

Хмельницький національний університет  
Факультет інформаційних технологій  
Кафедра комп'ютерних наук

## КВАЛІФІКАЦІЙНА РОБОТА

на тему Метод оцінювання відповідності хештегів коротким текстам засобами обробки природної мови

Рівень вищої освіти другий (магістерський)

Галузь знань 12 – Інформаційні технології


Шифр і назва галузі знань

Спеціальність 122 – Комп'ютерні науки


Шифр і назва спеціальності

Освітня програма Комп'ютерні науки

Назва освітньої програми

Виконала: студентка 2 курсу, група КНм-24-1  Ольга ТРОХИМЧУК  
Курс, група виконавця Підпис Ім'я, ПРІЗВИЩЕ

Керівник: к.т.н., доцент кафедри КН  Олександр ПАСІЧНИК  
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

Нормоконтроль: к.т.н., доцент кафедри КН  Руслан БАГРІЙ  
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

До захисту допускаю:

Зав. кафедри КН, д.т.н., професор



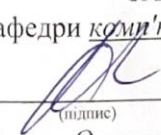
Олександр БАРМАК  
Ім'я, ПРІЗВИЩЕ

16 грудня 2025 р.

Хмельницький 2025

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
Факультет інформаційних технологій  
Кафедра комп'ютерних наук  
Освітній ступінь магістр  
Галузь знань 12 – Інформаційні технології  
Спеціальність 122 – Комп'ютерні науки

ЗАТВЕРДЖУЮ  
Завідувач кафедри комп'ютерних наук


  
(підпис)  
д.т.н., професор Олександр БАРМАК  
« 28 » серпня 2025 року


**ЗАВДАННЯ  
НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА**

1. Тема кваліфікаційної роботи магістра: «Метод оцінювання відповідності хештегів коротким текстам засобами обробки природної мови».
2. Завдання видано студенту Ользі ТРОХИМЧУК  
(Ім'я, ПРІЗВИЩЕ)
3. Керівник роботи доцент кафедри КН Олександр Пасічник  
(Ім'я, ПРІЗВИЩЕ)
4. Затверджені наказом університету від «25» серпня 2025 р. № 65.
5. Дата видачі завдання студенту: «28» серпня 2025 р.
6. Зміст пояснювальної записки (перелік задач) та вихідні дані:  
Метою кваліфікаційної роботи є підвищення достовірності та інтерпретованості оцінювання релевантності хештегів коротким текстовим повідомленням. Для досягнення мети слід вирішити такі задачі: аналіз наукових праць з обробки природної мови; формалізацію задачі оцінювання релевантності хештегів та побудову відповідної нейромережевої моделі; розроблення та реалізацію алгоритмів формування корпусу пар «текст- хештег-мітка» на основі реальних даних соціальних мереж; нейромережеве моделювання з використанням трансформерних архітектур для класифікації відповідності хештегів; застосування методів пояснюваного штучного інтелекту для обчислення атрибутів важливості фрагментів тексту й оцінювання їх узгодженості зі змістовим ядром хештегу; експериментальне дослідження запропонованого методу на тестових вибірках.

7. Календарний план виконання кваліфікаційної роботи:

№	Назва етапів (розділів) кваліфікаційної роботи бакалавра	Термін виконання	Примітка
1	Вибір напрямку дослідження та узгодження теми кваліфікаційної роботи з керівником, складання календарного графіка виконання роботи	вересень 2025	Виконано
2	Ознайомлення з предметною областю, аналіз існуючих методів і моделей, формулювання мети та завдань дослідження, визначення об'єкта й предмета дослідження	вересень 2025	Виконано
3	Розробка методу чи моделі для вирішення обраного завдання, опис архітектури рішення	жовтень 2025	Виконано
4	Програмна реалізація методу чи моделі	жовтень 2025	Виконано
5	Дослідження ефективності та експериментальна перевірка результатів, порівняння з відомими підходами	листопад 2025	Виконано
6	Написання пояснювальної записки, оформлення відповідно до вимог, врахування зауважень керівника	листопад 2025	Виконано
7	Підготовка презентаційних матеріалів та попередній захист	листопад 2025	Виконано
8	Перевірка пояснювальної записки на відповідність вимогам оформлення (нормоконтроль) та перевірка на академічну доброчесність. Отримання відгуку керівника та рецензії.	грудень 2025	Виконано
9	Публічний захист кваліфікаційної роботи	грудень 2025	Виконано

Виконавець: студентка гр. КНм-24-1  Ольга ТРОХИМЧУК  
Група виконавця Підпис Ім'я, ПРІЗВИЩЕ

Керівник: к.т.н., доц. каф. КН  Олександр ПАСІЧНИК  
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

## Реферат

Кваліфікаційна робота магістра присвячена вирішенню науково-технічної задачі автоматизованого оцінювання відповідності хештегів коротким текстам засобами обробки природної мови. Запропонований у роботі підхід до пояснюваного оцінювання релевантності хештегів, що інтегрує класифікаційні можливості сучасних трансформерних мовних моделей із метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу, спрямований на подолання зазначеного дисбалансу. Він дає змогу не лише кількісно оцінювати відповідність хештегів коротким текстам, а й якісно контролювати, на яких саме словах та фрагментах тексту ґрунтуються рішення моделі. Це підвищує довіру до результатів, створює підґрунтя для більш відповідального використання нейромережових технологій у аналізі соціальних мереж і відкриває можливості для виявлення нових, семантично коректних, але рідковживаних хештегів. Таким чином, дослідження є актуальним як з точки зору розвитку теорії пояснюваного штучного інтелекту в задачах текстової аналітики, так і з огляду на практичні потреби індустрії цифрового контенту.

**Актуальність теми.** Актуальність теми зумовлена стрімким зростанням обсягів коротких текстових повідомлень у соціальних мережах та цифрових комунікаційних сервісах, де хештеги фактично виконують роль спрощеної системи індексації, навігації й тематичної маршрутизації контенту. Від якості підбору та узгодженості хештегів із змістом повідомлення залежать видимість публікацій, коректність аналітики залученості аудиторії, ефективність маркетингових кампаній, репрезентативність досліджень громадської думки й інформаційних потоків. Нерелевантні, випадкові або маніпулятивні хештеги призводять до засмічення інформаційного простору, спотворення статистики переглядів і взаємодій, ускладнюють тематичний пошук і побудову достовірних моделей поведінки користувачів. За умов інформаційного перенасичення та високої динаміки онлайн-контенту ручний контроль за коректністю хештегів є нереалістичним, що зумовлює потребу в автоматизованих засобах їх оцінювання та відбору.

Попри значний прогрес у галузі обробки природної мови та застосування глибинних нейромережових моделей для класифікації текстів і рекомендації тегів, більшість наявних рішень орієнтовані на максимізацію прогностичної точності без

належного врахування прозорості та інтерпретованості рішень. Типові підходи ґрунтуються на співпадіннях у навчальній вибірці або близькості векторних подань тексту й хештегу, не дозволяючи перевірити, чи спирається модель на семантично обґрунтовані фрагменти тексту, чи лише відтворює випадкові кореляції. За таких умов існує ризик підкріплення хибних або надлишкових хештегів, що формально отримують високу оцінку моделі, але не мають прозорого зв'язку зі змістовим ядром повідомлення. Це обмежує довіру до нейромережових систем аналізу соціальних медіа, ускладнює їх використання в наукових дослідженнях, маркетингових рішеннях та сервісах контент-модерації.

**Мета і задачі роботи.** Метою кваліфікаційної роботи магістра є підвищення достовірності та інтерпретованості оцінювання релевантності хештегів коротким текстовим повідомленням шляхом використання пояснюваного нейромережового підходу, який поєднує трансформерну класифікацію з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

Для досягнення мети слід вирішити такі задачі:

- виконати аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах;
- здійснити аналіз сучасних нейромережових та лінгвістичних підходів до класифікації коротких текстів, рекомендації тегів і пояснюваного аналізу рішень моделей обробки природної мови;
- проаналізувати наявні набори даних соціальних мереж, обґрунтувати вибір датасету для формування пар «текст – хештег – мітка» та визначити показники якості оцінювання релевантності;
- розробити нейромережовий метод пояснюваного оцінювання релевантності хештегів, що поєднує трансформерну класифікацію з метриками узгодженості важливих фрагментів тексту зі змістовим ядром хештегу;
- виконати підготовку та анотування вибірки коротких текстів і хештегів на основі даних соціальних мереж з формуванням позитивних і негативних прикладів для донавчання моделі;
- здійснити програмну реалізацію запропонованого методу з використанням сучасних засобів глибинного навчання та пояснюваного штучного інтелекту;

– провести експериментальні дослідження ефективності запропонованого методу порівняно з базовими підходами, що ґрунтуються лише на класифікаційних оцінках або векторних поданнях;

– виконати аналіз отриманих результатів, оцінити вплив пояснюваних метрик на якість відсікання нерелевантних хештегів і виявлення семантично коректних, але рідкісних тегів, та сформулювати практичні рекомендації щодо застосування методу.

**Об’єкт дослідження.** Процес автоматизованого оцінювання релевантності хештегів для коротких текстових повідомлень у цифрових комунікаційних середовищах.

