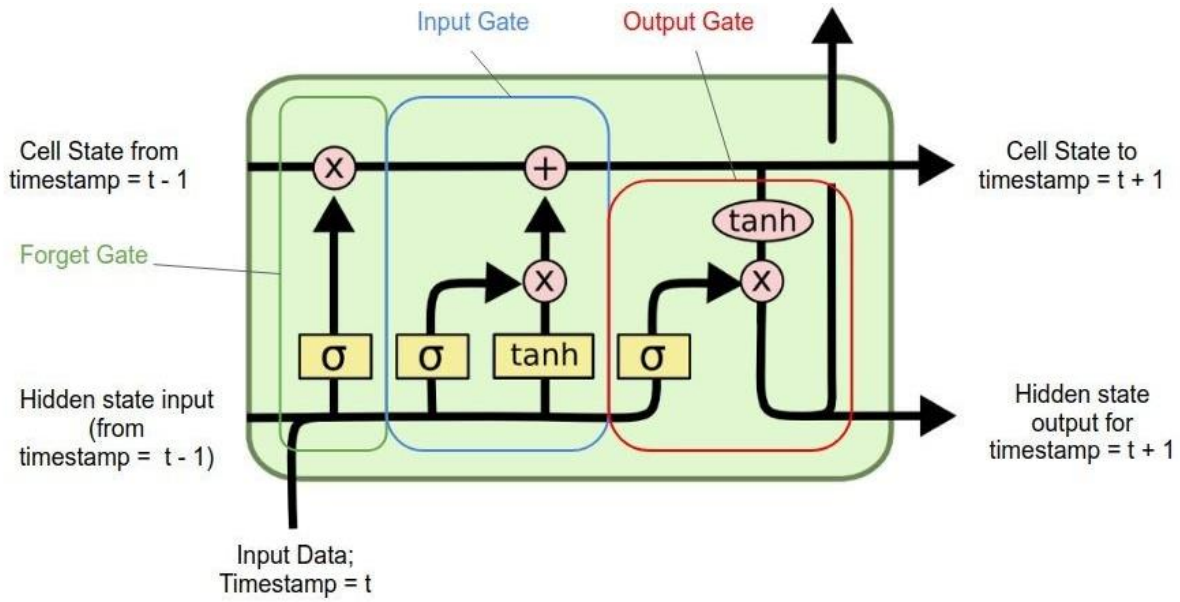

Рисунок 1.2 – Система «воріт» LSTM

Основним недоліком виступає фактор, що послідовне опрацювання обмежує можливості розпаралелювання і робить навчання більш повільним, ніж CNN або трансформери. Особливо на дуже довгих текстах можуть зазнавати труднощів з дуже складними залежностями порівняно з механізмами уваги.

1.7.2 Конволюційні нейронні мережі

Хоча CNN набули популярності завдяки переважанню в комп'ютерному зорі, їх також успішно застосовують у завданнях NLP, включно з аналізом настрою, розглядаючи текст як одновимірну послідовність вкраплень слів [11]. CNN ефективно визначають ключові фрази або n-грами в тексті, а здатність уловлювати локальні патерни робить їх стійкими до змін у розташуванні слів.

Замість згортки за пікселями CNN для роботи з текстом зазвичай застосовують згорткові фільтри (з різними розмірами вікна, наприклад, ті, що захоплюють 2, 3 або 4 слова за раз) за послідовністю вкраплень слів.

Кожен фільтр діє як детектор n-грам, навчаючись визначати важливі локальні патерни або особливості, що стосуються почуттів (рисунок 1.3). демонструє видобуття ознаки. Для кожного розміру вікна можна використовувати кілька фільтрів, щоб вловити різні типи патернів. Для вибору найбільш значущих ознак з усієї послідовності, часто використовуються операції максимального об'єднання. Цей вектор потім подається на шар із повним підключенням для класифікації.

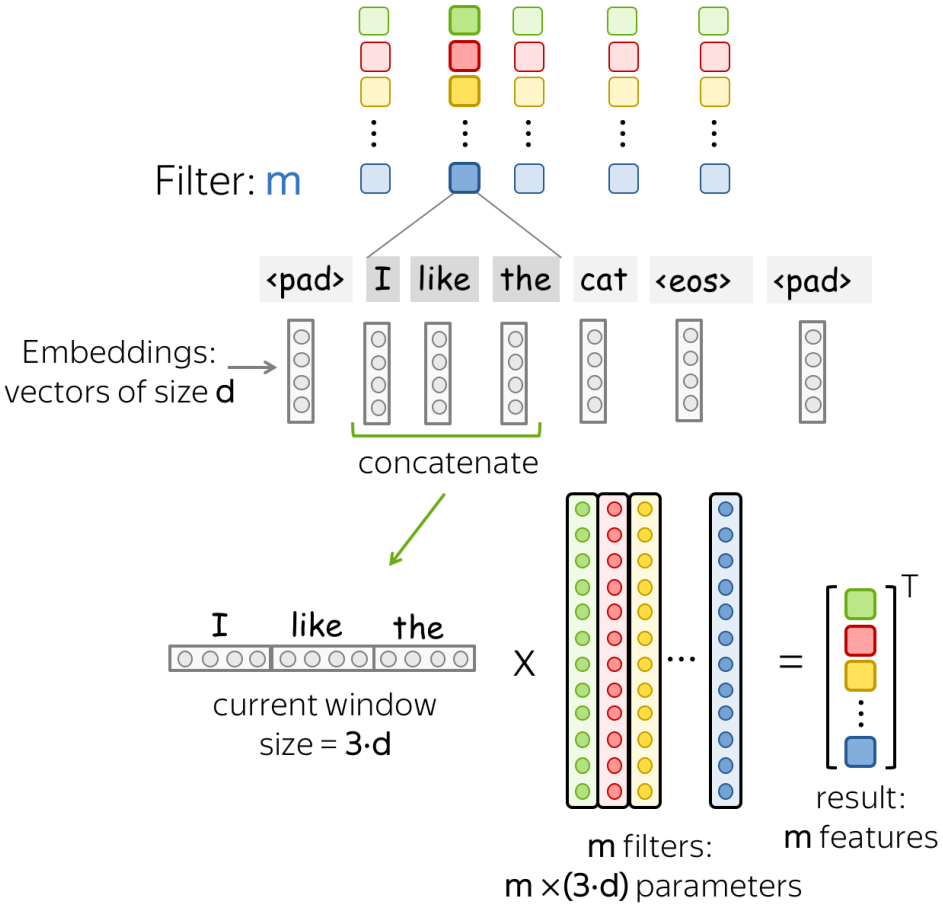

Рисунок 1.3 – Згортка CNN видобуває ознаку

Стандартні CNN за своєю природою не моделюють далекі залежності або послідовний порядок так само ефективно, як RNN. Вони в основному вловлюють локальну контекстну інформацію в межах розміру вікна фільтрації. Тому, часто комбінуються в моделі типу CNN-LSTM, щоб використовувати сильні сторони обох.

1.7.3 Трансформаторні моделі та механізм уваги

Трансформери, представлені в основоположній статті «Attention Is All You Need» [12], являють собою значний прорив. Вони стали домінуючою архітектурою для сучасного NLP. Основним нововведенням є механізм самоуваги (часто багатоголова самоувага). На відміну від RNN, які покладаються на послідовне опрацювання і приховані стани для передавання інформації, самоусвідомлення дає змогу моделі безпосередньо зважувати вплив усіх інших слів у вхідній послідовності під час обчислення уявлення для одного слова. Це дає змогу ефективно вловлювати далекі залежності, незалежно від відстані. Вона розраховує «бали уваги» між парами слів, показуючи, наскільки одне слово релевантне іншому в даному контексті. Таким чином модель надійніше розуміє складні відносини, включно із запереченням і модифікацією, і може визначати значення слів на основі повного контексту (наприклад, «bank» (берег) річки порівняно з фінансовим «bank»).

Трансформатори зазвичай використовують багатошарові багатоголові мережі із самонавіюванням і зворотним зв’язком у корені своєї архітектури. У них повністю відсутня рекурентність, що дає змогу значно розпаралелити й уксорити процес навчання.

Основоположною впливовою моделлю стала BERT, яка використовує частину архітектури шифратора. Вона попередньо навчена на величезних обсягах тексту з використанням таких завдань, як масковане моделювання мови та передбачення наступного речення.

Концепція маскування (MLM) має на увазі випадкове маскування деяких вхідних слів і навчання моделі передбаченню вихідних маскованих слів на основі навколишнього двонаправленого контексту [13].

Передбачення наступного речення – це наступне з двох основних завдань, що використовуються для попереднього навчання вихідної моделі BERT. Як показано на рисунку 1.4, BERT даються два речення, A і B.

Завдання полягає в тому, щоб передбачити, чи дійсно B йде одразу після A в тому самому документі.

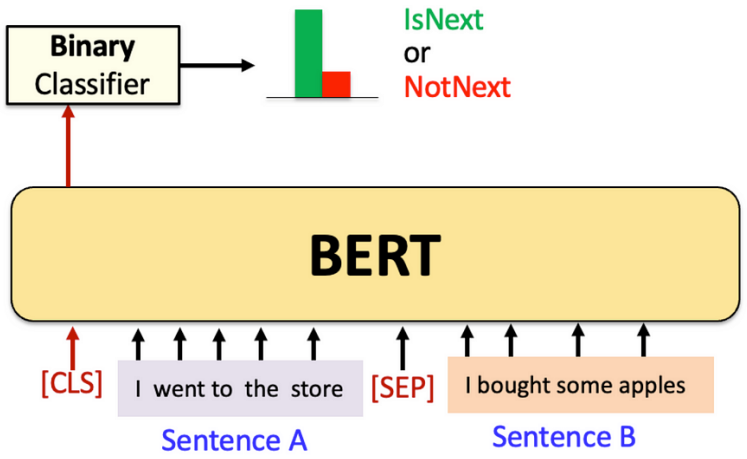

Рисунок 1.4 – Прогнозування наступного речення (BERT)

Таке попереднє навчання наділяє BERT глибоким розумінням структури та семантики мови. Для аналізу настроїв попередньо навчену модель BERT зазвичай піддають тонкому налаштуванню: поверх вихідних даних BERT додають шар класифікації (зазвичай це подання спеціальної лексеми [CLS]), і всю модель навчають упродовж невеликої кількості епох на конкретному наборі даних настроїв із мітками.

З'явилося безліч варіантів, що оптимізують попереднє навчання або архітектуру:

– RoBERTa: оптимізує стратегію попереднього навчання BERT (наприклад, динамічне маскування, більше даних, великі партії) для підвищення продуктивності;

– DistilBERT: використовує дистиляцію знань для створення компактнішої та швидшої версії BERT, зберігаючи водночас більшу частину її точності, що підходить для середовищ з обмеженими ресурсами;

– ALBERT: використовує методи зменшення параметрів для полегшення моделей;

– XLM-RoBERTa: попередньо навчений кількома мовами для розв’язання міжмовних задач;

– Сімейство GPT: моделі на основі декодера, які чудово справляються з генерацією тексту. Хоча вони зазвичай не налаштовуються так тонко, як BERT для класифікації, великі моделі GPT можуть виконувати аналіз настрою за допомогою нульових або декількох знімків підказки (даючи моделі інструкції та приклади в самій підказці).

Тонке налаштування попередньо навчених моделей-трансформерів, таких як BERT або RoBERTa, стало стандартним підходом для досягнення найвищих результатів у бенчмарках аналізу настрою. Вони забезпечують потужні контекстуалізовані уявлення, які ефективно відображають нюанси настрою. Вони здатні автомотично вивчати релевантні ознаки безпосередньо з даних, усуваючи необхідність у трудомісткому і потенційно неоптимальному ручному розробленні ознак, яке постійно потрібне в традиційному ML. Такі архітектури, як і LSTM, чудово справляються з уловлюванням складної контекстуальної інформації, порядку слів, далеких залежностей, заперечення, сарказму (певною мірою) і полісемії.

Хоча і трансформери вимагають попереднього навчання, але завдяки цьому на масивних немаркованих текстових базах моделі типу BERT набувають багатих лінгвістичних знань. Ці знання можуть бути ефективно перенесені на конкретні задачі, як-от аналіз настрою, шляхом тонкого налаштування, потребуючи значно менше помічених даних, порівняно з навчанням глибоких моделей з нуля.

1.7.4 Обмеження методів глибокого навчання

Хоча трансферне навчання знижує потребу в маркованих даних порівняно з навчанням у традиційних підходах, для оптимального налаштування все одно потрібна розумна кількість високоякісних,

специфічних для конкретного завдання маркованих даних. Навчання великих моделей з нуля може потребувати дуже багато даних. Зі зростанням моделі для глибокого навчання, особливо трансформерів, вимагають великих обчислювальних витрат, потужного обладнання (GPU/TPU) і значного часу. Продуктивність так само може бути чутливою до вибору архітектури (кількість шарів, прихованих блоків тощо) і гіперпараметрів навчання (швидкість навчання, розмір партії, алгоритм оптимізації), все вимагає ретельного налаштування й експериментів.

Проблеми інтерпретованості та упередженості. Проблема «чорної скриньки», або чому складна модель глибокого навчання зробила той чи інший прогноз настрою, часто буває не вирішена. Їхня внутрішня робота менш прозора, ніж у простіших моделей, таких як Naive Bayes або методи на основі лексикону. Відсутність можливості інтерпретації стає перешкодою в деяких завданнях, що вимагають пояснень або звітності. Такі методи, як візуалізація уваги, дають деяке уявлення, але не є повним рішенням. Тим паче, моделі глибокого навчання, навчені на великих, необроблених наборах даних (наприклад, веб-текстів), можуть засвоїти і закріпити суспільні забобони, присутні в даних, пов’язані зі статтю, расою, релігією тощо. Це може призвести до несправедливих або спотворених прогнозів настрою, якщо не боротися з цим за допомогою таких методів, як очищення даних або обмеження моделі.

Незважаючи на значний прогрес, навіть прогресивні моделі все ще не можуть впоратися з тонкими формами сарказму, іронії, культурними відсиланнями і здоровим глуздом, необхідним для ідеальної інтерпретації настроїв.

1.8 Аспектно-орієнтований аналіз настроїв

Хоча ранні методи аналізу настрою були спрямовані на класифікацію цілих документів або речень як позитивних, негативних або нейтральних,

цього часто виявлялося недостатньо. У захваленому відгуку про смартфон може міститися критика з приводу часу автономної роботи, а в негативному відгуку про готель – похвала доброзичливому персоналу. Щоб вловити ці нюанси, у цій галузі було розроблено більш детальний підхід.