**Предмет дослідження.** Нейромережеві моделі та методи пояснюваного оцінювання релевантності хештегів коротким текстам, що поєднує класифікаційні моделі обробки природної мови з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

**Методи дослідження:** методи математичної статистики (оцінювання параметрів моделей, перевірка статистичної значущості відмінностей, аналіз узагальнювальних та помилкових характеристик), нейромережеві методи та методи машинного навчання (побудова та навчання трансформерних моделей для бінарної класифікації відповідності хештегів текстам, налаштування гіперпараметрів, оптимізація функції втрат), методи обробки природної мови (токенізація, нормалізація, векторизація текстів, формування простору ознак для коротких повідомлень та хештегів), а також методи пояснюваного штучного інтелекту (застосування інтегрованих градієнтів та споріднених підходів для обчислення важливості окремих фрагментів тексту). Додатково використовуються методи комп’ютерного моделювання та чисельного експерименту для дослідження поведінки нейромережевих моделей на реальних даних соціальних мереж.

**Наукова новизна одержаних результатів** полягає у подальшому розвитку методів оцінювання відповідності хештегів коротким текстам шляхом поєднання нейромережевої класифікації з показниками пояснюваності. Запропоновано підхід, у межах якого релевантність хештегу визначається не лише за ймовірністю віднесення пари «текст-хештег» до класу «відповідає / не відповідає», отриманою трансформерною мовною моделлю, а й за мірою узгодженості внутрішніх пояснень моделі із семантичним ядром самого хештегу. На відміну від традиційних підходів,

що спираються виключно на відстані в просторах векторних подань або на бінарне рішення класифікатора, введено інтегральний показник релевантності, який синтезує оцінку моделі та частку найбільш інформативних для неї фрагментів тексту, лексично чи словотвірно пов'язаних зі змістовим ядром хештегу.

Додаткову складову новизни становить методика автоматизованого формування корпусу пар «підпис-хештег-мітка» на основі реальних дописів соціальної мережі з використанням негативної вибірки, а також подальший аналіз режимів роботи моделі, що поєднують високу або низьку впевненість класифікатора з високою або низькою узгодженістю пояснень. Показано, що запропонована пояснювана метрика дає змогу не лише відсікати нерелевантні та шумові хештеги, але й виявляти семантично коректні, однак рідкісні або відсутні у навчальній вибірці хештеги, які базовий класифікатор схильний відхиляти. У сукупності це формує новий підхід до оцінювання якості хештегів, у якому вона описується через прозорість та інтерпретованість рішень нейромережевої моделі, а не лише через її прогностичну точність.

**Апробація результатів кваліфікаційної роботи магістра та публікації.** Основні наукові й практичні результати роботи доповідались у доповіді «Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP» на XVII Всеукраїнській науково-практичній конференції «Актуальні проблеми комп'ютерних наук АПКН-2025» (м.Хмельницький) 14-15 листопада 2025 року.

**Структура та обсяг роботи.** Кваліфікаційна робота магістра складається з реферату, завдання, змісту, переліку скорочень, вступу, 4-х розділів, висновків, переліку посилань із 66 найменувань та 9 додатків. Обсяг основного тексту кваліфікаційної роботи магістра становить 86 сторінок. У роботі наведено 26 рисунків і 5 таблиць.

**Ключові слова:** хештег, релевантність, NLP, бінарна класифікація, пояснюваний штучний інтелект, BERT.

## Зміст

Перелік скорочень .....	4
Вступ.....	5
РОЗДІЛ 1 Дослідження сучасного стану автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням.....	9
1.1 Аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням.....	9
1.2 Аналіз сучасних нейромережових та лінгвістичних підходів до класифікації коротких текстів .....	14
1.3 Аналіз наукових праць для оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах .....	17
1.4 Постановка задачі.....	20
РОЗДІЛ 2 Метод оцінювання відповідності хештегів коротким текстам засобами NLP .....	22
2.1 Ідея підходу до оцінювання відповідності хештегів коротким текстам засобами NLP.....	22
2.2 Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP.....	23
2.3 Спосіб оцінювання кандидатних хештегів відповідно до повідомлення.....	25
2.4 Спосіб пояснюваного об'єднання оцінок та формування інтегральної міри релевантності .....	26
2.5 Формування та підготовка навчальних даних .....	28
Висновки до розділу 2 .....	30
РОЗДІЛ 3 Проектування інтелектуальної системи оцінювання відповідності хештегів коротким текстам .....	33
3.1 Вибір засобів розробки інтелектуальної системи.....	33
3.2 Проектування структури інтелектуальної системи .....	35
3.3 Проектування елементів інтелектуальної системи.....	38
3.4 Метрики нейромережевого оцінювання релевантності хештегів.....	44

Висновки до розділу 3 .....	46
РОЗДІЛ 4 Експериментальне дослідження методу оцінювання відповідності хештегів коротким текстам засобами NLP .....	49
4.1 Структура компонентів інтелектуальної системи оцінювання відповідності хештегів коротким текстам .....	49
4.2 Особливості створення інтелектуальної системи .....	53
4.3 Особливості тестування розробленої інтелектуальної системи.....	58
4.4 Експериментальна установка.....	66
4.5 Дослідження методу .....	68
Висновки до розділу 4 .....	75
Загальні висновки.....	76
Перелік посилань.....	79
Додатки	

## Перелік скорочень

Скорочення, термін, позначення	Пояснення
Word2Vec	Word to Vector (модель векторного подання слів)
GloVe	Global Vectors (глобальні вектори слів)
CNN	Згорткова нейронна мережа (Convolutional Neural Network)
LSTM	Довгострокова короткочасна пам'ять (Long Short-Term Memory)
GRU	Об'єднана рекурентна одиниця (Gated Recurrent Unit)
BERT	Bidirectional Encoder Representations from Transformers (двонаправлені подання з трансформерів)
fastText	Модель векторного подання слів від Facebook
TRANSIT	Метод рекомендації постів у соціальних мережах (назва моделі)
#REval	Семантична система оцінювання релевантності хештегів
COVID-19	Коронавірусна хвороба 2019 року

## Вступ

**Актуальність теми.** Актуальність теми зумовлена стрімким зростанням обсягів коротких текстових повідомлень у соціальних мережах та цифрових комунікаційних сервісах, де хештеги фактично виконують роль спрощеної системи індексації, навігації й тематичної маршрутизації контенту. Від якості підбору та узгодженості хештегів із змістом повідомлення залежать видимість публікацій, коректність аналітики залученості аудиторії, ефективність маркетингових кампаній, репрезентативність досліджень громадської думки й інформаційних потоків. Нерелевантні, випадкові або маніпулятивні хештеги призводять до засмічення інформаційного простору, спотворення статистики переглядів і взаємодій, ускладнюють тематичний пошук і побудову достовірних моделей поведінки користувачів. За умов інформаційного перенасичення та високої динаміки онлайн-контенту ручний контроль за коректністю хештегів є нереалістичним, що зумовлює потребу в автоматизованих засобах їх оцінювання та відбору.

Попри значний прогрес у галузі обробки природної мови та застосування глибинних нейромережових моделей для класифікації текстів і рекомендації тегів, більшість наявних рішень орієнтовані на максимізацію прогностичної точності без належного врахування прозорості та інтерпретованості рішень. Типові підходи ґрунтуються на співпадіннях у навчальній вибірці або близькості векторних подань тексту й хештегу, не дозволяючи перевірити, чи спирається модель на семантично обґрунтовані фрагменти тексту, чи лише відтворює випадкові кореляції. За таких умов існує ризик підкріплення хибних або надлишкових хештегів, що формально отримують високу оцінку моделі, але не мають прозорого зв'язку зі змістовим ядром повідомлення. Це обмежує довіру до нейромережових систем аналізу соціальних медіа, ускладнює їх використання в наукових дослідженнях, маркетингових рішеннях та сервісах контент-модерації.

**Мета і задачі роботи.** Метою кваліфікаційної роботи магістра є підвищення достовірності та інтерпретованості оцінювання релевантності хештегів коротким текстовим повідомленням шляхом використання пояснюваного нейромережового

підходу, який поєднує трансформерну класифікацію з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

Для досягнення мети слід вирішити такі задачі:

- виконати аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах;

- здійснити огляд сучасних нейромережових та лінгвістичних підходів до класифікації коротких текстів, рекомендації тегів і пояснюваного аналізу рішень моделей обробки природної мови;

- проаналізувати наявні набори даних соціальних мереж, обґрунтувати вибір датасету для формування пар «текст – хештег – мітка» та визначити показники якості оцінювання релевантності;

- розробити нейромережовий метод пояснюваного оцінювання релевантності хештегів, що поєднує трансформерну класифікацію з метриками узгодженості важливих фрагментів тексту зі змістовим ядром хештегу;

- виконати підготовку та анотування вибірки коротких текстів і хештегів на основі даних соціальних мереж з формуванням позитивних і негативних прикладів для донавчання моделі;

- здійснити програмну реалізацію запропонованого методу з використанням сучасних засобів глибинного навчання та пояснюваного штучного інтелекту;

- провести експериментальні дослідження ефективності запропонованого методу порівняно з базовими підходами, що ґрунтуються лише на класифікаційних оцінках або векторних поданнях;

- виконати аналіз отриманих результатів, оцінити вплив пояснюваних метрик на якість відсікання нерелевантних хештегів і виявлення семантично коректних, але рідкісних тегів, та сформулювати практичні рекомендації щодо застосування методу.