ABSA, також званий «пошуком думок на основі ознак», за термінологією Ху та Лю [14], глибший, ніж аналіз на рівні документів або речень. Його мета – виявити конкретні аспекти (особливості, атрибути або теми), обговорювані в тексті, і визначити настрій, виражений щодо кожного з цих аспектів. В основоположних працях мету часто формулюють як витяг повної структури думок, включно з об'єктом, його аспектом, вираженим настроєм, носієм думки і часом. Така детальна розбивка дає набагато багатше і практичніше розуміння думок.

Досягнення такого рівня деталізації включає в себе кілька взаємопов'язаних завдань, але два з них виділяються як основні, яким приділяється значна увага в дослідженнях:

– вилучення аспектів: визначення конкретних термінів або фраз у тексті, які представляють оцінювані аспекти. Наприклад, у фразі «Яскравість екрана приголомшлива, але клавіатура здається дешевою» аспектами є «яскравість екрана» і «клавіатура»;

– аспектна класифікація настрою: визначення полярності (позитивної, негативної або нейтральної) думки, вираженої щодо кожного ідентифікованого аспекту. У наведеному вище прикладі настрій стосовно «яскравості екрана» є позитивним, а стосовно «клавіатури» – негативним.

Хоча іноді аспекти, що цікавлять, можуть бути заздалегідь визначені користувачем (наприклад, аналіз думок щодо «часу автономної роботи» та «камери» для смартфонів), реальне завдання часто полягає в автоматичному виявленні цих аспектів із неструктурованого тексту.

Виділення аспектів, іноді зване виділенням термінів аспектів або виділенням цілей думок, є, ймовірно, одним із найскладніших етапів у конвеєрі ABSA. Він вимагає складних методів обробки природної мови для

точного визначення характеристик або об’єктів, які є предметом думки. Дослідники вивчали різні стратегії, які загалом поділяються на методи без нагляду, з наглядом і напівнаглядом, причому підходи глибокого навчання посідають дедалі більшу частку в цих категоріях.

1.8.1 Пошук закономірностей без попередніх позначок

Непідконтрольні методи намагаються визначити аспекти, не спираючись на попередньо позначені навчальні дані, часто шляхом пошуку статистичних закономірностей або лінгвістичних підказок. Одними з першопрохідців у цій галузі стали Ху і Лю [14], які використовували метод пошуку асоціативних правил. У їхньому підході іменники та словосполучення розглядалися як потенційні аспекти. Проаналізувавши, які іменники часто зустрічаються в реченнях, вони змогли визначити потенційні аспекти. Потім для уточнення списку використовували стратегії відсікання, а аспекти, які трапляються нечасто, іноді виявляли шляхом пошуку іменників, що примикають до відомих слів думки.

Попеску та Етціоні [15] прагнули підвищити точність таких частотних методів. Вони визнали, що не всі іменники, які часто зустрічаються, обов’язково є характеристиками продукту. Їхня методика передбачає обчислення показника точкової взаємної інформації між фразою-кандидатом і набором «дискримінаторів меронімії», характерних для класу товару (наприклад, «сканер має», «зі сканера»). Низький показник PMI вказував на те, що фраза-кандидат з меншою ймовірністю є істинним аспектом. Незважаючи на ефективність цього методу, його залежність від великих веб-запитів являє собою практичну проблему. Серед інших ранніх робіт, виконаних без контролю, можна назвати роботу Йі та ін. [16], які розробили алгоритми, засновані на мовних моделях і коефіцієнтах правдоподібності.

Зовсім недавно глибоке навчання відкрило нові можливості для виділення аспектів без контролю. Хе та ін. [17] запропонували модель на основі уваги, натхненну автоенкодерами, використовуючи механізм уваги для автоматичного фокусування на словах, що мають відношення до аспектів, і зменшення ваги нерелевантних контекстних слів, тим самим неявно визначаючи аспекти без прямого контролю.

1.8.2 Навчання на маркованих прикладах

Методи, засновані на спостереженні, використовують моделі машинного навчання, навчені на наборах даних, де аспекти були анотовані вручну. Ці підходи часто розглядають витяг аспектів як завдання маркування послідовності, подібно до розпізнавання іменованих сутностей (NER).

Популярністю користуються умовні випадкові поля. Якоб та І. Гуревич [18] продемонстрували ефективність CRF, використовуючи такі ознаки, як теги частини мови (POS) і шляхи залежностей. Чжан та ін. [19] розширили модель CRF, замінивши традиційні дискретні ознаки на безперервні вкраплення слів і включивши шар нейронної мережі, що дало змогу моделі навчатися багатшим уявленням. Також застосовували машини опорних векторів (SVM). Кесслер і Ніколов [20] навчили SVM-класифікатор спеціально для зв’язку виразів думки з їхніми правильними цільовими аспектами в реченні, використовуючи ознаки, отримані із синтаксичних і семантичних відносин. Дейв та ін. [21] також використовували SVM для класифікації відгуків після значного попереднього опрацювання тексту.

Приховані марковські моделі досліджували Цзінь та ін. [22] в їхній системі «OpinionMiner», яка розробила лексикалізовану структуру HMM, що включає різні лінгвістичні характеристики. Інша точка зору розглядає вилучення аспектів як проблему розв’язання кореференції тем, навчаючи

класифікатор, щоб визначити, чи відносяться два вирази думки до одного й того самого базового аспекту.

1.8.3 Розквіт глибокого навчання у вилученні аспектів

Моделі глибокого навчання подають великі надії завдяки своїй здатності автоматично вивчати складні уявлення ознак на основі даних. Глибокі двоспрямовані LSTM використовуються для спільного вилучення сутностей думки та їхніх зв’язків. Спеціальні LSTM-установки націлені на уловлювання взаємодій між аспектом і почуттям, тоді як інтегрування RNN і CRF для спільного вилучення аспектів і термінів думки. Додатково RNN також можуть застосовуватися для міждоменного вилучення аспектів.

Механізми уваги будуються на об’єднаних багатошарових моделях CMLA. Використовуючи блоки GRU з окремими механізмами уваги для аспектів і думок, вони полегшили спільне вилучення. Увага також відіграє ключову роль у деяких підходах без контролю. Глибокі CNN підходять для класифікації слів як аспектних або неаспектних, іноді інтегруючи лінгвістичні патерни.

Ефективні подання слів мають вирішальне значення. Для вивчення вкраплень на основі великих і зашумлених наборів даних використовуються напівспостережні методи та гібридні моделі. Інші вивчали вкраплення з урахуванням шляхів залежності для вилучення на основі CRF, або використовували увагу і метричне навчання для створення вкраплень, придатних для кластеризації схожих аспектів.

1.8.4 Категоризація аспектів

Крім простого вилучення аспектних термінів (наприклад, «картинка» або «фотографія»), існує й суміжне завдання – категоризація аспектів. Це передбачає об’єднання синонімічних або тісно пов’язаних термінів аспекту

в одну канонічну категорію (наприклад, об’єднання термінів «картинка», «фотографія» і «зображення» в категорію «Якість зображення»). Таке об’єднання дає змогу отримати зведення вищого рівня. У вирішенні цього завдання можуть допомогти такі методи, як кластеризація на основі вкраплень. Після того як аспекти визначено, наступним, уже розглянутим раніше, кроком є визначення полярності настрою (позитивного, негативного або нейтрального), вираженого стосовно кожного з них.

Долаючи розрив між неконтрольованим використанням лексикону і контрольованим навчанням, напівконтрольовані методи, такі як подвійне розповсюдження, використовують природні відносини між аспектами і словами думки. Починаючи з невеликого вихідного списку слів думки, система використовує синтаксичні шаблони для виявлення аспектів-кандидатів. Потім ці нові аспекти використовуються для визначення більшої кількості слів думки, і так по колу. Такий підхід дає змогу одночасно розширювати як список аспектів, так і лексикон думок.

1.9 Складнощі в аналізі настроїв

Аналіз настроїв, аналіз тональності тексту, комп’ютерне вивчення думок, емоцій та суб’єктивності в тексті, швидко перетворився з нішевої галузі досліджень на наріжну технологію в різних сферах – від маркетингових досліджень та обслуговування клієнтів до політології та моніторингу соціальних мереж. Привабливість та ефективність автоматичного вимірювання суспільних настроїв, розуміння споживчих вподобань та відстеження реакції на події не викликає сумнівів. За останні десятиліття було досягнуто значних успіхів: дедалі складніші алгоритми та величезні масиви даних сприяли прогресу. Однак, незважаючи на ці досягнення, шлях до справжнього розуміння нюансів людських почуттів, виражених через мову, залишається пов’язаним зі складними і постійними проблемами. Ці перешкоди охоплюють тонкощі самої мови, складнощі

збору та інтерпретації даних, а також постійно еволюціонуючі способи людської комунікації, особливо в цифровій сфері.

Один із найфундаментальніших викликів криється у самій природі людської мови. Мова – це мінлива, динамічна і глибоко контекстуальна сутність, яка часто чинить опір прямолінійній інтерпретації машинами. Неоднозначність є її постійним супутником; одне слово, фраза чи символ може нести дуже різне смислове навантаження залежно від навколишнього тексту та ширшого контексту. Візьмемо, наприклад, слово «хворий», яке може позначати хворобу (негативне значення) або бути сленговим позначенням чогось чудового (позитивне). Без глибокого розуміння контексту система аналізу настроїв може легко неправильно витлумачити його значення.

Ця складність збільшується, коли ми маємо справу з образною мовою. Люди є майстрами непрямої комунікації, використовуючи сарказм, іронію, метафори та гіперболи, щоб додати кольору, гумору чи акценту до своїх висловлювань. Сарказм, зокрема, являє собою серйозний виклик, оскільки він часто передбачає висловлювання, протилежне тому, що мається на увазі. Наприклад, «О, чудово, ще одна зустріч», швидше за все, виражає негативний настрій, незважаючи на те, що «чудово» звучить позитивно. Як досліджували Чжан та ін. [19], використовуючи глибокі нейронні мережі для виявлення сарказму в твітах, виявлення такої лінгвістичної акробатики вимагає моделей, здатних вийти за межі поверхневих значень слів і вловити тонкі контекстуальні підказки, можливо, навіть спираючись на історичні моделі взаємодії або риси особистості [23], включивши в них емоції та особистісні особливості. Складність автоматичного виявлення і правильної інтерпретації цих нелітературних виразів залишається значним бар’єром на шляху досягнення високої точності в реальних застосуваннях. Навіть якщо системи не навчені працювати з саркастичними даними, їхня продуктивність може різко знизитися, коли вони стикаються з ними, що підкреслює потребу в спеціалізованих підходах.

Проблема деталізації ще більше ускладнює ситуацію. Хоча класифікація всього документа може бути корисною, часто потрібне більш детальне розуміння, що призводить до аспектно-орієнтованого аналізу настроїв. Визначення конкретних аспектів або характеристик, що обговорюються (наприклад, «камера» або «час автономної роботи» телефону), а також думок, висловлених щодо кожного окремого аспекту, є складним завданням саме по собі. В одному відгуку можуть хвалити камеру і нарікати на час роботи акумулятора, що вимагає від систем розмежування цих думок. Це стає ще складнішим завданням, коли думки є неявними або аспекти не згадуються явно.

Тісно пов’язане з цим, але відмінне від нього виявлення позиції. Це передбачає визначення позиції автора (за, проти або нейтральна) щодо конкретної цілі або пропозиції, яка може навіть не бути безпосереднім об’єктом емоцій у тексті. Наприклад, твіт «Вагітні – це не просто ходячі інкубатори!» вказує на позицію легалізації абортів, навіть якщо «аборт» не є безпосереднім об’єктом висловлення думки. Як підкреслили Мохаммад, Собхані та Кириченко [24], це часто вимагає виведення зв’язків і розуміння ширших наслідків – завдання, з яким системи все ще борються, особливо коли ціль думки в тексті відрізняється від цілі, що цікавить для позиції.

Крім того, амбіції аналізу почуттів розширюються за межі простої полярності, щоб охопити багатший гобелен людських почуттів. Виявлення семантичних ролей почуттів та емоцій – визначення не лише того, що хтось відчуває емоцію, але й того, хто її відчуває (суб’єкт), що її викликало (стимул), і, можливо, причина або ступінь – це межа, яка підштовхує системи до більш фреймового розуміння емоційних подій. Аналогічно, вихід за межі позитивного/негативного для класифікації ширшого спектру дискретних емоцій, таких як радість, гнів, смуток, страх і здивування, або навіть таких вимірів, як збудження, представляє свій власний набір викликів, вимагаючи більш нюансованих лексиконів і моделей.

Самі дані, які є основою сучасних систем аналізу настроїв, також несуть у собі безліч проблем. Цифровий текст, з якого часто видобувають настрої, особливо в соціальних мережах, характеризується розрідженістю та зашумленістю. Абревіатури, творчі помилки, сленг, емодзі та незакінчені речення – все це нестримно поширюється, що ускладнює ефективну обробку традиційними інструментами NLP. Така «розрідженість даних» може суттєво вплинути на точність класифікатора. Крім того, обмеження конфіденційності може призвести до неповноти даних; наприклад, відсутність інформації про геолокацію може перешкоджати аналізу, який має на меті зрозуміти регіональні відмінності в настроях.

Сам процес анотування даних для навчання та оцінювання пов’язаний із суб’єктивністю та складністю. Прохання до анотаторів позначити текст як позитивний, негативний або нейтральний, хоча й здається простим, може призвести до неузгодженостей. Що означає «позитивний»? Це емоційний стан мовця чи позитивна оцінка події? Як слід коментувати речення, що описує успіх однієї спортивної команди над іншою, якщо його зміст повністю залежить від прихильності читача? Нейтральне висвітлення дуже емоційних подій (наприклад, «Війна призвела до появи мільйонів біженців») ставить ще одну дилему: чи є мовець нейтральним, чи притаманний події негатив є домінуючим настроєм, який потрібно передати? Сарказм, цитати та ретвіти ще більше каламутять воду, оскільки часто незрозуміло, чи схвалює мовець висловлену думку. Ці питання потребують чітких, але не надто складних рекомендацій щодо анотування, і навіть тоді досягнення високої узгодженості між анотаторами може бути важкодосяжним.

Залежність моделей аналізу настроїв від предметної області є ще однією добре задокументованою проблемою. Класифікатор, навчений на рецензіях на фільми, з його специфічною лексикою та поширеними виразами настрою, часто погано працює, коли його застосовують до фінансових новин або медичних текстів. Слова, що виражають почуття, та

їхні конотації можуть кардинально відрізнятися в різних доменах; «непередбачуваний» може бути негативним для акцій, але позитивним для сюжету кінофільму. Це вимагає або навчання моделі для конкретного домену, що є ресурсномістким, або розробки надійних методів адаптації до домену, які все ще залишаються активною сферою досліджень.

Багатомовний аналіз настроїв набуває все більшого значення, але він також створює унікальні перешкоди. Прямий переклад тексту за допомогою машинного перекладу може призвести до втрати або зміни настрою, оскільки нюанси та культурні конотації важко зберегти. Багатьом мовам бракує багатих лексиконів на позначення почуттів та анотованих корпусів, доступних для англійської мови, що ускладнює створення надійних систем. Саме вираження настрою може відрізнятися структурно і культурно в різних мовах, а це означає, що методи, ефективні для англійської мови, не можуть бути безпосередньо перенесені на інші мови. Хоча деякі дослідження, як-от Саламе, Мохаммад і Кириченко [25], присвячені арабським соціальним мережам, показують, що переклад може давати конкурентоспроможні результати, проблема точного відображення настроїв у мовному розмаїтті світу залишається значною, особливо для мов з обмеженими ресурсами, де моделі глибокого навчання можуть запропонувати шлях уперед, як це досліджували Ахтар та ін. [26] для гінді.

1.9.1 Зловмисні користувачі та шахрайські практики

Зростання кількості спаму в соціальних мережах та фейкових відгуків є значною проблемою. Приватні особи або компанії можуть публікувати неправдиві відгуки, щоб просувати власні продукти або дискредитувати конкурентів. Як зазначають Джиндал і Лю [27], ці фейкові відгуки часто дуже важко відрізнити від справжніх, просто прочитавши їх, що робить автоматизоване виявлення критично важливим, але складним завданням. Проблема полягає у виявленні тонких підказок, які відрізняють автентичний

досвід від сфабрикованих наративів, що часто вимагає аналізу не лише самого тексту, але й поведінки рецензентів, мережевих патернів і розподілу рейтингів.

Ця проблема ускладнюється поширенням соціальних ботів – автоматизованих акаунтів, створених для імітації людських користувачів у соціальних мережах. Боти в соціальних мережах можуть бути запрограмовані генерувати і поширювати думки в масовому масштабі, штучно роздуваючи уявну популярність або негативне ставлення до продуктів, осіб чи політичних ідей. Відрізнити справжні людські настрої від згенерованого шуму має вирішальне значення для надійності будь-якого аналізу настроїв, який спирається на соціальні медіа. Це також вимагає складних механізмів виявлення ботів, які працюють у тандемі з класифікаторами настроїв.

Забігаючи наперед, можна сказати, що «Opinion Mining» стикається з кілька ширшими викликами, які з’являються все частіше. Онлайн-світ змінюється: розвивається мова, з’являється новий сленг, змінюються теми для обговорення, змінюється база користувачів. Це вимагає динамічних систем аналізу настроїв, які можуть адаптуватися до цих часових змін, а не покладатися на статичні моделі, навчені на історичних даних. Зростаюча доступність різнорідних джерел даних – тексту, зображень, відео, аудіо – відкриває нові можливості і водночас кидає виклик. Справжнє розуміння настроїв може вимагати об’єднання інформації з різних модальностей, наприклад, аналіз настроїв відео на YouTube шляхом аналізу транскрипту, візуальних підказок у відео та тону голосу в аудіозаписі. Це вимагає вдосконалених методів мультимодального аналізу. Розуміння впливу соціальних мереж – як зв’язки між користувачами формують настрої, як поширюються думки і наскільки впливовими є окремі користувачі – вимагає виходу за рамки ізольованого аналізу окремих дописів. Інтеграція мережевого аналізу з аналізом настроїв може дати набагато глибшу і точнішу картину колективної думки.

Пошук продовжує виходити за рамки простої полярності, щоб охопити весь спектр людського впливу. Хоча виявлення базових позитивних, негативних або нейтральних настроїв є фундаментальним завданням, зростає потреба в системах, здатних ідентифікувати ширший спектр дискретних емоцій (таких як радість, гнів, страх, здивування) та їхню різну інтенсивність. Для цього потрібні більш досконалі лексикони емоцій і моделі, здатні розпізнавати тонкі емоційні сигнали.

2 ТЕОРЕТИЧНЕ ПІДҐРУНТЯ ТА ОБРАННЯ РЕФЕРЕНТНИХ МОДЕЛЕЙ ДЛЯ ВИДОБУТКУ ДУМОК

Незважаючи на те, що аналіз охопив досить різноманітний спектр методів і рішень, реалізація не може бути досконалою і поєднувати лише переваги запропонованих часом методів. З огляду на оцінку актуальності способів аналізу складних аспектів суспільного дискурсу, розберемо моделі, що можна представити різними методологічними підходами, від широкої класифікації за полярністю до тонкого розпізнавання настроїв і емоцій на основі аспектів. Потенційні можливості та застосування, що випливатимуть з реалізації, можуть досліджуватися в різних секторах і з різними необхідними відхиленнями.

2.1 Попередня оцінка полярності публікації новини

Для загальної класифікації настроїв були обрані моделі, здатні визначати загальну позитивну, негативну або нейтральну полярність тексту, що є основоположною частиною аналізу для розуміння суспільного настрою не тільки в цілому, але також відображеного і сформованого в дискурсі ЗМІ. Новинні статті часто містять складні лінгвістичні структури, розгорнуті розповіді та нюансовані вирази, які все частіше вимагають використання складних аналітичних моделей. Вибір моделей для цього завдання був заснований на їх перевірених можливостях в області розуміння природної мови та контекстуальної інтерпретації.

Моделі на основі трансформерів (наприклад, BERT, RoBERTa, DeBERTa) були обрані завдяки їх передовій архітектурі, яка використовує механізми уваги для вловлювання далеких залежностей і контекстуальних нюансів у тексті. На відміну від традиційних моделей, які обробляють текст послідовно, трансформери можуть одночасно оцінювати значимість різних слів у реченні або документі, що робить їх дуже ефективними для часто

довгих і структурно складних новинних статей. Їх попередня підготовка на великих текстових корпусах забезпечує глибоке базове розуміння мови.

FinBERT є спеціалізованим варіантом BERT, спеціально налаштованим для текстів з фінансової сфери (наприклад, фінансові новини, звіти компаній). Її включення підкреслює важливість адаптації до конкретної області для досягнення точної інтерпретації настрою в спеціалізованих областях публічного дискурсу. Фінансова мова часто містить термінологію, в якій полярність настрою може значно відрізнятися від загального вживання (наприклад, «зобов'язання», «ризик»). FinBERT розроблений для правильної інтерпретації цих специфічних для даної області нюансів.

Двонаправлені мережі з довгою короткочасною пам'яттю (BiLSTM) обрані як надійна альтернатива для обробки послідовних даних. Хоча BiLSTM, як правило, менш обчислювально-ємні, ніж великі трансформаторні моделі, вони ефективно вловлюють інформацію як з попередніх, так і з наступних слів у послідовності. Це робить їх придатними для аналізу настрою в коротших новинних фрагментах, заголовках або твітах, і вони служать цінним еталоном для порівняння з трансформаторними архітектурами.

Стратегічний вибір такого різноманітного набору моделей полегшує порівняльну оцінку ефективності, дозволяючи оцінити їх ефективність для текстів різної довжини, різного ступеня лінгвістичної складності і з наявністю специфічної для даної області лексики. Використання цих моделей у проекті не є суворим або якимось чином обмеженим, такий вибір лише забезпечує оптимальний комплексний підхід до аналізу тональності.

Можливість систематично оцінювати полярність новинних публікацій хоч і здається не настільки важливим завданням, але вона є стандартом, з яким можуть порівнюватися конкретні думки. Також це все ще відкриває значні можливості в багатьох секторах, наприклад, у корпоративному секторі це включає управління репутацією бренду за

допомогою постійного моніторингу публікацій, спрямованих на компанію, що дозволяє розробляти проактивні PR-стратегії або проводити ринкову розвідку шляхом розуміння настроїв ЗМІ щодо конкурентів або тенденцій у галузі для прийняття стратегічних рішень. Для фінансових ринків стандартне застосування включає інтеграцію оцінок настроїв в алгоритмічні торгові моделі, а також формування інвестиційної стратегії на основі аналізу переважаючих. У політології та державному управлінні такий аналіз допомагає в аналізі сприйняття політики шляхом оцінки настроїв ЗМІ щодо нового законодавства, моніторингу виборчих кампаній шляхом відстеження настроїв щодо кандидатів і уточнення повідомлень для інформаційних кампаній. А в соціологічних дослідженнях аналіз новинних джерел дозволяє відстежувати упередженість ЗМІ шляхом порівняння настроїв у різних новинних агентствах.

2.2 Загальна оцінка полярності комментарію

Ця система передбачає завдання, яке, як видається, має найбільшу вагу з точки зору аналізу настроїв суспільства. Як і для новин, необхідно визначити загальну полярність настрою коментарів користувачів, зокрема з соціальних мереж, які характеризуються неформальною мовою, стислістю та унікальними лінгвістичними особливостями. Тексти в соціальних мережах представляють собою особливі випробування, включаючи сленг, смайлики, хештеги, орфографічні помилки і нестандартну граматику. Моделі для цієї системи були обрані за їх здатність ефективно обробляти особливості коротких і простих дописів.

Twitter RoBERTa – це варіант моделі RoBERTa, попередньо навченої на великому корпусі твітів. Завдяки спеціальній попередній підготовці вона особливо добре розуміє нюанси мови соціальних мереж, в тому числі інтерпретує смайлики і хештеги як індикатори настрою. Надійне розуміння

контексту має вирішальне значення для усунення неоднозначності в коротких, часто двозначних коментарях.

Архітектури BiLSTM з механізмами уваги ефективні для послідовного тексту. Додавання саме цього механізму уваги розширює їх можливості, дозволяючи динамічно оцінювати важливість різних слів або токенів в коментарі. Це особливо корисно в текстах соціальних мереж, де ключові слова, що несуть настрій, можуть бути перемішані з нерелевантними токенами.

Методи на основі лексикону (наприклад, VADER, TextBlob) використовують заздалегідь визначені словники слів і пов’язані з ними оцінки настрою, а також правила для обробки інтенсифікаторів, заперечень та інших лінгвістичних конструкцій. VADER (Valence Aware Dictionary and sEntiment Reasoner) спеціально налаштований для текстів у соціальних мережах. Хоча і пошук якісних словників є випробуванням, лексико-орієнтовані підходи є обчислювально ефективними і забезпечують високий ступінь інтерпретованості, слугуючи цінною базою або компонентом в ансамблевих системах, застосування яких не варто упускати.

Ансамблеві стратегії передбачають об’єднання прогнозів з декількох різних моделей. Ансамблеві методи використовуються для підвищення загальної надійності та точності шляхом агрегування різних аналітичних точок зору, що може пом’якшити індивідуальні слабкі сторони складових моделей, приводячи до більш надійної класифікації сентименту, особливо для галасливих і динамічних текстів у соціальних мережах.

2.3 Поглиблення в емоційний стан коментатора

Це завдання вже виходить за рамки простої оцінки полярності і класифікує текст, зокрема коментарі користувачів, що нас цікавлять, за заздалегідь визначеним набором основних категорій емоцій: смуток, радість, любов, гнів, страх і здивування. Для розпізнавання емоційного

стану потрібні моделі, здатні розрізняти більш тонкі лінгвістичні сигнали, що вимагає як високої точності, так і ефективності. Тут, крім таких беззаперечно ефективних моделей, як Twitter RoBERTa, мережі BiLSTM і ансамблеві моделі, буде ще використовуватися DistilBERT.

Саме спрощена версія BERT, що зберігає значні можливості розуміння мови, при цьому будучи більш ефективною з точки зору обчислень, робить її придатною для великомасштабної класифікації емоцій як індивідуально, так і в ансамблі.

Вибір цих моделей дозволяє застосовувати багатогранний підхід до розпізнавання емоцій. Розуміння емоційного підтексту публічних коментарів дає більш глибоке уявлення про ситуацію. Крім створення продукції та контенту, ЗМІ та політики, систему для обчислення емоційного спектру можна застосувати і в багатьох інших сферах. В управлінні клієнтським досвідом це дозволяє поліпшити підтримку клієнтів, виявляючи такі емоції, як розчарування, для індивідуальних відповідей і проактивного відновлення обслуговування. Цікаве застосування в області психічного здоров’я і благополуччя, де аналіз сукупних емоційних тенденцій може служити індикатором суспільного психічного благополуччя (з урахуванням етичних міркувань) і допомагати в управлінні динамікою груп підтримки. Також потенційно, в області взаємодії людини і комп’ютера це може бути інтегровано в системи штучного інтелекту для більш емпатичної взаємодії.

2.4 Виявлення ставлення до певних прокоментованих аспектів

Аспектний аналіз настроїв виходить за рамки загальної полярності і включає в себе виявлення конкретних аспектів (особливостей, тем або сутностей) в тексті і визначення настрою, вираженого по відношенню до кожного з них, щоб забезпечити детальне розуміння думок. ABSA вимагає моделей, здатних як виявляти релевантні аспекти, так і класифікувати

пов’язані з ними настрої, часто в межах одного речення. Це вимагає складного розуміння контексту і можливостей структурованого прогнозування.

Основними будуть моделі на основі BERT/DeBERTa для вилучення аспектних термінів (ATE), класифікації аспектних настроїв (ASC) і наскрізного ABSA (end-to-end). Універсальні моделі Transformer, такі як BERT і DeBERTa, використовуються для ATE з метою вилучення аспектних термінів, а для ASC – з метою класифікації настроїв по відношенню до кожного аспекту. Наскрізний же ABSA передбачає одночасне навчання однієї моделі з заздалегідь заготовленими аспектами для виконання обох завдань.

LoRA (Low-Rank Adaptation) буде необхідний для тонкого налаштування великих моделей. LoRA – це ефективна з точки зору параметрів техніка тонкого налаштування, яка дозволяє адаптувати великі моделі до ABSA шляхом навчання тільки невеликої кількості додаткових параметрів, що знижує обчислювальні витрати.

Традиційні моделі маркування послідовностей, де архітектура BiLSTM-CRF поєднує BiLSTM з шаром CRF, можуть успішно використовуватися для ефективного структурованого прогнозування в ATE. Також можна використовувати автономні CRF. Вони включені для порівняльного аналізу.

ABSA надає високодетальні аналітичні дані, перетворюючи загальні відгуки в корисну інформацію. В області розробки та управління продуктами вона пропонує відгуки на рівні функцій для поліпшення продуктів і конкурентного аналізу шляхом розбору продуктів конкурентів на рівні аспектів.