**Об’єкт дослідження.** Процес автоматизованого оцінювання релевантності хештегів для коротких текстових повідомлень у цифрових комунікаційних середовищах.

**Предмет дослідження.** Нейромережеві моделі та методи пояснюваного оцінювання релевантності хештегів коротким текстам, що поєднує класифікаційні моделі обробки природної мови з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу

**Методи дослідження:** методи математичної статистики (оцінювання параметрів моделей, перевірка статистичної значущості відмінностей, аналіз узагальнювальних та помилкових характеристик), нейромережеві методи та методи машинного навчання (побудова та навчання трансформерних моделей для бінарної класифікації відповідності хештегів текстам, налаштування гіперпараметрів, оптимізація функції втрат), методи обробки природної мови (токенізація, нормалізація, векторизація текстів, формування простору ознак для коротких повідомлень та хештегів), а також методи пояснюваного штучного інтелекту (застосування інтегрованих градієнтів та споріднених підходів для обчислення важливості окремих фрагментів тексту). Додатково використовуються методи комп'ютерного моделювання та чисельного експерименту для дослідження поведінки нейромережевих моделей на реальних даних соціальних мереж.

**Наукова новизна одержаних результатів** полягає у подальшому розвитку методів оцінювання відповідності хештегів коротким текстам шляхом поєднання нейромережевої класифікації з показниками пояснюваності. Запропоновано підхід, у межах якого релевантність хештегу визначається не лише за ймовірністю віднесення пари «текст-хештег» до класу «відповідає / не відповідає», отриманою трансформерною мовною моделлю, а й за мірою узгодженості внутрішніх пояснень моделі із семантичним ядром самого хештегу. На відміну від традиційних підходів, що спираються виключно на відстані в просторах векторних подань або на бінарне рішення класифікатора, введено інтегральний показник релевантності, який синтезує оцінку моделі та частку найбільш інформативних для неї фрагментів тексту, лексично чи словотвірно пов'язаних зі змістовим ядром хештегу.

Додаткову складову новизни становить методика автоматизованого формування корпусу пар «підпис-хештег-мітка» на основі реальних дописів соціальної мережі з використанням негативної вибірки, а також подальший аналіз

режимів роботи моделі, що поєднують високу або низьку впевненість класифікатора з високою або низькою узгодженістю пояснень. Показано, що запропонована пояснювана метрика дає змогу не лише відсікати нерелевантні та шумові хештеги, але й виявляти семантично коректні, однак рідкісні або відсутні у навчальній вибірці хештеги, які базовий класифікатор схильний відхиляти. У сукупності це формує новий підхід до оцінювання якості хештегів, у якому вона описується через прозорість та інтерпретованість рішень нейромережевої моделі, а не лише через її прогностичну точність.

#### **Апробація результатів кваліфікаційної роботи магістра та публікації.**

Основні наукові й практичні результати роботи доповідались у доповіді «Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP» на XVII Всеукраїнській науково-практичній конференції «Актуальні проблеми комп'ютерних наук АПКН-2024» (м.Хмельницький) 14-15 листопада 2025 року.

**Структура та обсяг роботи.** Кваліфікаційна робота магістра складається з реферату, завдання, змісту, переліку скорочень, вступу, 4-х розділів, висновків, переліку посилань із 66 найменувань та 9 додатків. Обсяг основного тексту кваліфікаційної роботи магістра становить 86 сторінок. У роботі наведено 26 рисунків і 5 таблиць.

## **РОЗДІЛ 1 Дослідження сучасного стану автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням**

### **1.1 Аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням**

Проблема автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням належить до кола завдань, у яких технічні механізми позначення інформації тісно переплетені з практиками масової комунікації та колективного конструювання смислів [1]. Щоб коректно формалізувати цю задачу, необхідно спочатку розкрити сутність хештегу як соціально-технічного явища, простежити його походження, окреслити основні функції й типові відхилення від ідеальної ролі, а вже потім переходити до обговорення можливостей автоматизації оцінювання його доречності щодо конкретного тексту.

У найзагальнішому вигляді хештегом називають текстову мітку, що починається з символу «#» і використовується для позначення теми, події, кампанії, об'єкта чи певного смислового поля, із яким користувач прагне пов'язати своє повідомлення [2]. З формально-технічної точки зору це послідовність символів без пробілів, яку платформа розпізнає як спеціальний токен і автоматично перетворює на гіперпосилання або елемент індексації. З комунікативної точки зору хештег є елементом метамови: він не просто повторює слова з тексту, а конденсує в собі уявлення автора про те, про що цей допис, до якої ширшої теми, спільноти чи дискурсу він належить і в який потік повідомлень має бути включений [3].

Історично практика використання подібних міток виникла у спільнотах інтернет-чатів та групових обговорень, де символ «#» застосовувався для позначення каналів, технічних команд, тематичних гілок. Сучасне розуміння хештегу як універсального маркера теми сформувалося у мікроблогових сервісах і соціальних мережах, де зростання обсягів коротких повідомлень зробило необхідним простий та інтуїтивний механізм самостійного маркування контенту користувачами. Надалі ця практика була адаптована практично всіма основними платформами, а сам хештег

перетворився на один із ключових елементів користувацького інтерфейсу, аналітики та механізмів пошуку [4].

Роль хештегів у сучасних цифрових середовищах є багатовимірною. Для кінцевого користувача вони виступають способом тематичної навігації: перехід за тегом відкриває стрічку публікацій, пов'язаних із певною темою чи подією, і дозволяє увійти в дискусію, навіть якщо автори окремих дописів не знайомі між собою [5]. Хештеги підтримують самоідентифікацію та належність до спільноти, дозволяють маркувати підтримку соціальних кампаній, висловлювати позицію у суспільно-політичних питаннях, використовувати спільні меми й внутрішні коди групи [6]. Для контент-креаторів і брендів хештеги є інструментом управління видимістю, таргетування на певні аудиторії, побудови впізнаваних кампаній (рисунок 1.1). Для самих платформ це механізм індексації та кластеризації контенту, на основі якого будуються тематичні стрічки, підбір рекомендацій, виявлення трендових тем і подій. Для дослідників і аналітиків хештеги слугують маркерами тем, рухів і кампаній, які дозволяють реконструювати структуру дискурсу, відстежувати інформаційні хвилі та оцінювати реакції спільнот [7].



Рисунок 1.1 – Приклад текстового посту та хештегів до нього [8]

У ідеальному випадку хештег має бути змістово узгодженим із текстом повідомлення. Це означає, що тема, подія, об'єкт або явище, яке позначається тегом, дійсно присутні в дописі як суттєва складова його змісту, а не як побічна асоціація.

Наприклад, позначення допису про побудову фінансових дашбордів тегом, що вказує на фінансову аналітику, візуалізацію даних або конкретний інструмент, відповідає очікуваній функції тегу. Натомість позначення такого ж допису популярними, але загальними маркерами, які не розкривають теми, або тегами, пов'язаними з іншими сферами (подорожі, кулінарія, розваги), суперечить інтуїтивному розумінню релевантності [9].

На практиці спостерігається цілий спектр некоректних або низькорелевантних використань хештегів. Часто автори додають до будь-яких публікацій великі набори універсальних «модних» тегів, які мають високу частоту використання, але мало повідомляють про конкретний зміст. Такі загальні позначки можуть стосуватися «успіху», «натхнення», «життя», «мотивації» тощо і виконують радше роль емоційного маркування або засобу підвищення охоплення, ніж інформативних дескрипторів [10]. Інший тип некоректності – так зване «перехоплення» чужих тем, коли тег, пов'язаний із певною акцією, трагедією, суспільно важливою кампанією, використовується для промоції стороннього комерційного чи розважального контенту. У таких випадках хештег формально присутній, але його змістове поле й очікування аудиторії різко розходяться з реальною тематикою допису. Можливі також ситуації, коли тег використовується через непорозуміння: автор вважає, що певний маркер означає одне, тоді як у спільноті він усталено асоціюється з іншою темою [11].

Наслідки некоректного використання хештегів проявляються на кількох рівнях. Для окремого користувача це зниження якості пошуку та навігації: тематичні стрічки наповнюються контентом, який лише номінально позначений відповідним тегом, але не містить очікуваної інформації [12]. Для автора публікацій неправильні або надмірно загальні теги можуть призвести до зміщення аудиторії: допис потрапляє до користувачів, які не зацікавлені в його змісті, а ті, кому він би був корисний, його не бачать. Для платформ і аналітичних систем це виражається у «засміченні» показників, пов'язаних із популярністю тем, ефективністю кампаній, взаємодією з певними групами. Дослідження, що спираються на вибірки постів за тегами, можуть давати хибні висновки про структуру дискурсу, емоційні реакції або поведінку

користувачів, якщо частка нерелевантного чи маніпулятивного контенту в таких вибірках значна [13]. В аспекті модерації та управління інформаційним простором некоректні хештеги ускладнюють виявлення справді проблемного або чутливого контенту, маскують його серед маси «шумових» публікацій.