Додатковою особливістю для ABSA і не тільки, буде розпізнавання іменованих сутностей (NER). Воно являє собою ідентифікацію та класифікацію ключових інформаційних одиниць у тексті, таких як особи, організації, місця, дати, грошові значення та інші заздалегідь визначені

категорії (наприклад, продукти, інше). Хоча NER і не є прямим методом аналізу думок, він є важливою технікою попередньої обробки і додатковою технікою, особливо для забезпечення надійного аспектного аналізу настроїв (ABSA) і контекстуалізації думок. Правильний NER має вирішальне значення для точного визначення суб'єктів або об'єктів, про які висловлюються думки. У цьому плані моделі вибираються за їх здатністю точно ідентифікувати межі сутностей і правильно класифікувати їх в різних текстових контекстах.

BiLSTM обробляють текст послідовно, захоплюючи контекстну інформацію, цінну для NER, оскільки класифікація сутностей часто залежить від навколишніх слів. Або архітектура, що доповнює BiLSTM шаром CRF, який моделює залежності між прогнозованими тегами сутностей, що робить її дуже ефективною для NER, враховуючи достовірність послідовностей тегів. І звичайно моделі на основі трансформерів часто налаштовуються на наборах даних, специфічних для NER і можуть успішно впоратися з неоднозначністю сутностей.

NER відіграє важливу допоміжну роль в аналізі тональності, особливо для ABSA, шляхом ідентифікації конкретних об'єктів думок. Він дозволяє контекстуалізувати думки для ABSA, пов'язуючи тональність з конкретними сутностями (наприклад, пов'язуючи позитивну тональність для «екрану» з «iPhone») і усуваючи неоднозначність аспектів у складних реченнях з декількома сутностями. Крім прямої підтримки аналізу тональності, більш широкі області застосування NER включають одну з найважливіших частин вилучення інформації і побудову бази знань, категоризацію і рекомендацію контенту. А на додаток може включати і анонімізацію даних шляхом маскування конфіденційних сутностей, а ще й підтримку дотримання нормативних вимог і електронного виявлення шляхом ідентифікації у великих наборах документів. Надаючи інформацію про «хто» і «що» в дискурсі, NER значно підвищує точність і корисність подальших завдань аналізу тональності.

3 РОЗРОБКА ЕМПІРИЧНИХ МОДЕЛЕЙ, ОЦІНЮВАННЯ ТА ВИВЕДЕННЯ АРХІТЕКТУР

3.1 Набори даних для аналізу загальних настроїв

Щоб забезпечити широку застосовність, для аналізу настрою на основі полярності використовували кілька наборів формальних і неформальних текстових даних. Для аналізу новин використовували набір Financial PhraseBank і набори даних із новинних завдань SemEval. Ці набори даних значно різняться за розміром, лінгвістичним стилем і специфікою галузі, це дає змогу ретельно оцінити стійкість моделей у різних новинних контекстах. Прикладом для огляду послужить набір Financial PhraseBank, що містить вузькоспеціалізовану лексику (рисунок 3.1), яка стосується фінансових настроїв.

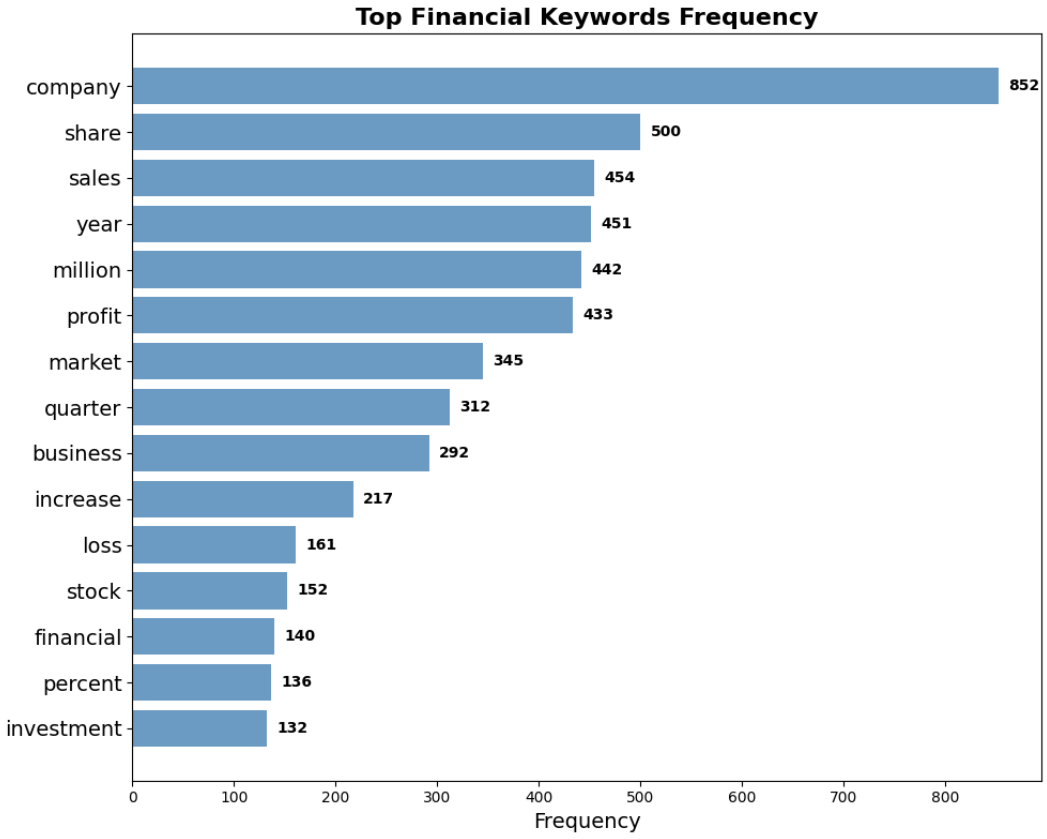

Рисунок 3.1 – Аналіз ключових слів фінансового домену

Проєкт з аналізу настрою коментарів використовує велику колекцію наборів даних соціальних мереж, що охоплює близько 292 000 попередньо оброблених зразків, детально зображено на рисунку 3.2. Основні набори даних включають: Twitter Combined, Twitter Social, Reddit Sentiment, US Airline Twitter Sentiment, Apple Twitter Sentiment.

Ці набори даних дуже важливі для вловлювання нюансів неформальної лексики, емодзі, хештегів і специфічних особливостей платформи, поширених у соціальних мережах.

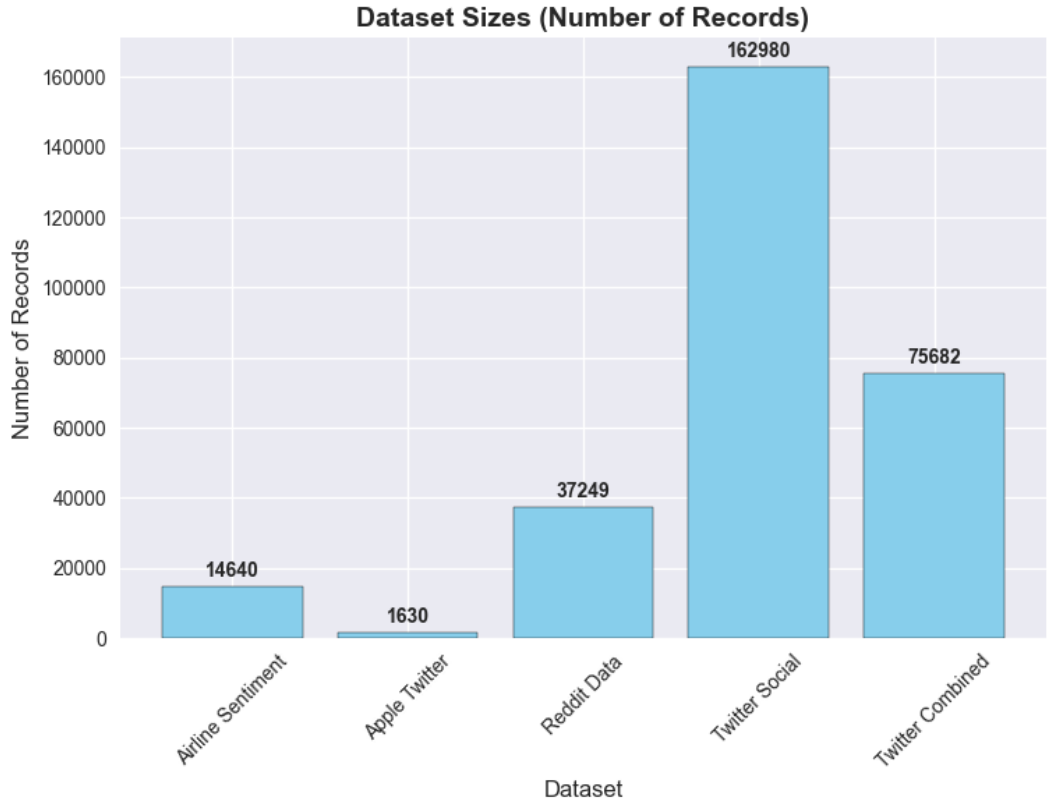

Рисунок 3.2 – Порівняння наборів даних для оцінки настроїв у коментарях

3.2 Набори даних для аспектного аналізу та аналізу емоцій

В основі аспектно-орієнтованого аналізу настроїв (ABSA) використовуються відомі набори даних SemEval ABSA за 2014-2016 роки. Ці набори текстів є золотим стандартом у цій галузі. Кожен набір даних

ретельно організований і розділений за певними доменами, такими як «ноутбуки» або «ресторани». Ця продумана структура дає змогу навчати як широкі, так і вузькоспеціалізовані моделі, які розуміють унікальну мову і вирази почуттів, характерні для кожної сфери.

Дані додатково збагачуються чіткими анотаціями «аспектних термінів», зображених на рисунку 3.3 (наприклад, «термін служби батареї» для ноутбука або «сервіс» для ресторану).

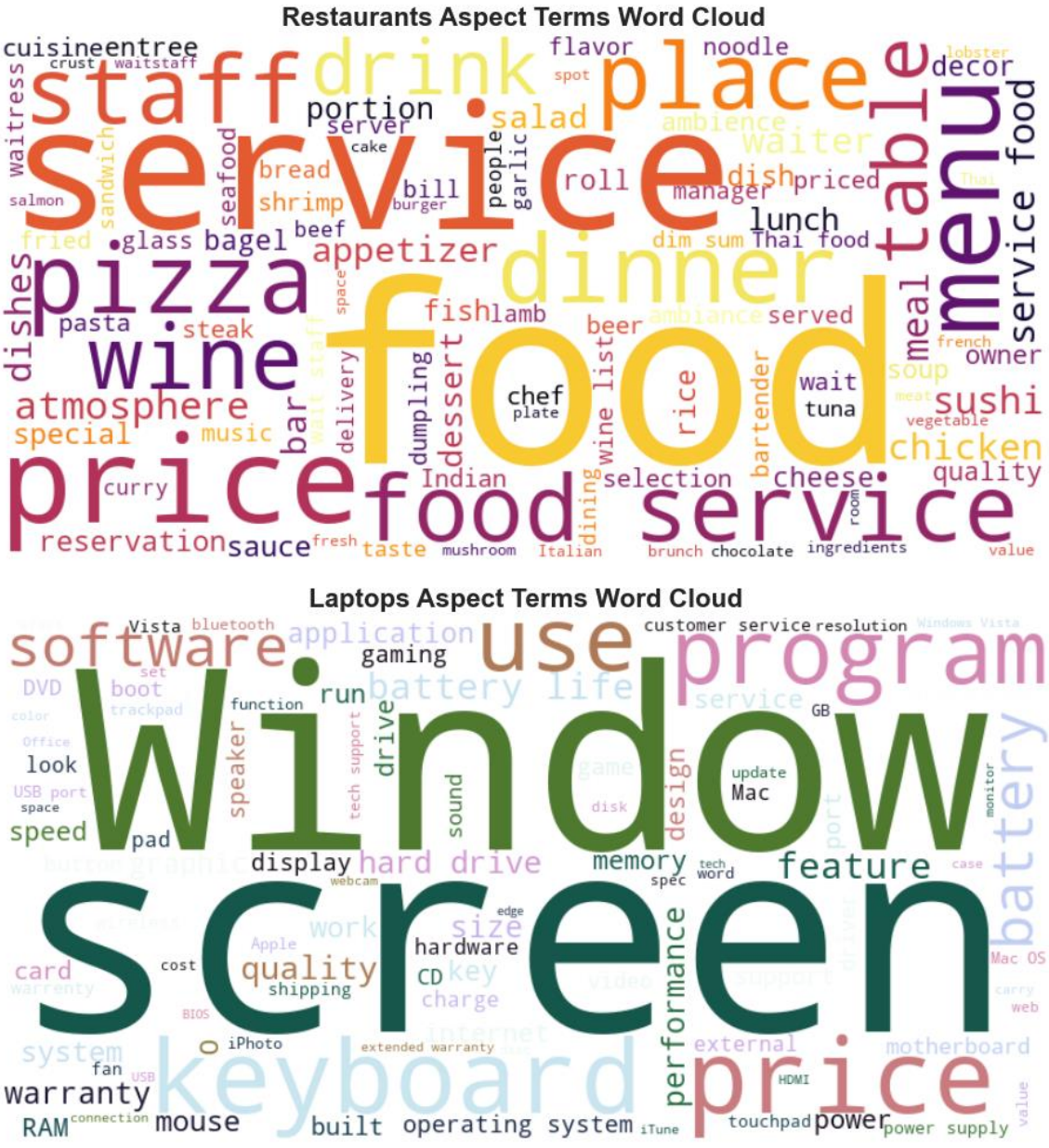

Рисунок 3.3 – Хмари слів для доменних наборів даних ABSA

Концепція розпізнавання емоцій розроблена для розуміння нюансів емоційного ландшафту в тексті. Набір даних являє собою ретельно анотовану збірку повідомлень у Twitter із шістьма основними емоціями: гнів, страх, радість, кохання, смуток і здивування (рисунок 3.4). Необхідно також розраховувати на роботу з різними джерелами даних, включно з власними, попередньо анотованими колекціями або загальнодоступними наборами даних, спеціально створеними для аналізу емоцій. Ця гнучкість дасть змогу створювати надійні моделі, здатні ідентифікувати та розрізняти тонкі емоційні потоки, які керують людською експресією.