Поза межами автоматизованих систем людина оцінює релевантність хештегу, поєднуючи декілька когнітивних операцій. Спочатку здійснюється семантичне зчитування тексту повідомлення: виділяються ключові теми, об'єкти, дії, визначається, про що саме йдеться, що є головним, а що другорядним. Далі інтерпретується сам хештег: розкладається на складові слова або корені, співвідноситься зі знанням про те, як його зазвичай розуміють у спільноті, з якими темами пов'язують, у яких контекстах бачать [14]. Потім відбувається зіставлення цих двох уявлень: чи відповідає позначка тому, що насправді сказано в тексті, чи описує вона основну тему, важливий аспект, другорядну, але все ж присутню характеристику, чи, навпаки, привносить зовнішній смисловий план, що не впливає з повідомлення. Важливою є також оцінка ролі тегу у структурі допису: чи є він тематичним маркером, чи тільки сигналізує настрій, жанр («жарт», «мем»), або ж виконує суто технічну функцію, наприклад позначає рубрику блогу.

На рівні практичної діяльності редакцій, маркетингових команд або дослідницьких груп формуються неформальні критерії такої оцінки. Релевантним зазвичай вважається хештег, який називає те, про що йдеться в тексті, або той аспект ситуації, який автор явно прагне висвітлити; при цьому надмірне використання дуже загальних або модних тегів без прямого зв'язку зі змістом розглядається як небажана практика [15]. Часто вводяться внутрішні рекомендації: не використовувати маркери, що не згадуються або не впливають із тексту; уникати підміни тематичних тегів суто емоційними; не «перехоплювати» хештеги інших кампаній без прямого змістового зв'язку. Усе це базується на здоровому глузді, знанні предметної області та розумінні очікувань аудиторії [16].

Автоматизація оцінювання релевантності хештегів спирається на те, що як самі теги, так і короткі тексти мають низку формальних і змістових властивостей, придатних для обробки засобами інформаційних технологій. По-перше, хештеги

мають чітку синтаксичну форму: починаються з фіксованого символу, за яким іде послідовність знаків без пробілів. Це дозволяє надійно виявляти та виділяти їх у тексті за простими правилами, навіть без складної мовної обробки. По-друге, внутрішня структура тегу піддається декомпозиції: його можна розбити на окремі слова за допомогою виявлення змін регістру, підкреслювань, числових вставок, характерних шпарин у написанні, а потім зіставити з лексикою повідомлення, словниками, термінологічними базами. По-третє, і тексти, і хештеги є елементами великих масивів даних, де реєструються частоти їхнього вживання, співвживання одне з одним, позиціонування в часі та просторі. Таким чином, для кожного тегу можна емпірично оцінити, з якими словами, темами, подіями він зазвичай пов'язується, а для кожного тексту – які маркери для нього є типовими [17].

Ці властивості відкривають можливість для побудови алгоритмів, які, не претендуючи на повне відтворення людського розуміння, здатні наближено оцінювати доречність хештегів. На найпростішому рівні автоматизація може зводитися до лексичного зіставлення: система перевіряє, чи містить текст слова або словосполучення, що відповідають компонентам тегу або пов'язані з ними через словникові відношення [18]. Більш розвинені підходи можуть враховувати статистичні профілі: якщо хештег у більшості випадків з'являється разом із певною групою лексем або у межах конкретної тематичної категорії, а у даному тексті такі ознаки відсутні, доречно припускати можливу нерелевантність. Можуть використовуватися й онтологічні або класифікаційні структури, які пов'язують теги з ієрархією тем, подій, об'єктів, що дозволяє співвідносити зміст тексту з більш загальними чи конкретними поняттями [19].

З огляду на ці міркування, хештеги можна розглядати як структуровані маркери, що одночасно належать до мовного та даного контексту, і саме ця подвоєна природа робить їх зручними об'єктами для автоматизованої обробки. З одного боку, вони достатньо формалізовані, щоб їх можна було однозначно виділяти, зберігати, порівнювати, будувати на їх основі таблиці частот і співвживань, аналізувати динаміку. З іншого боку, вони несуть змістові навантаження, що прив'язують їх до тем, подій і практик реального життя, а отже, пов'язані з текстовими й контекстними

характеристиками публікацій. Саме на перетині цих двох вимірів, формальної структури та змістового поля, виникає можливість автоматизувати оцінювання релевантності, перетворюючи неформалізовані інтуїції користувачів на алгоритмічно оброблювані правила, показники та критерії.

## **1.2 Аналіз сучасних нейромережевих та лінгвістичних підходів до класифікації коротких текстів**

Класифікація коротких текстів: дописів у соціальних мережах, твітів, заголовків новин, повідомлень месенджерів, розглядається як окремий клас задач обробки природної мови, що має специфічні обмеження й виклики. На відміну від довших документів, короткі повідомлення містять небагато слів, характеризуються високою лексичною розрідженістю, насичені сленгом, аббревіатурами, емодзі, елементами мовної гри й сильно залежать від позатекстового контексту [20]. Оглядіві праці з текстової класифікації демонструють еволюцію методів від «мілких» статистичних моделей до глибоких нейромережевих архітектур і підкреслюють, що саме у випадку коротких текстів роль контекстуальних подань і спеціалізованих моделей зростає особливо суттєво [21, 22].

Традиційні лінгвістичні та статистичні підходи базуються на представленні тексту у вигляді вектора ознак, отриманих із поверхневої й морфологічної структури повідомлення [23]. Найпоширенішими є моделі «мішка слів» і «мішка n-грам», у яких текст описується частотами окремих слів або послідовностей слів/символів, часто зі зважуванням за схемою TF-IDF [24]. Типовий конвеєр включає токенізацію, нормалізацію, лематизацію, вилучення стоп-слів, іноді, розмітку частин мови, після чого отримані вектори подаються на вхід логістичній регресії, наївному байєсівському класифікатору, методу опорних векторів, випадковим лісам тощо [25]. Такі моделі залишаються конкурентними на невеликих і середніх корпусах, однак їхня якість істотно залежить від ручної інженерії ознак і середовища, в якому формувались повідомлення, що є критичним для коротких і шумних текстів соціальних мереж.

Окремий напрямок становлять словникові та лексиконні методи, де текст класифікується за його профілем у заздалегідь визначених словниках: тематичних, емоційних, стилістичних. Для кожного повідомлення підраховується кількість лексем, які належать до певних категорій, і на цій основі робиться висновок про тональність, належність до теми чи стилю [26]. Такі підходи широко використовуються в аналізі тональності й соціальних дослідженнях, оскільки забезпечують інтерпретованість критеріїв. Проте в контексті коротких текстів, що швидко змінюються й поповнюються новими мемами, жаргоном і неформальними формами, словникові ресурси не встигають актуалізуватися, через що частина змістовно важливих одиниць залишається поза полем зору моделі.

Компроміс між простотою класичних методів і необхідністю враховувати семантичні зв'язки забезпечують підходи, що використовують розподільні векторні подання слів. У таких моделях (word2vec, GloVe, інші word embeddings) кожне слово відображається у векторному просторі так, що лексеми з подібним контекстом мають близькі координати [27]. Подання короткого тексту формується як середнє або зважене поєднання векторів слів, після чого застосовуються лінійні класифікатори чи невеликі нейромережі. Характерним прикладом є fastText, де додатково враховуються символні n-грами, що дає змогу краще працювати з рідкісними й морфологічно складними словами та забезпечує високу швидкодію при збереженні якості на рівні глибинних моделей [28]. Саме швидкість та здатність працювати з дуже великими корпусами роблять fastText важливим базовим орієнтиром для задач класифікації коротких текстів.

Суттєвий етап розвитку методів пов'язаний із впровадженням згорткових і рекурентних нейронних мереж для аналізу тексту [29]. У класичній роботі Й. Кіма запропоновано використати одношарову згорткову нейронну мережу поверх попередньо навчених векторів слів для класифікації речень, де згортки з фільтрами різної ширини дозволяють виявляти локальні шаблони стійкі n-грами, характерні для певних класів [30]. Результати показали, що навіть відносно проста згорткова архітектура перевершує більшість традиційних методів на широкому наборі задач класифікації коротких текстів. Подальші роботи розвивали цю ідею, комбінуючи

згортки з кодуванням символів, частин мови та іншими лінгвістичними ознаками, що особливо корисно у випадку коротких, насичених інформативними маркерами повідомлень [22].

Рекурентні нейронні мережі, зокрема LSTM та GRU, були запропоновані як засоби моделювання послідовного характеру тексту, здатні враховувати порядок слів і довгострокові залежності. Двонапрямні архітектури дозволяють одночасно брати до уваги попередній і наступний контекст, що важливо для відтінків значення у коротких фразах. Введення механізмів уваги дає змогу виділяти найбільш важливі фрагменти тексту для класифікації й частково інтерпретувати рішення моделі. Порівняльні огляди свідчать, що такі моделі значно покращують якість класифікації коротких текстів порівняно з класичними підходами, особливо в задачах аналізу тональності, виявлення спаму, розпізнавання наміру користувача [21, 22].