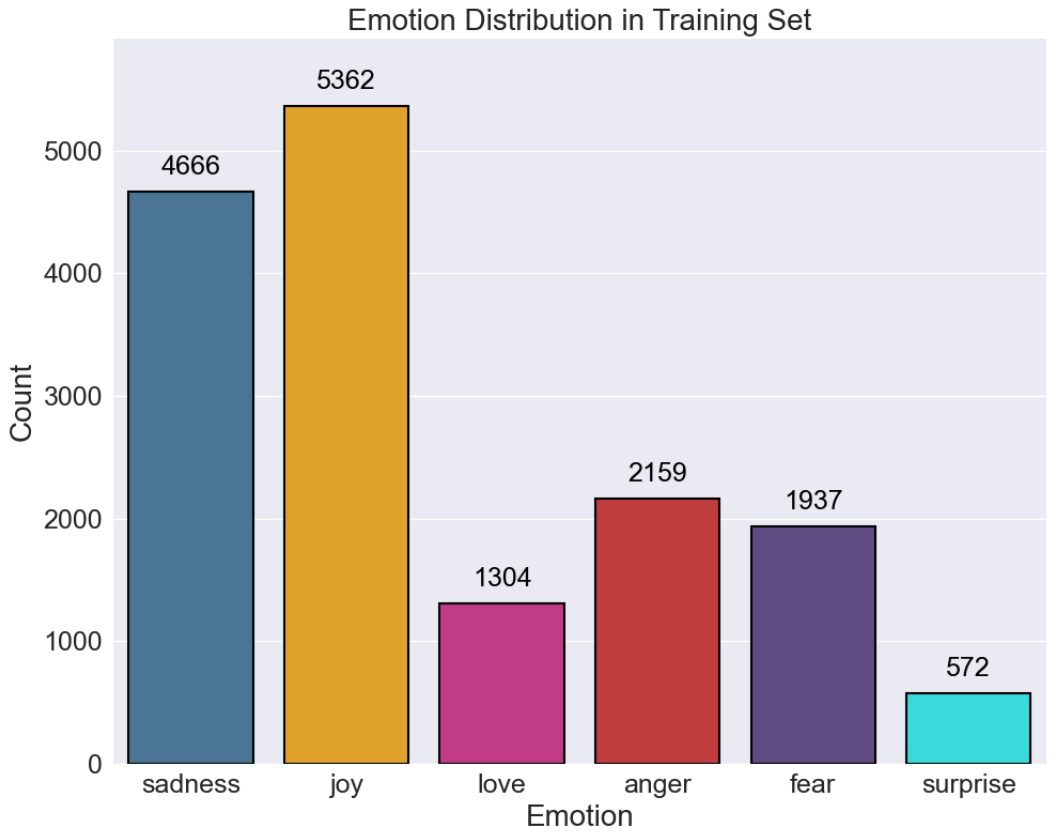

Рисунок 3.4 – Розподіл емоційних міток

В основі вилучення сутностей (NER) лежить обширний корпус списку POS-тегів для кожного слова в реченні і списку NER-тегів для кожного слова в реченні (наприклад, особи, організації, місця та інші сутності). Ці дані є основоположними для навчання моделей, здатних ідентифікувати

конкретні суб’єкти або об’єкти думок у тексті, а також є важливими для складніших аналітичних завдань, забезпечуючи важливий контекст, необхідний для зв’язку настроїв та емоцій із реальними об’єктами, до яких вони відносяться.

3.3 Попередня обробка даних та розробка характеристик

Ефективне попереднє опрацювання даних і розробка ознак – важливі попередні етапи для створення якісних моделей для аналізу настроїв. Очищаючи зашумлені дані, нормалізуючи різноманітні текстові вирази та структуруючи вхідні дані для спеціалізованих архітектур, попереднє опрацювання є невід’ємним компонентом, що суттєво впливає на продуктивність моделей. Загальний вплив на набори даних можна побачити, наприклад, у порівняльній статистиці необроблених і оброблених даних з погляду довжини тексту, згадок аспектів і розподілу настроїв, як показано на прикладі набору даних відгуків про ноутбуки на рисунку 3.5.

Основні методи очищення та нормалізації є фундаментальними для всіх проектів. Базові процедури очищення, як-от видалення тегів HTML під час обробки статей, зібраних із веб-сторінок і обробка розділових знаків, є стандартними. Маючи справу з неформальним контентом, як коментарі, створеним користувачами, важлива більш просунута попередня обробка, специфічна для соціальних медіа. Це передбачає систематичну обробку таких елементів, як @згадки, #хештеги і URL-адреси, які можна видалити або замінити загальними маркерами. Емодзі, поширені в соціальних мережах, також обробляються окремо, з опціями, що включають їхнє перетворення на текстові подання (наприклад, «усміхнене обличчя» для 😂) або їх видалення, залежно від архітектури нижчої моделі та аналітичної мети.

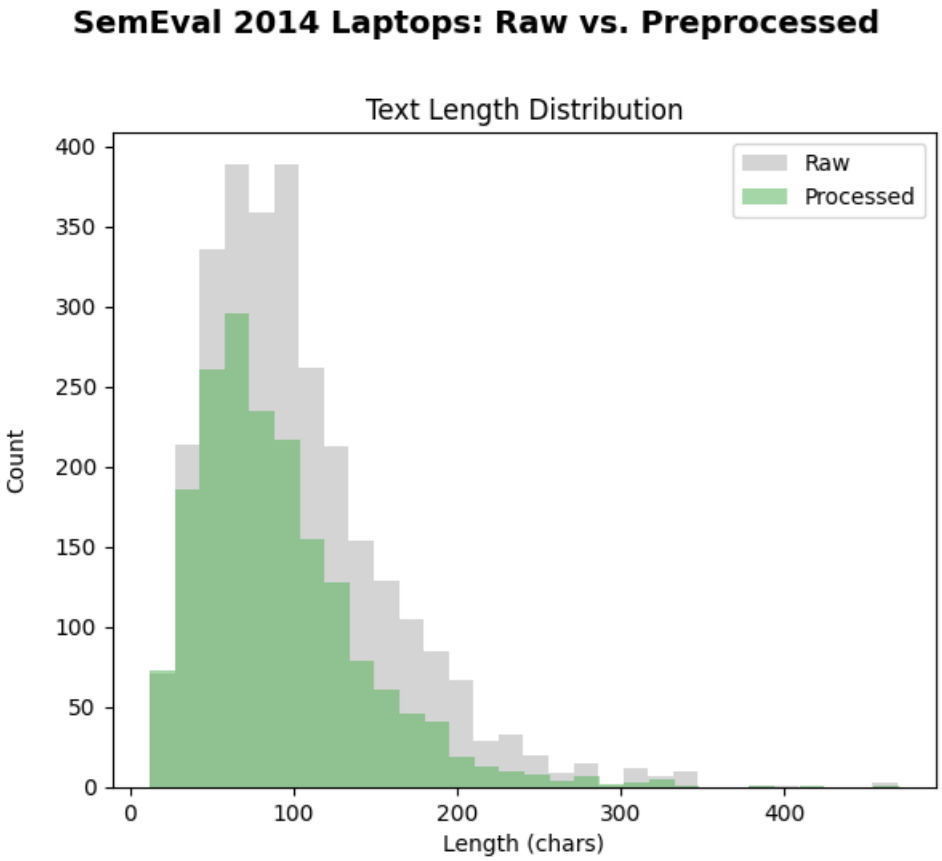

Рисунок 3.5 – Розподіл довжини тексту після попередньої обробки в наборі даних про ноутбуки

Для стандартизації тексту та зменшення убогості словникового запасу, особливо в даних соціальних мереж, широко застосовуються додаткові методи нормалізації тексту. До них належать розширення скорочень (наприклад, «don’t» до «do not»), виправлення орфографічних помилок для зменшення впливу друкарських помилок і неформального написання слів, а також нормалізація регістру (зазвичай перетворення всього тексту в нижній регістр). Кількісний вплив цих етапів очищення та нормалізації на показано на таблиці 3.1 та таблиці 3.2, де підкреслюється, як попереднє опрацювання може змінити метрики тексту, як-от середня кількість символів і слів у різних наборах даних соціальних мереж (наприклад, Airline Sentiment, Twitter Social Sentiment). Видалення стоп-слів (або загальних слів, таких як «the», «is», «in») і

стеммінг (скорочення слів до їхньої кореневої форми, наприклад, «running» до «run») також є ключовими техніками. Імпелементовані для зменшення шуму та розмірності ознак, наприклад, класи NewsPreprocessor та TweetPreprocessor у проєкті аналізу тональності новин та схожі з параметрами, які можна настроювати, для розпізнавання емоцій.

Таблиця 3.1 – Зменшення обсягу символів у наборах даних про настрої

	original_char_avg	clean_char_avg	char_reduction_pct
Airline Sentiment	103.82063	99.148125	4.501873
Apple Twitter Sentiment	95.341718	90.582512	4.991735
Reddit Sentiment	180.909796	181.434985	-0.290304
Twitter Social Sentiment	124.173518	123.267689	0.729487
Twitter Combined Sentiment	108.129568	110.844114	-2.510456

Таблиця 3.2 – Зменшення обсягу слів у наборах даних про настрої

	original_word_avg	clean_word_avg	word_reduction_pct
Airline Sentiment	17.653415	18.278228	-3.539332
Apple Twitter Sentiment	14.520245	15.966133	-9.957735
Reddit Sentiment	29.327579	29.550423	-0.759842
Twitter Social Sentiment	20.079948	20.009315	0.351763
Twitter Combined Sentiment	19.084049	20.219253	-5.948443

Токенізація, або процес розбиття тексту на окремі слова або підсловесні одиниці, є основоположним етапом практично для всіх завдань NLP, зокрема й для завдань з аналізу настроїв у новинах і NER. Після токенізації важливе значення має довжина послідовності, особливо це важливо для моделей із фіксованими обмеженнями за розміром вхідних даних. Для трансформаторних моделей, BERT або RoBERTa, які мають максимальну довжину вхідної послідовності (наприклад, 512 токенів), токенізація має життєво важливе значення під час обробки довгих документів. Тому статті систематично розбиваються з використанням механізмів для збереження контекстуальної цілісності. Вплив попередньої обробки на розподіл довжини тексту візуалізовано на зображенні на рисунку 3.6. Аналогічним чином, у проєкті NER реалізовано стратегії заповнення для опрацювання змінної довжини послідовностей у патчах.

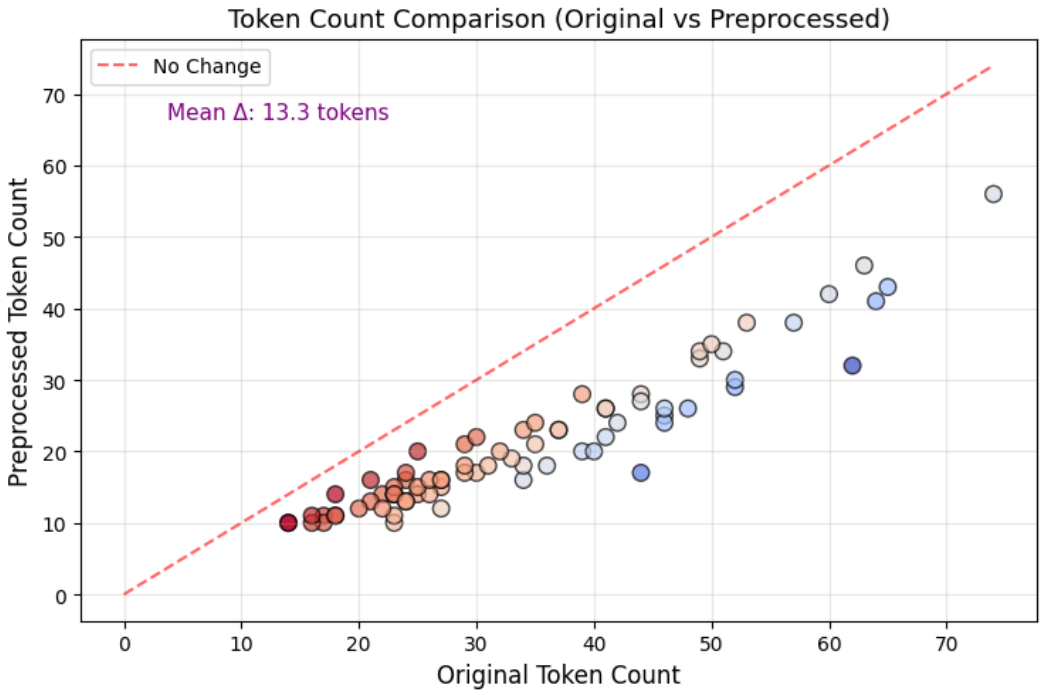

Рисунок 3.6 – Кількість токенів до та після попередньої обробки (Financial PhraseBank)

Удосконалена структуризація даних та інжиніринг ознак мають ще більше підвищити продуктивність. Створення словника, що використовується TweetPreprocessor, створює зв'язний і обмежений простір ознак із потенційно зашумлених і різноманітних лексиконів соціальних мереж. У таких завданнях, як NER, набори даних часто демонструють значний дисбаланс класів (наприклад, переважання токенів сутностей «Outside», рисунок 3.7). Для NER ще один ключовий аспект – збагачення характеристик. Де інтеграція попередньо навчених вкладень слів (наприклад, GloVe, FastText) забезпечує багаті семантичні відправні точки для слів. Також варто використовувати характеристики на рівні символів, щоб допомогти моделям опрацьовувати слова, яких немає в словнику, і розпізнавати морфологічні патерни, що вказують на іменовані сутності.