Найбільш глибокі зміни в підходах до класифікації текстів пов'язані з появою трансформерних моделей, насамперед BERT (Bidirectional Encoder Representations from Transformers). BERT реалізує глибинні двонапрямні подання мовних одиниць, що попередньо навчаються на великих неанотованих корпусах за допомогою задач маскуванню токенів і прогнозування сусідніх речень [31]. Після попереднього навчання модель може бути донавчена для конкретної задачі класифікації, додаючи лише простий вихідний шар, при цьому демонструючи state-of-the-art результати на широкому спектрі задач, включно з короткими текстами. Важливо, що BERT природно підтримує формат пари висловлювань: на вхід подається два сегменти тексту, розділені спеціальним токеном, а модель навчається розпізнавати відношення між ними. У контексті задачі оцінювання релевантності хештегу це дозволяє розглядати пару «короткий текст-хештег» як єдиний вхід і безпосередньо навчати модель розрізняти узгоджені та неузгоджені пари.

Незважаючи на домінування трансформерних моделей, у спеціалізованих задачах класифікації коротких текстів досліджуються також гібридні архітектури, що поєднують розподільні подання, згорткові й рекурентні блоки, а також різні механізми уваги. Наприклад, для коротких заголовків і повідомлень пропонуються двоканальні архітектури, які одночасно обробляють послідовності слів та символічні

подання, щоб краще справлятися з орфографічними варіаціями та сленгом [32]. У роботах, присвячених активному навчанню, додатково розглядається питання вибору найбільш інформативних коротких текстів для розмітки, що особливо важливо за обмежених ресурсів і високої вартості ручної анотації [33].

Таким чином, сучасний ландшафт підходів до класифікації коротких текстів поєднує класичні лінгвістичні й статистичні методи з глибинними нейромережевими архітектурами різної складності. Для задачі оцінювання релевантності хештегів особливо перспективними є саме ті моделі, які можуть спільно обробляти пару «текст-тег» і враховувати їх контекстні подання, насамперед трансформери типу BERT, доповнені, за потреби, більш легкими згортковими або fastText-подібними моделями як базові орієнтири для порівняння. Це створює методологічне підґрунтя для побудови нейромережових рішень, здатних працювати з короткими повідомленнями соціальних мереж, у тому числі в задачі пояснюваного оцінювання релевантності хештегів.

### **1.3 Аналіз наукових праць для оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах**

У сучасних дослідженнях проблема оцінювання релевантності хештегів коротким текстовим повідомленням зазвичай розглядається через призму задачі автоматичної рекомендації хештегів, де людські мітки виступають «золотою еталонною» розміткою. У низці робіт релевантність ототожнюється з фактом входження рекомендованого хештегу до множини хештегів, фактично використаних автором допису, а якість моделей оцінюється за класичними метриками точності, повноти,  $F_1$ -міри або hit-rate. Огляд для Twitter та Sina Weibo, опублікований у журналі Future Internet, систематизував ці підходи, виділивши контент-орієнтовані, контекстуальні та графові методи, але при цьому релевантність не розглядається як окрема семантична категорія, а лише як збіг з існуючою розміткою [34].

Текстоцентричні нейромережеві моделі останніх років формулюють задачу як багатоміткову класифікацію коротких повідомлень. У праці Kirti Jain та Rajni Jindal

запропоновано гібридну модель на основі трансформерних векторних подань (BERT) та рекурентної мережі LSTM для автоматичного добору хештегів до твітів про COVID-19 у мовах з обмеженими ресурсами, де під релевантністю фактично розуміється збіг вихідного набору рекомендованих тегів з набором, обраним користувачем [35]. Подібний підхід спостерігається і в роботі P. Chakrabarti та співавторів, де запропоновано метод TRANSIT, орієнтований на підвищення популярності постів у соціальних мережах за рахунок урахування ключових слів допису, популярності користувача та поточних трендів, однак якість рекомендацій також вимірюється за відповідністю авторським хештегам без явного поділу на релевантні та нерелевантні з погляду змісту [15].

Вагомий блок робіт стосується мультимодальних моделей, що поєднують текстову та візуальну інформацію, і тим самим безпосередньо торкаються співвіднесення хештегів зі змістом повідомлення. У статті D. Roġar запропоновано гібридну модель аналізу зображень для рекомендації хештегів, яка використовує глибинні згорткові мережі для витягання візуальних ознак та їх поєднання з текстовими характеристиками, орієнтуючись на узгодженість тегів із зображенням у сервісах на кшталт Instagram [36]. У праці M. M. Y. Khalil та співавторів запропоновано трансформерний підхід до кросмодального подання даних для задачі прогнозування хештегів: текст і зображення кодуються спільною моделлю з механізмами уваги, що дозволяє краще узгоджувати тег з обома модальностями, однак і тут релевантність вимірюється через правильність передбачення набору тегів, а не через окремий критерій змістової відповідності [37].

Окрема лінія досліджень фокусується на виявленні саме нерелевантних або «інконгруентних» хештегів. У статті S. Dadgar та M. Neshat розглядається гібридна мультимодальна глибинна модель для визначення інконгруентності хештегів у соціальних мережах: текстовий компонент на основі трансформерної моделі й візуальний компонент на основі згорткової мережі використовуються для виявлення випадків, коли хештег суттєво не відповідає змісту зображення та супровідного підпису [38]. Хоча така постановка задачі наближається до вимірювання релевантності, модель оцінює відповідність на рівні глобального рішення

«конгруентний/інконгруентний», без детального аналізу того, на які саме фрагменти тексту чи які семантичні підмножини хештегу вона спирається.

Роботи, пов'язані з аналізом якості хештегів у візуальних середовищах, демонструють, що авторські мітки далеко не завжди відображають зміст, і тим самим ставлять під сумнів використання їх як безумовного «золотого стандарту» релевантності. Зокрема, дослідження S. Giannoulakis та співавторів для Instagram на основі гістограм кольорів зображень та наборів хештегів показало, що лише незначна частка тегів дійсно корелює з візуальними характеристиками контенту, тоді як решта виконує переважно функції самопросування, гумору або участі в трендах [39]. Це підкреслює, що проста відповідність між рекомендованими хештегами й тими, які фактично використовуються авторами, не завжди може розглядатися як коректний показник релевантності до змісту повідомлення.

У напрямі багатомовної та персоналізованої рекомендації хештегів сучасні дослідження пропонують графові та мережеві архітектури, де тег розглядається як вершина в графі взаємодії користувачів, постів і тематичних кластерів. У роботі Shubhi Bansal та колег представлено модель, яка реалізує персоналізовану рекомендацію хештегів для низькоресурсних індійських мов, використовуючи графову глибинну неймережу для поєднання мовних, семантичних та соціальних характеристик, однак і в цьому випадку релевантність інтерпретується через збіг з авторськими тегами та покращення показників рекомендаційної системи [40]. Такі підходи важливі для врахування контексту користувача та багатомовності, проте вони також працюють із релевантністю як із непрямую величиною, похідною від історичних даних.

Проблема саме оцінювання релевантності вийшла на передній план у роботах, присвячених удосконаленню методів валідації моделей. У статті A. Alsini, D. Huynh та A. Datta запропоновано семантичну систему оцінювання #REval, яка критикує традиційну практику жорсткого зіставлення рекомендованих і «істинних» хештегів та демонструє, що така схема ігнорує семантичну близькість між синонімічними або тематично спорідненими мітками [41]. У межах #REval вводиться метрика, що враховує семантичну відстань між тегами на основі їх векторних подань, а також

аналізується чутливість оцінки до обраного способу побудови ембедингів. Хоча ця робота робить важливий крок від поверхневих метрик до семантично обґрунтованих показників, вона все ще розглядає релевантність у площині подібності між тегами, а не у зв'язку між тегом і конкретними фрагментами тексту чи зображення.

Таким чином, аналіз сучасних праць показує, що більшість нейромережових і мовних підходів до роботи з хештегами опосередковано трактують релевантність як здатність моделі відтворити авторські набори тегів. Текстові та мультимодальні моделі навчаються на історичних даних, де теги вважаються апріорно правильними, а якість формально вимірюється за точністю й повнотою відносно цієї розмітки. Окремі дослідження, що фокусуються на виявленні інконгруентних хештегів або аналізі реальної відповідності тегів змісту зображень, демонструють, що така постановка не завжди коректно відображає змістову релевантність. Семантичні рамки оцінювання на кшталт #REval частково виправляють ситуацію, але зосереджуються здебільшого на подібності самих тегів. У цьому контексті відкритим залишається напрям, пов'язаний із побудовою метрик, які поєднують ймовірнісну оцінку класифікатора «релевантний/нерелевантний» з аналізом того, наскільки обґрунтування моделі спирається саме на ключові фрагменти тексту, що семантично відповідають ядру хештегу. Такий підхід забезпечує більш змістовне трактування релевантності й узгоджується з сучасними вимогами до пояснюваності інтелектуальних систем у сфері цифрових комунікацій.