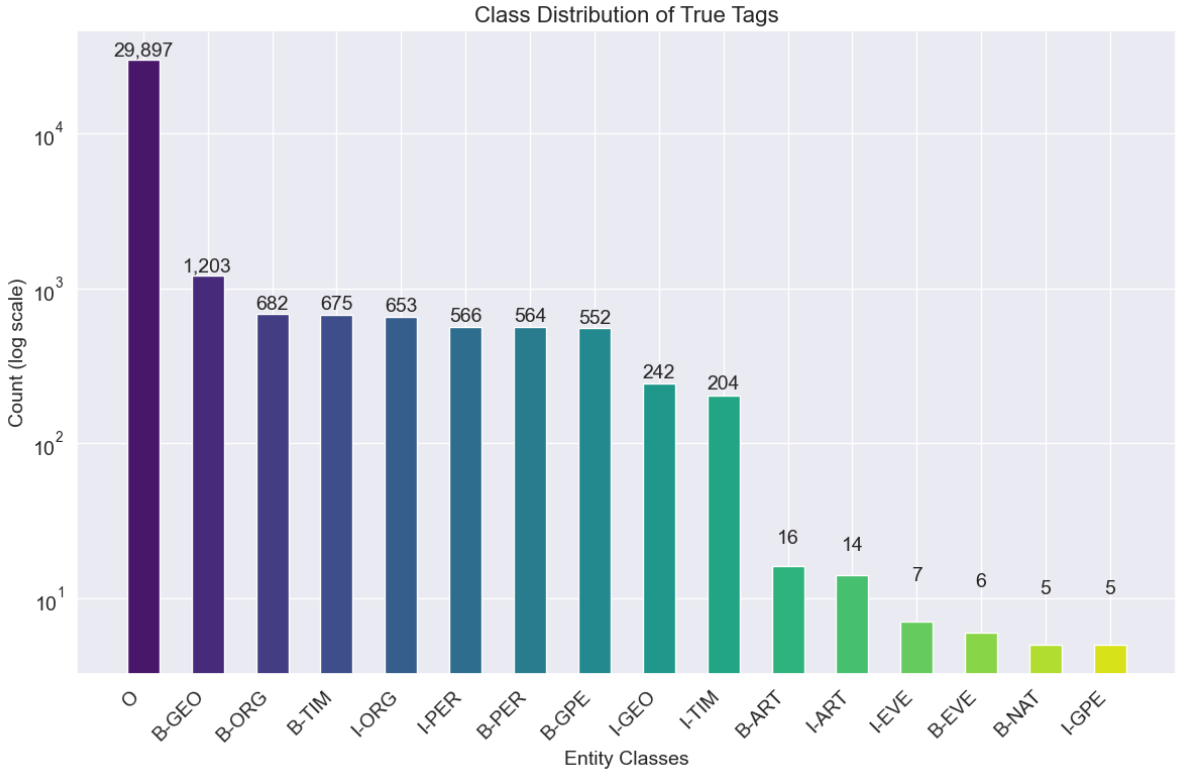

Рисунок 3.7 – Розподіл тегів за класами

Зрештою, для складних сценаріїв навчання або даних, специфічних для певної галузі, часто потрібне спеціалізоване опрацювання даних. Наприклад, для ABSA використовується модульний і розширюваний конвеєр попередньої обробки, розроблений для адаптації до різних галузей, таких як огляди продуктів або послуг. Це дуже важливо, оскільки характер аспектів і вираження емоцій можуть значно відрізнятися в різних областях. Важливим є перетворення вихідного тексту в структурований формат з ідентифікованими аспектами, сентиментами і тегами BIO (Beginning, Inside, Outside). Крім того, конвеєр ABSA містить у собі настроювані функції пакетної обробки (Collate Functions). Вони особливо необхідні для сценаріїв багатозадачного навчання, як-от у моделях типу «end-to-end», які одночасно вчаться витягувати аспекти та класифікувати їхній настрій.

3.4 Архітектури моделей та стратегії навчання

В основі перетворення необробленого тексту на змістовні висновки лежить складне розроблення моделей машинного навчання та дисципліновані стратегії їхнього навчання. З цією метою від оцінки настрою новин до вилучення іменованих сутностей, зроблено наголос на робочий процес, керований конфігурацією. Цей підхід слугує простою основою для експериментів, даючи змогу ретельно визначати та відстежувати критичні параметри за допомогою файлів у форматі YAML, наприклад deberta_ate_with_ner.yaml (додаток А) або, як ensemble_config.yaml (додаток Б). Такий систематичний процес дає змогу легко налаштовувати тренувальні прогони, порівнювати різні налаштування та забезпечує можливість відтворення експериментів, сприяючи прозорому та спільному дослідницькому середовищу.

Що стосується самих моделей, то основною стратегією в цьому контексті є тонке налаштування архітектур на основі трансформерів. Ці моделі мають широке розуміння мови, і завдання полягає в тому, щоб

адаптувати їхні знання до специфічних нюансів конкретних наборів даних і поставлених цілей. Процес тонкого налаштування – ретельна процедура, заснована на стандартних найкращих практиках глибокого навчання. Ретельно налаштовуються гіперпараметри, досліджуються різні графіки швидкості навчання (що включають фази розігріву для полегшення навчання моделі, а потім спаду для уточнення навчання). Також застосовується градієнтне обрізання для забезпечення стабільного і плавного навчання (рисунок 3.8).

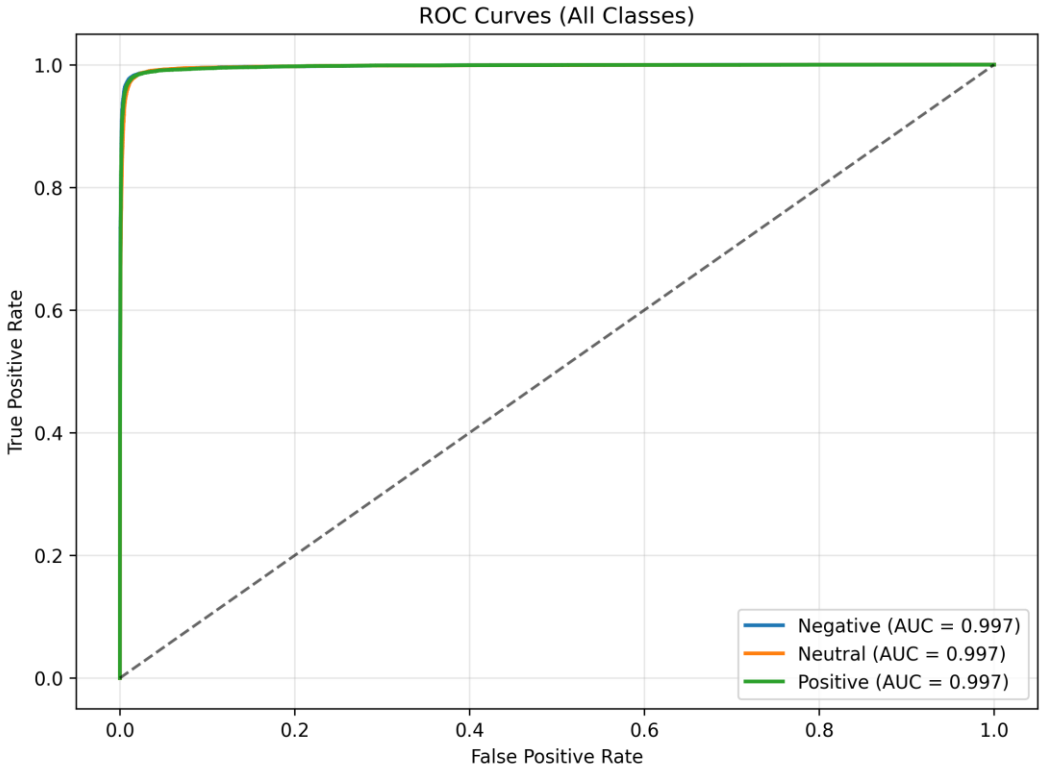

Рисунок 3.8 – Графік ROC-кривих для багатокласової класифікації новинних настроїв

Щоб запобігти запам’ятовуванню моделлю навчальних даних (overfitting) на шкоду узагальненню використовується рання зупинка. Для деяких завдань, що вимагають великих обчислювальних витрат, наприклад, у проєкті розширеного вилучення сутностей, також

використовується навчання зі змішаною точністю. Ця техніка використовує комбінацію чисел низької точності (float16) і стандартної точності (float32) під час навчання, що дає змогу значно прискорити процес і скоротити об’єм пам’яті без істотної шкоди для точності – значна перевага під час роботи з такими великими і складними моделями.

Поряд із трансформаторами, мережі з двонаправленою довготривалою пам’яттю (BiLSTM) залишаються універсальним і ефективним вибором, з’являючись у наборах інструментів для вивчення настрою новин, настрою коментарів, розпізнавання емоцій і NER. BiLSTM відмінно справляються із захопленням послідовної інформації, зчитуючи текст як вперед, так і назад, щоб створити багате розуміння контексту. Їхні можливості часто розширюються, використовуючи механізми уваги. Увагу можна уявити як прожектор, що дає змогу моделі фокусуватися на найважливіших словах або фразах у реченні під час ухвалення рішення, а не розглядати всі слова однаково. Щоб забезпечити BiLSTM сильну відправну точку, особливо під час роботи з невеликими наборами даних, їхні уявлення слів часто ініціалізуються за допомогою попередньо навчених вкраплень, таких як GloVe. Ці вкраплення забезпечують багату семантичну основу, а отже, моделі не потрібно заново вивчати значення кожного слова. Візуальне уявлення того, як конкретно влаштована одна з моделей на основі BiLSTM, що потенційно містить у собі вкраплення, увагу та інші шари, можна побачити на архітектурній схемі, представленій в додатку В. Для завдань, пов’язаних із визначенням послідовностей міток у тексті, як-от визначення аспектних термінів в ABSA або NER, архітектури BiLSTM часто доповнюються шаром умовного випадкового поля (CRF). BiLSTM може добре передбачати ймовірність належності кожного слова до певного класу (наприклад, «Початок імені людини»), але шар CRF йде на крок далі. Він розглядає всю послідовність передбачених міток і допомагає переконатися, що вони мають сенс разом (наприклад, мітка «Внутрішня організація» повинна логічно слідувати за міткою «Початок організації», а

не за міткою «Місцезнаходження»). Це доповнення призводить до більш послідовних і точних передбачень для цих завдань структурованого висновку. Відчутні результати такої складної моделі NER, яка може поєднувати BiLSTM і шар CRF, можна побачити в її продуктивності за допомогою матриці плутанини, як показано на рисунку 3.9.

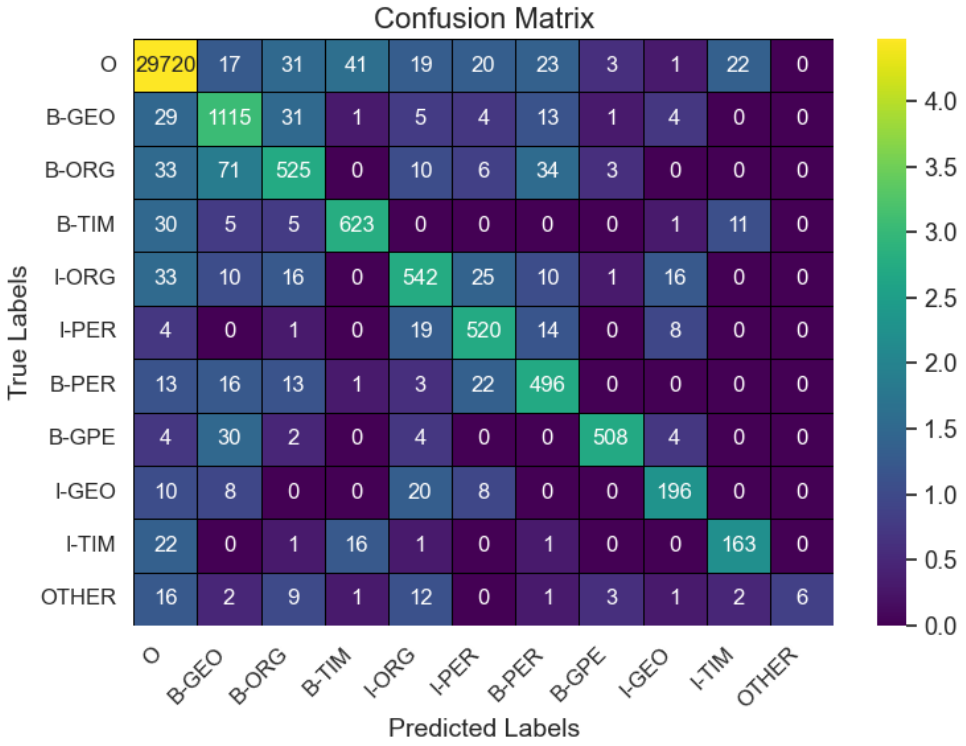

Рисунок 3.9 – Матриця плутанини для тегів NER

Визнаючи, що жодна архітектура моделі не є універсальним рішенням, ансамблеві моделі також посідають чільне місце, особливо в проєктах із розпізнавання настрою та емоцій у коментарях. Філософію тут можна описати як «мудрість натовпу». Навчаються кілька різних моделей як трансформатор, BiLSTM і навіть простіший підхід на основі лексикону, а потім їхні індивідуальні прогнози об’єднуються. Таке об’єднання, за допомогою механізмів зваженого усереднення або голосування за більшістю голосів, часто призводить до остаточного прогнозу, що є

надійнішим і точнішим, ніж могла б досягти будь-яка окрема модель. Стратегія ансамблю допомагає пом’якшити неминучі похибки та недоліки, властиві кожному окремому підходу.

Подальше вдосконалення передбачає інтеграцію окремих можливостей моделі за допомогою розширення функціональних можливостей, яскравим прикладом чого є інтеграція NER-ATE. Цей підхід використовує результати попередньо навченої моделі NER для поліпшення продуктивності вилучення термінів аспекту. Лінгвістична мотивація цього є очевидною: багато термінів аспекту, таких як назви продуктів або брендів, також є іменованими сутностями. Теоретично, надаючи явну інформацію про сутності моделі ATE, модель може краще розрізнити межі аспектів і поліпшити узагальнення. Цей процес передбачає модифікацію конвеєра даних для попереднього прогнозування тегів NER для кожного токена у вхідному тексті. Потім ці теги вбудовуються та надаються моделі ATE як додаткова вхідна особливість об’єднана з основними вбудованими токенами. Це збагачує вхідне представлення моделі, надаючи інформацію з навченої моделі NER. Ця техніка є формою трансферного навчання, яка може бути особливо корисною в сценаріях адаптації до домену або з обмеженими ресурсами.

Здебільшого весь життєвий цикл навчання організовується спеціальними скриптами навчання. Ці скрипти керують циклами навчання, ретельно записуючи докладні журнали прогресу і метрики продуктивності (зазвичай зберігаються в каталогах logs/) і старанно зберігають контрольні точки моделі через регулярні проміжки часу або щоразу, коли модель демонструє поліпшену продуктивність на валідаційному наборі (каталоги models/). Ретельний і продуманий процес навчання, реєстрації та ітеративного вдосконалення розробляють моделі, ефективність яких можна чітко зрозуміти за допомогою докладних оцінок, як-от звіти про класифікацію (рисунок 3.10), матриці плутанини (рисунок

3.11) і ROC-криві (рисунок 3.8), що забезпечує впевненість у їхній здатності розшифровувати складний гобелен людської мови.