#### **1.4 Постановка задачі**

Метою кваліфікаційної роботи магістра є підвищення достовірності та інтерпретованості оцінювання релевантності хештегів коротким текстовим повідомленням шляхом використання пояснюваного нейромережового підходу, який поєднує трансформерну класифікацію з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

Для досягнення мети слід вирішити такі задачі:

– виконати аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах;

– здійснити аналіз сучасних нейромережових та лінгвістичних підходів до класифікації коротких текстів, рекомендації тегів і пояснюваного аналізу рішень моделей обробки природної мови;

– проаналізувати наявні набори даних соціальних мереж, обґрунтувати вибір датасету для формування пар «текст – хештег – мітка» та визначити показники якості оцінювання релевантності;

– розробити нейромережовий метод пояснюваного оцінювання релевантності хештегів, що поєднує трансформерну класифікацію з метриками узгодженості важливих фрагментів тексту зі змістовим ядром хештегу;

– виконати підготовку та анотування вибірки коротких текстів і хештегів на основі даних соціальних мереж з формуванням позитивних і негативних прикладів для донавчання моделі;

– здійснити програмну реалізацію запропонованого методу з використанням сучасних засобів глибинного навчання та пояснюваного штучного інтелекту;

– провести експериментальні дослідження ефективності запропонованого методу порівняно з базовими підходами, що ґрунтуються лише на класифікаційних оцінках або векторних поданнях;

– виконати аналіз отриманих результатів, оцінити вплив пояснюваних метрик на якість відсікання нерелевантних хештегів і виявлення семантично коректних, але рідкісних тегів, та сформулювати практичні рекомендації щодо застосування методу.

## **РОЗДІЛ 2 Метод оцінювання відповідності хештегів коротким текстам засобами NLP**

### **2.1 Ідея підходу до оцінювання відповідності хештегів коротким текстам засобами NLP**

Ідея підходу полягає в тому, щоб трактувати релевантність хештегу короткому тексту не лише як відповідь нейромережевого класифікатора, а як поєднання цієї відповіді з формалізованим аналізом того, чому модель вважає тег доречним. На першому етапі з реального датасету формується вибірка пар «текст-хештег» з позитивними й негативними прикладами, на основі якої донавчається попередньо тренована мовна модель типу BERT [42]. Вона виступає оцінювачем відповідності і для кожної пари повертає ймовірність того, що хештег відповідає змісту тексту. Таким чином фіксується латентне семантичне відношення між текстом і тегом, вбудоване у параметри трансформерної моделі.

Другий ключовий елемент полягає у пояснюваності цієї оцінки. Для кожної пари «текст-хештег» аналізується внесок окремих слів підпису в рішення моделі (атрибуції важливості), а сам хештег розкладається на семантичне ядро, базові змістові компоненти. Далі обчислюється, наскільки слова, які модель вважає найважливішими для класу «релевантний», лексично та словотвірно пов'язані з цим ядром. Отримана міра узгодженості показує, чи справді нейромережа «дивиться» на правильні з погляду людини фрагменти тексту, чи її рішення зумовлені випадковими або побічними ознаками. Підсумкова оцінка релевантності формується як інтегральний показник, що поєднує ймовірність класифікатора та узгодженість пояснення з ядром хештегу з урахуванням параметра впливу нейромережевої складової. Таким чином метод одночасно враховує прогностичну здатність моделі та прозорість її рішень і дозволяє не лише ранжувати хештеги за релевантністю, а й надавати інтерпретовані пояснення для кожного з них.

## 2.2 Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP

На схемі (рисунок 2.1) подано узагальнений підхід пояснюваного оцінювання релевантності хештегів коротким текстам, що складається з блоку вхідних даних, трьох послідовних кроків обробки та блоку вихідних результатів.

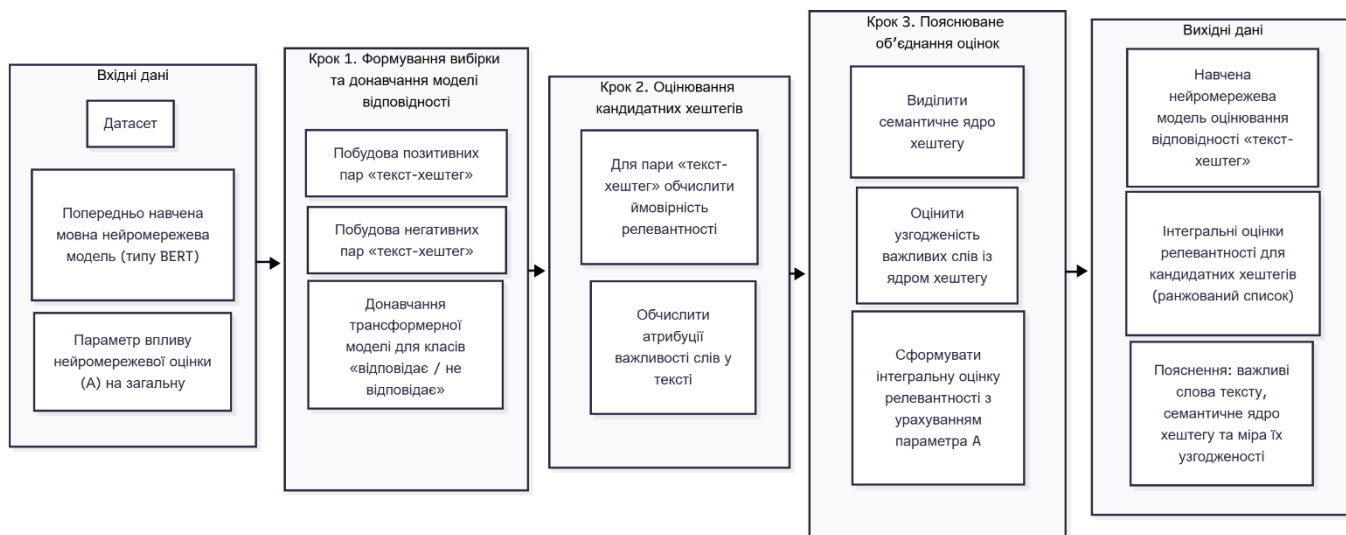


Рисунок 2.1 – Схема та етапи підходу до оцінювання відповідності хештегів коротким текстам засобами NLP

У блоці «Вхідні дані» подано три основні ресурси, на яких побудовано підхід. По-перше, це датасет коротких текстів із наявними хештегами, з якого формуються навчальні та тестові вибірки. По-друге, попередньо навчена мовна нейромережева модель типу BERT, що використовується як базовий трансформер для подання тексту і хештегу. По-третє, параметр впливу нейромережевої оцінки на загальну (позначений як  $A$ ), який визначає, наскільки сильно вихід класифікатора буде впливати на інтегральну оцінку порівняно з пояснюваною частиною.

Крок 1 «Формування вибірки та донавчання моделі» відповідає за побудову класифікатора відповідності. На основі датасету формується множина позитивних пар «текст-хештег» (реальні теги, вказані автором) і негативних пар (штучно підібрані інші хештеги, які не використовувалися в даному дописі). Ці пари подаються на вхід

попередньо навченої трансформерної моделі, яка донавчається на бінарну задачу класифікації з класами «відповідає» / «не відповідає». Результатом цього кроку є нейромережева модель, здатна для довільної пари «текст-хештег» оцінювати ймовірність їхньої відповідності.

Крок 2 «Оцінювання кандидатних хештегів» використовує донавчену модель для аналізу конкретних кандидатів. Для кожної пари «текст-хештег» модель обчислює ймовірність релевантності (нейромережеву оцінку відповідності). Паралельно за допомогою методів пояснюваного ШІ [43, 44] (атрибуцій важливості слів у тексті) визначається, які саме токени підпису зробили найбільший внесок у рішення «хештег релевантний». На виході цього кроку для кожного кандидата маємо числову оцінку ймовірності та розподіл важливості слів у тексті.

Крок 3 «Пояснюване об'єднання оцінок» забезпечує інтерпретацію та фінальне ранжування. Спочатку для кожного хештегу виділяється його семантичне ядро: тег очищується від службових символів, розкладається на змістові компоненти (підслова, корені, частини складених слів), які розглядаються як базові смислові одиниці. Далі за атрибуціями з попереднього кроку оцінюється, наскільки важливі для моделі слова тексту лексично й словотвірно пов'язані з цим ядром; на цій основі розраховується показник узгодженості. Нарешті, нейромережева ймовірність релевантності та показник узгодженості об'єднуються в інтегральну оцінку релевантності, причому внесок першої регулюється параметром  $A$ . Отримана оцінка відображає як впевненість моделі, так і змістову обґрунтованість її рішення.

У блоці «Вихідні дані» зведено результати роботи всього підходу, а саме: навчена нейромережева модель оцінювання відповідності «текст-хештег», яку можна застосовувати до нових даних; інтегральні оцінки релевантності для множини кандидатних хештегів у вигляді ранжованого списку для кожного тексту; пояснювальна інформація: перелік найважливіших слів тексту, виділене семантичне ядро кожного хештегу та числова міра їх узгодженості, що дозволяє інтерпретувати, чому саме система вважає конкретний тег більш або менш доречним.

### 2.3 Спосіб оцінювання кандидатних хештегів відповідно до повідомлення

Нижче (рисунку 2.2) наведено детально опис способу оцінювання окремого кандидатного хештегу відносно конкретного короткого текстового повідомлення, який реалізує другий та третій кроки підходу (п.2.3) і забезпечує як числову, так і пояснювану оцінку релевантності.

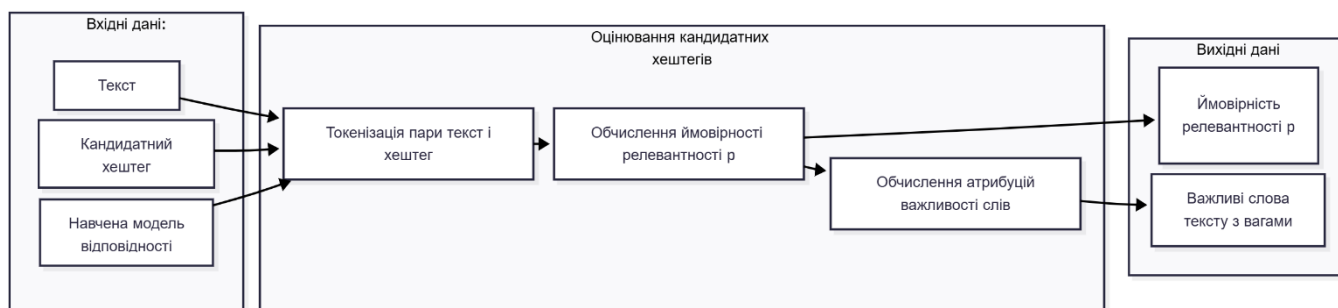


Рисунок 2.2 – Схема процесу оцінювання кандидатних хештегів

Вхідними даними є текст короткого допису, кандидатний хештег, релевантність якого необхідно оцінити, а також навчена нейромережева модель відповідності «текст-хештег». Текст і тег інтерпретуються як єдина логічна пара, що подається на обробку. Спершу відбувається реалізація блоку токенізації [45]: текст і хештег об'єднуються у формат вхідної послідовності трансформерної моделі, після чого перетворюються на послідовність ідентифікаторів токенів із відповідними масками уваги та, за потреби, індикаторами сегментів. На цьому кроці виконується узгодження вихідних даних із вимогами архітектури моделі, що забезпечує коректне подання як текстового контексту, так і тегу.

Далі токенізована пара передається до блоку обчислення ймовірності релевантності. Навчена модель генерує вектор логітів для класів «відповідає» та «не відповідає», який перетворюється на ймовірнісний розподіл. Значення, що відповідає класу «хештег відповідає тексту», інтерпретується як нейромережева оцінка релевантності  $p$ . Ця величина фіксує ступінь впевненості моделі у тому, що заданий хештег є змістово узгодженим із конкретним повідомленням, і надалі використовується як ключовий числовий показник якості кандидатного тегу.

Паралельно з обчисленням ймовірності активується блок атрибуцій важливості слів. Для тієї ж токенованої пари «текст-хештег» за допомогою методу інтегрованих градієнтів визначається внесок кожного токена тексту в збільшення ймовірності класу «відповідає». Отримані значення нормуються, службові токени відсікаються, а для лексичних одиниць тексту формується розподіл ваг важливості, який відображає, які саме слова відіграли найсуттєвішу роль у рішенні моделі.

Таким чином, вихідними даними підсистеми є, з одного боку, числова оцінка релевантності  $p$ , а з іншого – пояснення у вигляді набору слів тексту з відповідними вагами важливості. Ці результати далі використовуються на наступному етапі методу для побудови інтегральної оцінки відповідності кандидатних хештегів.

#### **2.4 Спосіб пояснюваного об'єднання оцінок та формування інтегральної міри релевантності**

Для пояснюваного об'єднання оцінок та формування інтегральної міри релевантності використовується інформація, отримана на кроці 2 підходу (п.2.2), а саме ймовірність релевантності пари «текст-хештег» та розподіл важливості слів у тексті, доповнена самою формою хештегу і параметром впливу нейромережевої оцінки. Відповідна структурна схема процесу наведена на рисунку 2.3.

Вхідними даними є кандидатний хештег, числове значення ймовірності релевантності  $p$ , обчислене донавченою моделлю, а також множина слів тексту з вагами важливості, отриманими за результатами атрибуцій. Додатковим керувальним параметром виступає коефіцієнт  $A$ , який задає вагу нейромережевої оцінки в інтегральній мірі релевантності. Ці величини у сукупності описують як «довіру» моделі до хештегу, так і те, на які саме фрагменти тексту вона спирається.

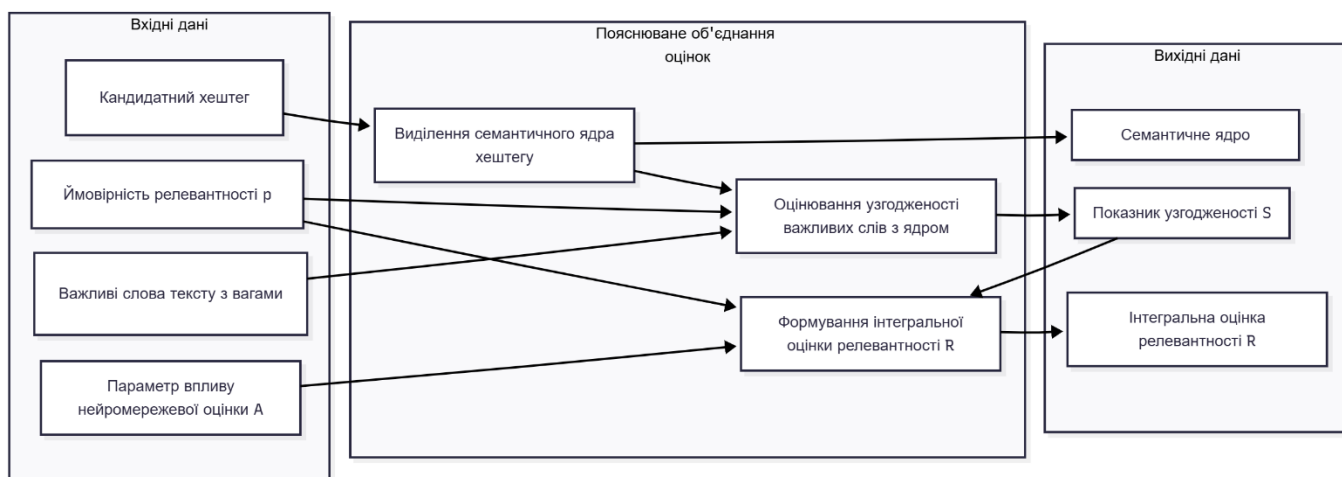


Рисунок 2.3 – Схема способу пояснюваного об'єднання оцінок та формування інтегральної міри релевантності

Спершу відбувається виділення семантичного ядра хештегу [46]. Тег очищується від службового символу та інших технічних маркерів, після чого розкладається на змістові компоненти, що можуть відповідати окремим словам, підсловом або частинам складених лексем. Отримані фрагменти нормуються, зводяться до базових форм і розглядаються як множина ключових смислових одиниць хештегу. Саме з ними у подальшому співвідносяться важливі слова тексту, що дозволяє інтерпретувати, який зміст повинен відображати релевантний тег.

Далі здійснюється оцінювання узгодженості між поясненням моделі та семантичним ядром хештегу. На цьому кроці використовується розподіл важливості слів: аналізуються ті лексеми тексту, які отримали найбільші ваги за результатами атрибуцій [47]. Для кожного такого слова перевіряється наявність лексичного або словотвірного зв'язку з компонентами ядра, зокрема повні збіги, спільні корені та включення одного фрагмента в інший. На основі цього аналізу розраховується показник узгодженості  $S$ , який інтерпретується як відносна частка важливих слів, змістово пов'язаних із ядром хештегу. Високе значення  $S$  свідчить про те, що модель, ухвалюючи рішення про релевантність, опирається на семантично доречні фрагменти тексту, тоді як низьке значення вказує на можливу випадковість або побічний характер використаних ознак.

Заключним етапом є формування інтегральної оцінки релевантності  $R$ . Для цього поєднуються дві складові: нейромережева ймовірність  $p$  і показник узгодженості  $S$ . Ваговий параметр  $A$  визначає, наскільки сильно на підсумкову оцінку впливає безпосередньо вихід класифікатора порівняно з внеском пояснюваної частини. За високих значень  $A$  більше значення має ймовірність  $p$ , тоді як за менших значень  $A$  підсилюється роль показника  $S$ . У результаті формується інтегральна міра  $R$ , яка одночасно відображає ступінь впевненості моделі в релевантності хештегу та ступінь змістової обґрунтованості цього рішення з погляду узгодженості важливих слів із семантичним ядром теґу.

Вихідними даними кроку є семантичне ядро хештегу, числовий показник узгодженості  $S$  та інтегральна оцінка релевантності  $R$  для кожної пари «текст-кандидатний хештег». Ці величини надалі використовуються для формування ранжованих списків хештегів, відсікання очевидно нерелевантних теґів і побудови інтерпретованих пояснень щодо причин присвоєння тій чи іншій мітці високої або низької оцінки.

## 2.5 Формування та підготовка навчальних даних

Розглядуваний у роботі датасет Instagram Reach Data [48] репрезентує емпіричні дані про ефективність дописів в Instagram для одного окремого облікового запису автора amankharwal.official. Набір даних опублікований на платформі Kaggle [49] як відкрите джерело для аналізу охоплення дописів та побудови моделей прогнозування їхнього «reach» і взаємодій користувачів. Кожен рядок відповідає одному посту, а змінні описують як кількість показів і взаємодій, так і супровідний текстовий контент у вигляді підпису та використаних хештегів. Сукупний обсяг даних становить близько 120 спостережень, що дає змогу розглядати цей набір як компактний, але цілісний зріз контентної активності конкретного освітнього/аналітичного блогу.