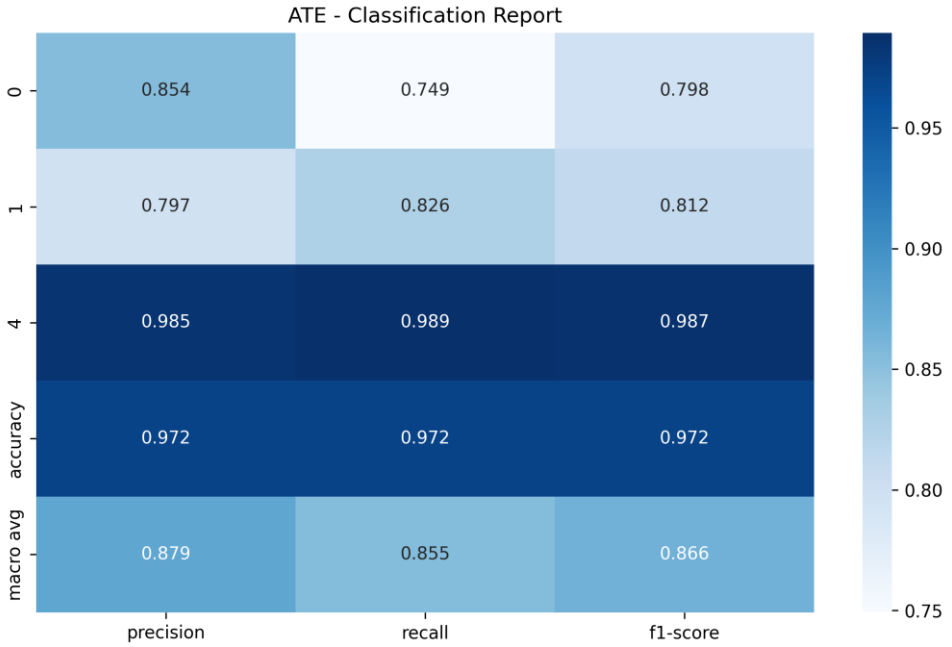

Рисунок 3.10 – Класифікаційний звіт для моделі ATE

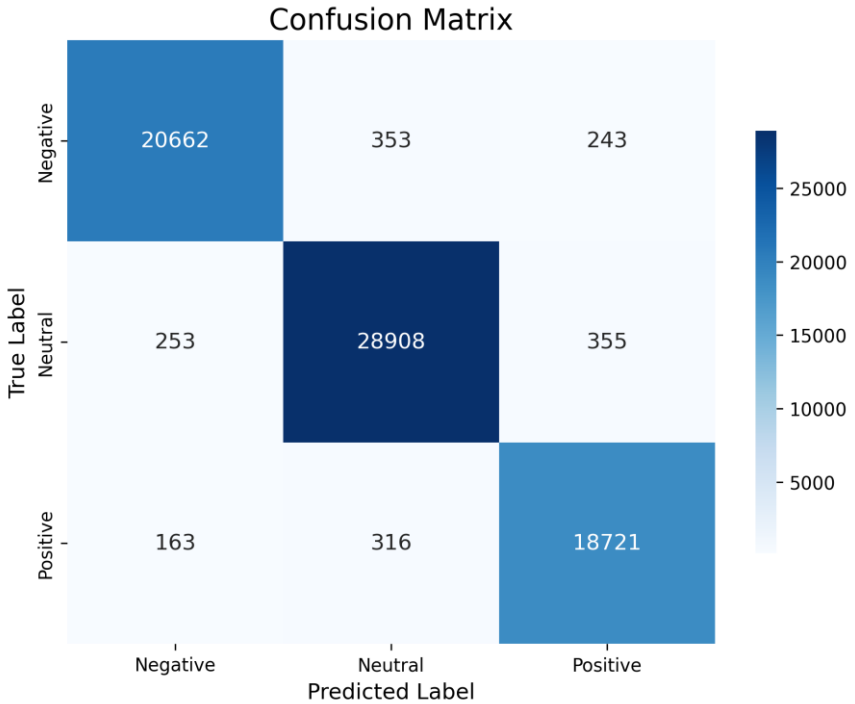

Рисунок 3.11 – Матриця плутанини для оцінки настроїв у коментарях

4 ПРАКТИЧНА РЕАЛІЗАЦІЯ

Теоретичні основи та архітектурні проекти моделей знаходять своє практичне застосування в чітко визначених робочих процесах. Реалізація являє собою невеликі інженерні системи, призначені для відтворення та налагодження експериментів, а потім і практичних висновків. Загальний робочий процес схожий у всіх обраних областях. Починаючи з оптимізованих процесів налаштування, які полегшуються скриптами для конфігурації середовища та управління залежностями. Це гарантує певну узгодженість, що перехід від одного проекту до іншого буде простим і звичайним процесом. Після налаштування важливим етапом є підготовка даних, деякі дані були описані в роботі і готові до використання. Для нових даних використовуються можливо перевантажені, але наочні, спеціально розроблені файли типу Jupyter Notebook для попередньої обробки. Процес конвеєра відповідає за перетворення необробленого, часто зашумленого тексту з різних джерел, таких як соціальні мережі або новинні статті, в чисті, нормалізовані і токенізовані формати для використання.

Навчання та оцінка моделей управляються за допомогою надійної, настроюваної інфраструктури. Конфігураційні файли YAML послідовно використовуються для визначення всіх параметрів навчання (швидкість навчання, розмір партій, гіперпараметри і т.д.) і додатково необхідних залежностей як посилання на словники або модель NER для ATE. Саме навчання здійснюється за допомогою спеціальних скриптів, які керують основним циклом навчання, реєструють метрики і зберігають контрольні точки моделі. Подальша оцінка проводиться з такою ж строгістю. Спеціальні скрипти генерують комплексні звіти зі стандартними метриками класифікації, такими як точність, F1-показник, точність і повнота. Ці кількісні показники доповнюються наочними візуалізаціями, які дають більш детальне уявлення про продуктивність моделі. Така оцінка дозволяє чітко відстежувати продуктивність і об’єктивно порівнювати моделі.

Крім оцінки, всі проекти оснащені функціями прогнозування, доступними для виконання інференції за новими, небаченими даними. Підтримується обробка окремих екземплярів і пакетна. Ключовим аспектом також є врахування не тільки точності, але і продуктивності та інтерпретованості. За необхідності включені оптимізації: компіляція XLA для прискорення лінійної алгебри, навчання зі змішаною точністю для балансу швидкості та використання пам’яті, ефективні методи обробки даних, такі як кешування та попередня вибірка. Щоб подолати розрив між прогнозуванням і розумінням, використовуються функції інтерпретованості моделей, такі як LIME (Local Interpretable Model-agnostic Explanations), які допомагають пояснити, чому модель зробила той чи інший прогноз.

4.1 Прогнозування полярності та емоцій

Для завдань, пов’язаних із загальною класифікацією настроїв та емоцій, навчені моделі обробляють необроблені текстові вхідні дані та генерують первинний результат: категорійну мітку (наприклад, «позитивний», «сум») та відповідний показник впевненості, що відображає впевненість моделі у своєму прогнозі.

У проекті з аналізу настроїв у новинах один текстовий рядок або файл, що містить кілька рядків, обробляється за допомогою спеціального скрипта прогнозування. Для даного тексту, такого як «International tourist spending in Europe seen up 11% this year, report says», модель спочатку застосовує свій конвеєр попередньої обробки, а потім передає отриманий вхідний сигнал через навчену нейронну мережу. У цьому прикладі видається прогноз з високим ступенем достовірності «позитивний» (p = 0,94). Щоб зрозуміти логіку моделі, налаштована візуалізація за допомогою LIME. Результат, показаний на рисунку 3.12, виділяє конкретні токени та їх вплив на цей прогноз. Можна помітити, що такі слова, як «seen up» (з найбільшою значимістю), логічно підштовхнули прогноз до позитивного настрою. Хоча

незначний негативний внесок таких токенів, як «report» або «Europe», може здатися суперечливим інтуїції, але не обов'язково вказує на фундаментальний недолік в контекстуальному розумінні моделі.

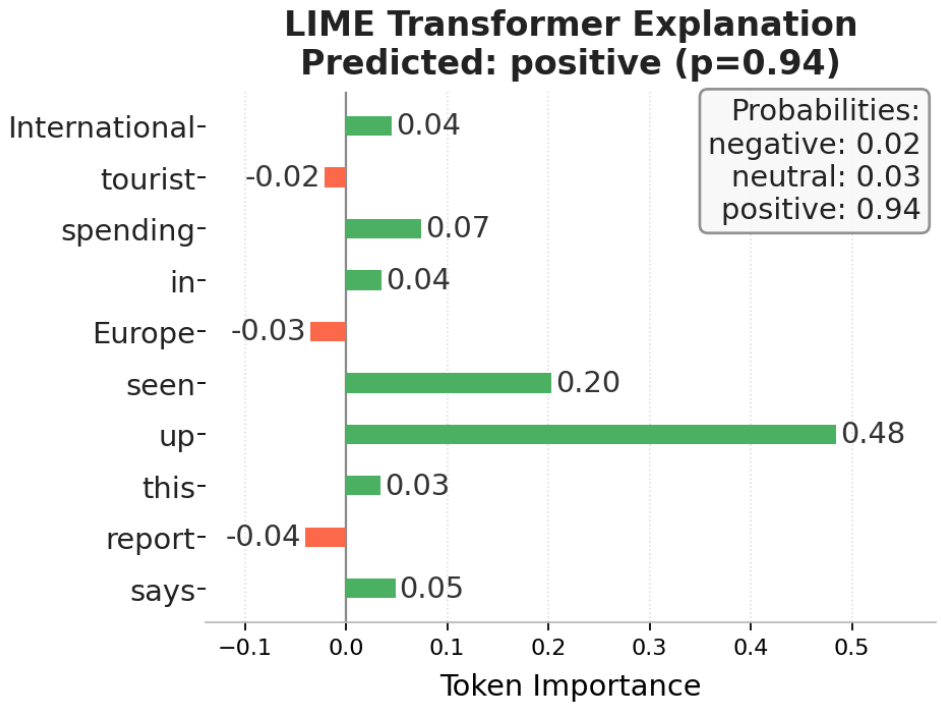

Рисунок 4.1 – Пояснення LIME для прогнозування моделі трансформера

Аналогічний робочий процес застосовується до аналізу настрою коментарів, де скрипт прогнозування приймає текстові входи, що представляють собою дискусію в соціальних мережах і подібних платформах. Для коментаря Reddit, такого як «Some are saying it's the greatest continent of all time.», спочатку застосовується спеціалізована попередня обробка, а потім, для прикладу, модель BiLSTM може з дуже високим ступенем впевненості класифікувати його як «позитивний» (p = 0,99). Пояснення LIME для цього прогнозу (рисунок 3.13) приписує переважну частину позитивного настрою слову «greatest». Але не варто забувати, що в

моделях, які враховують послідовність, одне і те ж слово може використовуватися в іншому контексті, маючи іншу вагу на рішення.

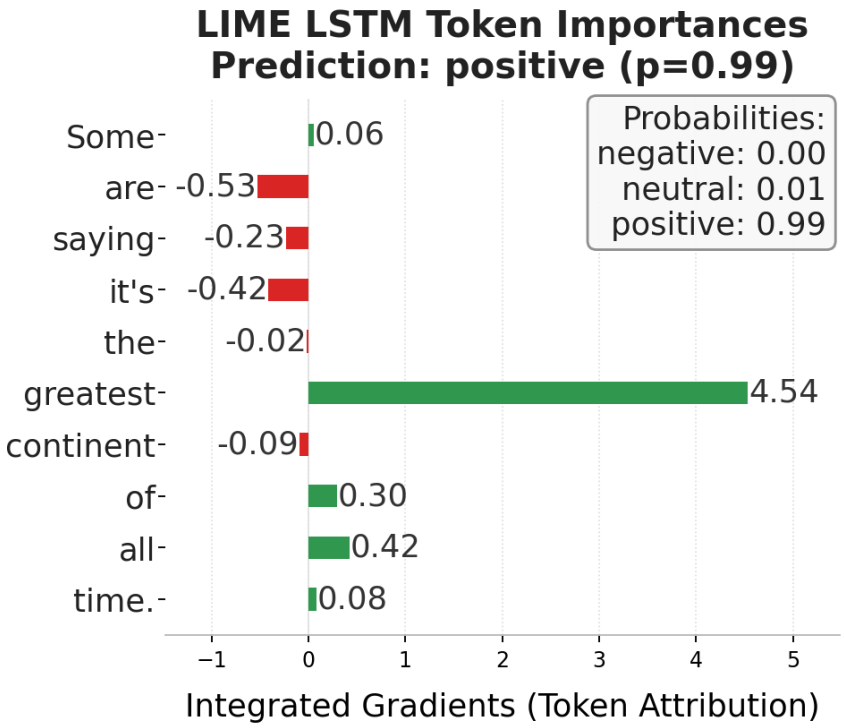

Рисунок 4.2 – Пояснення LIME для прогнозування моделей LSTM

Розпізнавання емоцій розширює цей процес до багатокласового завдання, класифікуючи текст за одним із шести емоційних ярликів. Продуктивність і достовірність різних архітектур моделей тут може значно відрізнятися. Для простого тексту «This is great to hear.» моделі DistilBERT не складає труднощів правильно визначити емоцію. Що можна детальніше розглянути в аналізі важливості слів (рисунок 3.14), токени «great» і «this» були найбільш впливовими токенами для рішення в даному контексті. Однак більш прості моделі можуть зіткнутися з неоднозначністю. Автономна модель BiLSTM в ході експериментів часто видавала розпливчасті ймовірності по всіх класах емоцій, що вказує на недостатню достовірність.

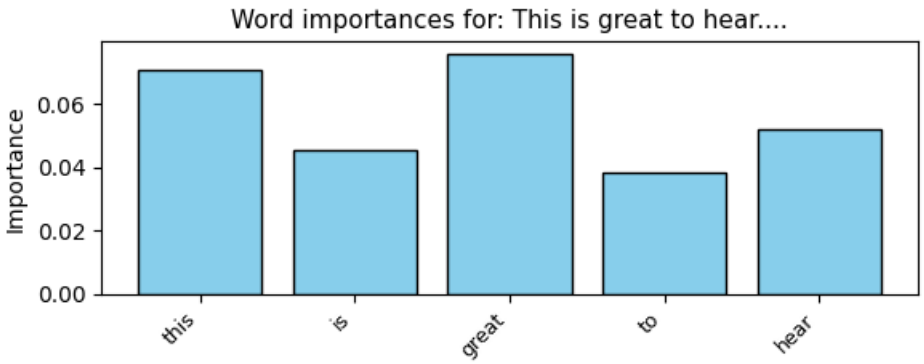

Рисунок 4.3 – Графік значущості слів для заданого текстового введення

Саме тут ансамблеві методи демонструють свою цінність. Для більш складного речення взято приклад: «Sadly, some european Americans like myself had to forgo the trip back home this summer.». Як видно на рисунку 3.15, для цього тексту компонент DistilBERT ансамблю з високим ступенем ймовірності передбачає «смуток», в той час як компонент TwitterRoBERTa дає більш змішаний прогноз, також показуючи деякі ознаки «любові» і «страху». Об’єднуючи ці голоси, ансамбль може прийти до більш тонкого і часто більш точного остаточного прогнозу «суму».

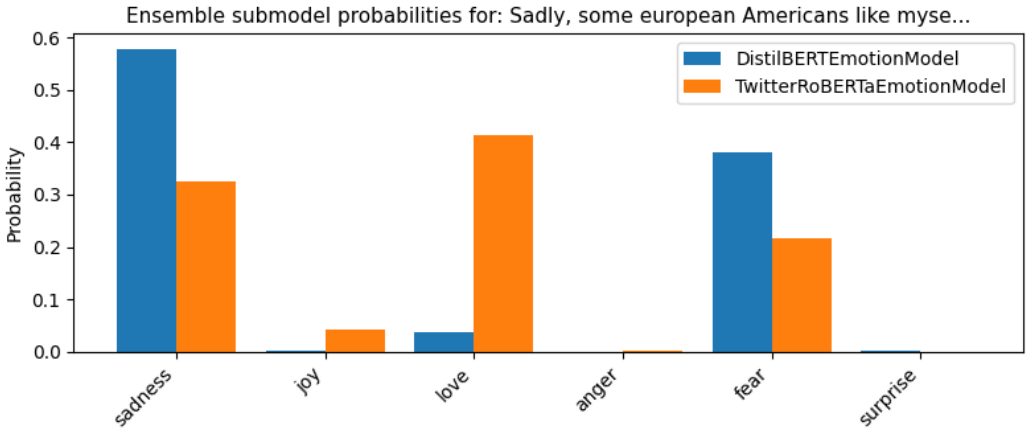

Рисунок 4.4 – Результати ансамблевого голосування

4.2 NER та ABSA на практичному рівні

Крім простої класифікації, для завдань, що вимагають структурованих результатів, таких як визначення конкретних об’єктів думки або іменованих сутностей, потрібні більш складні моделі. Ці моделі не просто класифікують весь текст, вони його аналізують.

NER служить основою для такого поглибленого аналізу. Він не виконує аналіз тональності, а виділяє «хто» і «що» в тексті. При отриманні рядка типу «Apple Inc. was founded by Steve Jobs in California.», модель NER обробляє текст і видає послідовність позначених сутностей (таблиця 3.3).

Таблиця 4.1 – Результат ідентифікації сутностей моделлю NER

Token	Tag	Entity
Apple	B-ORG	START ORG
Inc.	I-ORG	END ORG
was	O	
founded	O	
by	O	
Steve	B-PER	START PER
Jobs	I-PER	END PER
in	O	
California.	B-GEO	END GEO

Результат ідентифікує «Apple Inc.» як організацію (ORG), «Стів Джобс» як особистість (PER) і «Каліфорнія» як геополітичну сутність (GEO), часто використовуючи схему маркування BIO (Beginning, Inside, Outside) для позначення меж кожної сутності. Цей структурований висновок модель ABSA може потім використовуватися для пошуку думок, конкретно пов’язаних з сутністю «Apple Inc.», ідентифікованою моделлю

NER, що дозволяє проводити набагато більш цілеспрямований і всебічний аналіз.

Це безпосередньо веде до аспектного аналізу настроїв (ABSA), мета якого – надати найбільш деталізовану інформацію. Конвеєр прогнозування ABSA приймає коментарі користувачів і видає структуровані пари «аспект-настрій». Наприклад, для пропозиції типу «Їжа була відмінна, але обслуговування було повільним» ідеальна модель ABSA від початку до кінця видає результат [(«їжа», «позитивний»), («обслуговування», «негативний»)]. Це вимагає складної внутрішньої логіки для зіставлення вихідних результатів моделі, таких як теги послідовності і ймовірності класифікації, з інтерпретованими парами. Переваги комплексної моделі ABSA стають очевидними при порівнянні різних підходів. Для більш довгого коментаря, такого як «Не дивно. Замість того, щоб здійснити довгоочікувану поїздку до Нью-Йорка на мій 40-й день народження, я взяв гроші і минулого місяця поїхав на тиждень до Сицилії, а в кінці липня поїду в якесь інше місце в Європі. У нас фантастичний континент. Найкращий!», модель класифікації аспектів і тональності (ASC), що працює самостійно, може присвоїти тільки один «позитивний» ярлик всьому тексту. Хоча це і вірно, але це грубий аналіз. На відміну від цього, інтегрована модель, яка спочатку виконує вилучення термінів аспектів (ATE) перед класифікацією (ATE+ASC), може надати набагато багатший результат. Для того ж тексту вона визначить конкретні аспекти, про які йдеться – «Нью-Йорк», «Сицилія», «Європа» і «континент» – і присвоїть кожному з них «позитивний» відтінок. Це значний стрибок у глибині аналізу, перехід від загального відчуття до детальної карти конкретних думок, що є кінцевою метою тонкого аналізу настроїв.

ВИСНОВКИ

У цій кваліфікаційній роботі було досліджено обчислювальні методи аналізу думок з метою застосування у конкретних цілях або створення основ для системи, щоб оцінювати вплив громадської думки.

Головний висновок полягає в тому, що одного підходу до аналізу настроїв недостатньо. Для формальних новин потрібні великі моделі Transformer (наприклад, BERT), щоб зрозуміти складний контекст, тоді як для неформальних соціальних медіа потрібні спеціалізовані моделі (наприклад, Twitter-RoBERTa) та вдосконалене очищення даних. Для складних завдань, таких як розпізнавання емоцій, поєднання різних моделей у ансамблі забезпечує більш надійні результати. Щоб отримати детальну інформацію, аналіз повинен виходити за межі простих позитивних/негативних оцінок і знаходити думки щодо конкретних особливостей. Для цього спочатку необхідно визначити ключові сутності, що підтверджує необхідність модульного, багатоетапного процесу. Це дослідження досягає своєї мети, розбиваючи велику проблему «оцінки громадської думки» на менші частини. визначаючи основні вимоги до повноцінних систем: повинні бути багаторівневою, обробляти різні типи даних із використанням спеціального попереднього оброблення та використовувати спеціалізовані моделі для кожного завдання, забезпечуючи чіткий план моніторингу громадського дискурсу.

Це дослідження успішно вирішило складну проблему оцінки громадської думки, розробивши надійні, спеціалізовані методи для кожної визначеної підпроблеми. Воно створює всеосяжний план багаторівневої аналітичної системи, призначеної для обробки різноманітних типів даних із застосуванням спеціальної попередньої обробки та адаптованих моделей. Ці фундаментальні висновки забезпечують чіткий і багатообіцяючий шлях до кінцевої мети – комп’ютерної оцінки реального впливу суспільних настроїв.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1) Balahur A., Mihalcea R., Montoyo A. Computational approaches to subjectivity and sentiment analysis: present and envisaged methods and applications. *Computer speech & language*. 2014. Vol. 28, no. 1. P. 1–6. URL: https://doi.org/10.1016/j.csl.2013.09.003 (date of access: 21.04.2025).

2) Saad S., Saberi B. Sentiment analysis or opinion mining: a review. *International journal on advanced science, engineering and information technology*. 2017. Vol. 7, no. 5. P. 1660. URL: https://doi.org/10.18517/ijaseit.7.4.2137 (date of access: 22.04.2025).

3) Hatzivassiloglou V., McKeown K. R. Predicting the semantic orientation of adjectives. *The eighth conference*, Madrid, Spain, 7–12 July 1997. Morristown, NJ, USA, 1997. URL: https://doi.org/10.3115/979617.979640 (date of access: 21.04.2025).

4) Dave K., Lawrence S., Pennock D. M. Mining the peanut gallery. *The twelfth international conference*, Budapest, Hungary, 20–24 May 2003. New York, New York, USA, 2003. URL: https://doi.org/10.1145/775152.775226 (date of access: 21.05.2025).

5) Nasukawa T., Yi J. Sentiment analysis. *The international conference*, Sanibel Island, FL, USA, 23–25 October 2003. New York, New York, USA, 2003. URL: https://doi.org/10.1145/945645.945658 (date of access: 22.05.2025).

6) Popowich F., Moghaddam S. Opinion polarity identification through adjectives. 21 November 2010. *arXiv:1011.4623*.

7) Jang, J. W., Kwon, D. K., & Kim, J. E. (2018). Identifying opinion holders and opinion objects utilizing linguistic cues. *Institute of British and American Studies, 42,* 147–171. https://doi.org/10.25093/jbas.2018.42.147 (date of access: 24.05.2025).

8) Esuli A., Sebastiani F. Determining term subjectivity and term orientation for opinion mining. In Proceedings EACL-06, the 11rd Conference of

the European Chapter of the Association for Computational Linguistics, Trento, IT

9) Lexicon-Based methods for sentiment analysis / M. Taboada et al. *Computational linguistics*. 2011. Vol. 37, no. 2. P. 267–307. URL: https://doi.org/10.1162/coli_a_00049 (date of access: 26.05.2025).

10) 10.1. Long Short-Term Memory (LSTM) – Dive into Deep Learning 1.0.3 documentation. *Dive into Deep Learning – Dive into Deep Learning 1.0.3 documentation*. URL: https://d2l.ai/chapter_recurrent-modern/lstm.html (date of access: 26.06.2025).

11) Lagrari F.-E., Elkettani Y. Traditional and deep learning approaches for sentiment analysis: a survey. *Advances in science, technology and engineering systems journal*. 2021. Vol. 6, no. 5. P. 1–7. URL: https://doi.org/10.25046/aj060501 (date of access: 26.05.2025).

12) Attention Is All You Need / A. Vaswani et al. 2017. URL: https:// https://arxiv.org/abs/1706.03762 .

13) Understanding NLP algorithms: the masked language model. *Coursera*. URL: https://www.coursera.org/articles/masked-language-model (date of access: 27.05.2025).

14) Hu M., Liu B. Mining and summarizing customer reviews. The 2004 ACM SIGKDD international conference, Seattle, WA, USA, 22–25 August 2004. New York, New York, USA, 2004. URL: https://doi.org/10.1145/1014052.1014073 (date of access: 28.05.2025).

15) Popescu A.-M., Etzioni O. Extracting product features and opinions from reviews. The conference, Vancouver, British Columbia, Canada, 6–8 October 2005. Morristown, NJ, USA, 2005. URL: https://doi.org/10.3115/1220575.1220618 (date of access: 27.05.2025).

16) Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques / J. Yi et al. Third IEEE international conference on data mining, Melbourne, FL, USA. URL: https://doi.org/10.1109/icdm.2003.1250949 (date of access: 28.05.2025).

17) An unsupervised neural attention model for aspect extraction / R. He et al. *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: long papers)*, Vancouver, Canada. Stroudsburg, PA, USA, 2017. URL: https://doi.org/10.18653/v1/p17-1036 (date of access: 28.05.2025).

18) Jakob N., Gurevych I. Extracting opinion targets in a single and cross-domain setting with conditional random fields. Proceedings of the 2010 conference on empirical methods in natural language processing / ed. by Hang Li, L. Màrquez. Cambridge, MA, 2010.

19) Zhang M., Zhang Y., Vo D. T. Neural networks for open domain targeted sentiment. Proceedings of the 2015 conference on empirical methods in natural language processing, Lisbon, Portugal. Stroudsburg, PA, USA, 2015. URL: https://doi.org/10.18653/v1/d15-1073 (date of access: 28.05.2025).

20) Kessler J., Nicolov N. Targeting sentiment expressions through supervised ranking of linguistic configurations. *Proceedings of the international AAAI conference on web and social media*. 2009. Vol. 3, no. 1. P. 90–97. URL: https://doi.org/10.1609/icwsm.v3i1.13948 (date of access: 29.05.2025).

21) Dave K., Lawrence S., Pennock D. M. Mining the peanut gallery. *The twelfth international conference*, Budapest, Hungary, 20–24 May 2003. New York, New York, USA, 2003. URL: https://doi.org/10.1145/775152.775226 (date of access: 05.05.2025).

22) Jin W., Ho H. H., Srihari R. K. OpinionMiner. *The 15th ACM SIGKDD international conference*, Paris, France, 28 June – 1 July 2009. New York, New York, USA, 2009. URL: https://doi.org/10.1145/1557019.1557148 (date of access: 29.05.2025).

23) A deeper look into sarcastic tweets using deep convolutional neural networks / S. Poria et al. 2016. URL: https://arxiv.org/abs/1610.08815.

24) Mohammad S. M., Sobhani P., Kiritchenko S. Stance and sentiment in tweets. *ACM transactions on internet technology*. 2017. Vol. 17, no. 3. P. 1–23. URL: https://doi.org/10.1145/3003433 (date of access: 30.05.2025).

25) Salameh M., Mohammad S., Kiritchenko S. Sentiment after translation: a case-study on arabic social media posts. *Proceedings of the 2015 conference of the north american chapter of the association for computational linguistics: human language technologies*, Denver, Colorado. Stroudsburg, PA, USA, 2015. URL: https://doi.org/10.3115/v1/n15-1078 (date of access: 30.05.2025).

26) A multilayer perceptron based ensemble technique for fine-grained financial sentiment analysis / M. S. Akhtar et al. *Proceedings of the 2017 conference on empirical methods in natural language processing*, Copenhagen, Denmark. Stroudsburg, PA, USA, 2017. URL: https://doi.org/10.18653/v1/d17-1057 (date of access: 01.06.2025).

27) Jindal N., Liu B. Opinion spam and analysis. *The international conference*, Palo Alto, California, USA, 11–12 February 2008. New York, New York, USA, 2008. URL: https://doi.org/10.1145/1341531.1341560 (date of access: 01.06.2025).

28) An approach to the selection of behavior patterns autonomous intelligent mobile systems / O. Zolotukhin et al. *2021 IEEE 8th international conference on problems of infocommunications, science and technology (PIC S&T)*, Kharkiv, Ukraine, 5–7 October 2021. 2021. URL: https://doi.org/10.1109/picst54195.2021.9772110 (date of access: 01.06.2025).

29) Zolotukhin, O., Filatov, V., Yerokhin, A., Kudryavtseva, M.: The methods for the prediction of climate control indicators in the Internet of Things systems. CEUR Workshop Proc. (2021). URL: https://doi.org/10.5281/zenodo.14526027 (date of access: 01.06.2025).

30) Methods of intellectual analysis of processes in medical information systems / V. O. Filatov et al. *Information extraction and processing*. 2020. Vol. 2020, no. 48. P. 92–98. URL: https://doi.org/10.15407/vidbir2020.48.092 (date of access: 01.06.2025).