Структура датасету поєднує кілька груп ознак. Базовим інтегральним показником є змінна «Impressions», яка інтерпретується як кількість переглядів допису й використовується як наближення до досяжності або охоплення посту. До неї прив'язані чотири змінні-розкладки, що фіксують джерела формування цього охоплення: перегляди з домашньої стрічки («From Home»), з хештегів («From Hashtags»), з вкладки «Explore» («From Explore») та з інших каналів («From Other»). Така декомпозиція дає змогу відокремлювати органічну взаємодію постійної аудиторії від додаткового трафіку, що надходить через механізми рекомендацій та пошуку за тегами. Приклад даних датасету наведено на рисунку 2.4.

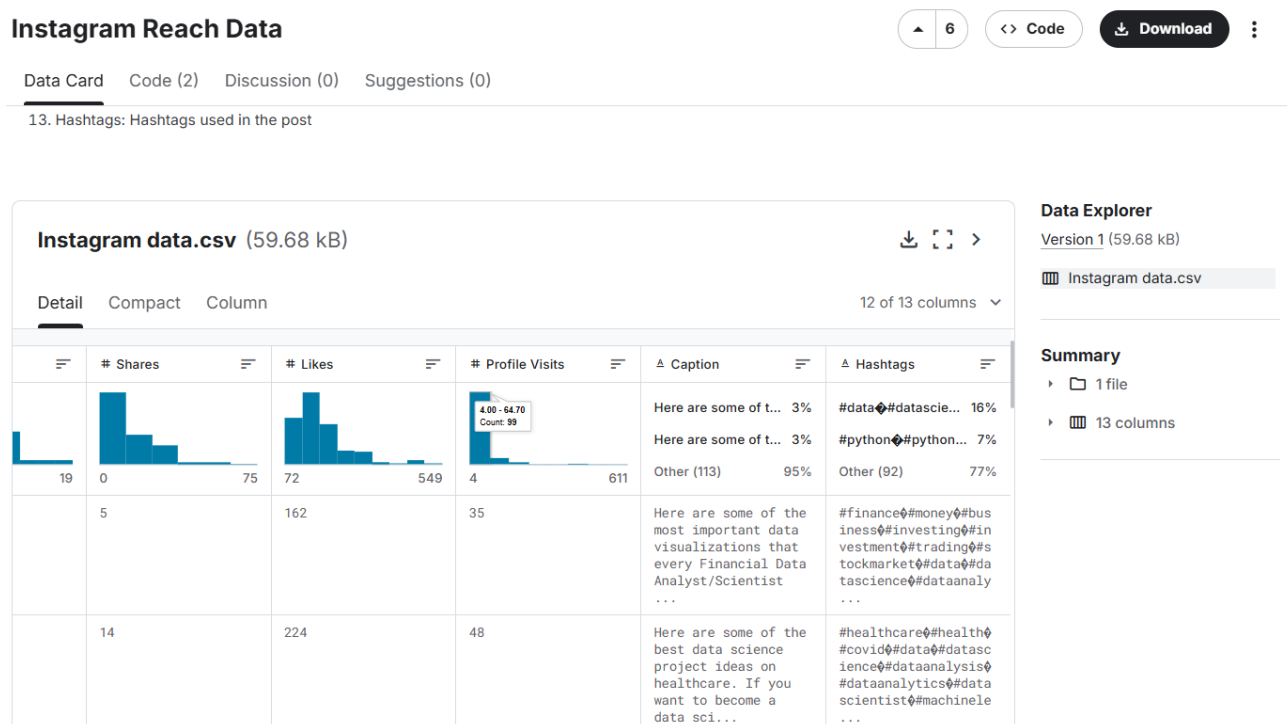


Рисунок 2.4 – Приклад записів датасету

Окрема група змінних описує безпосереднє залучення аудиторії. До неї належать кількість збережень допису («Saves»), коментарів («Comments»), поширень («Shares») та вподобань («Likes»). Ці показники відображають різні форми взаємодії користувачів із контентом і слугують як для класичного аналізу кореляцій із охопленням, так і для побудови регресійних або нейромережових моделей,

орієнтованих на прогнозування імпресій за відомими значеннями лайків, збережень чи поширень.

Ще одна важлива підсистема ознак пов'язана з переходами на профіль і конверсією в нових підписників. Змінна «Profile Visits» фіксує кількість відвідувань профілю, згенерованих конкретним дописом, тоді як «Follows» відображає число нових підписників, яких вдалося залучити через цей пост. Ці два показники часто аналізуються спільно, зокрема для обчислення коефіцієнта конверсії «Follows/Profile Visits» та дослідження того, наскільки певний тип контенту здатний не лише генерувати перегляди, а й стимулювати підписку на акаунт.

На відміну від багатьох «числових» маркетингових датасетів, Instagram Reach Data включає й текстову складову. Для кожного посту наявний стовпець «Caption», що містить повний текст підпису, та стовпець «Utilized Hashtags», у якому збережено список хештегів, використаних автором. Наявність цих полів робить набір особливо цінним для задач обробки природної мови, оскільки дозволяє поєднувати кількісні метрики охоплення та взаємодії з аналізом лінгвістичного змісту постів і маркувальної ролі хештегів. Саме зв'язок між текстом підпису, вибором тегів і подальшим охопленням лежить в основі багатьох досліджень, що використовують цей датасет для побудови предиктивних і пояснюваних моделей.

З методологічного погляду важливо, що дані репрезентують один конкретний акаунт із відносно стабільною тематикою дописів, орієнтованих на машинне навчання та аналіз даних. Це зумовлює певні обмеження щодо генералізованості висновків на інші домени або стилі ведення сторінок, однак водночас зменшує міжавторську та міжтемпоральну гетерогенність, дозволяючи зосередитися на внутрішніх закономірностях взаємодії між змістом посту, використаними хештегами та показниками охоплення.

## **Висновки до розділу 2**

У підсумку другого розділу обґрунтовано цілісну концепцію оцінювання відповідності хештегів коротким текстам, у якій поєднано потенціал сучасних мовних

нейромережових моделей з вимогами до інтерпретованості рішень. На концептуальному рівні релевантність хештегу розглянуто не як бінарну ознаку, а як результат двокомпонентного процесу: статистичної оцінки відповідності пари «текст-хештег» у спільному семантичному просторі та аналізу того, наскільки ця оцінка спирається на змістові фрагменти тексту, пов'язані із семантичним ядром тегу. Такий підхід дозволив перейти від «чорної скриньки» нейромережевої класифікації до пояснюваної інтегральної міри релевантності, що відбиває як упевненість моделі, так і змістову обґрунтованість її рішення.

Запропонований метод охоплює повний цикл обробки даних: від формування пар «текст-хештег» на основі реального датасету до побудови інтегральної оцінки відповідності для кожного кандидатного тегу. На першому етапі здійснюється донавчання мовної моделі на задачі класифікації «відповідає» / «не відповідає», що дає змогу одержувати ймовірнісну оцінку релевантності для довільної пари «допис-хештег». На другому етапі ця оцінка доповнюється поясненням у вигляді розподілу важливості слів тексту, отриманого за допомогою інтегрованих градієнтів. На третьому етапі відбувається співставлення важливих слів із семантичним ядром тегу та розрахунок показника узгодженості, після чого нейромережева ймовірність і ця пояснювана складова об'єднуються в єдину інтегральну міру релевантності. Внаслідок цього метод не лише ранжує хештеги, а й надає людині інтерпретовані аргументи щодо кожного рішення.

Окремим результатом розділу є формалізація двох ключових допоміжних процедур: способу оцінювання одного кандидатного хештегу відносно конкретного тексту та способу пояснюваного об'єднання оцінок. Перша процедура забезпечує узгоджене подання пари «текст-хештег» на вході донавченої моделі й одержання нейромережевої ймовірності релевантності разом із розподілом важливості слів. Друга процедура перетворює ці дані на пояснювану інтегральну оцінку через виділення семантичного ядра хештегу, розрахунок показника узгодженості між ядром і важливими словами тексту та кероване ваговим параметром поєднання цієї інформації з виходом класифікатора. Така декомпозиція робить метод гнучким і

модульним, що важливо для його подальшої модифікації та перенесення на інші домени.

Значущим практичним елементом є детальний аналіз і адаптація датасету Instagram Reach Data до задачі оцінювання релевантності хештегів. Показано, що поєднання кількісних метрик охоплення, показників залучення аудиторії та текстових полів «Caption» і «Utilized Hashtags» створює адекватну основу для навчання та валідації запропонованого методу. Водночас акцентовано обмеженість генералізації висновків через те, що дані стосуються одного акаунта зі специфічною тематикою. Це накладає природні межі на інтерпретацію результатів, але водночас зменшує гетерогенність і дозволяє сконцентруватися на внутрішніх закономірностях взаємодії між текстом, хештегами та охопленням.

Наукова новизна розділу полягає у введенні інтегрованої міри релевантності, яка поєднує ймовірнісний вихід мовної моделі та формалізовану міру узгодженості пояснення з семантичним ядром хештегу. На відміну від підходів, що спираються виключно на відстані в просторах ознак або на «плоскі» класифікаційні рішення, запропонований метод контролює не тільки результат, а й механізм його отримання: релевантним вважається тег, для якого модель демонструє високу впевненість і водночас опирається на лексеми, змістово пов'язані з ядром тегу. Такий підхід відповідає сучасним тенденціям у галузі пояснюваного штучного інтелекту й може бути розглянутий як внесок у розроблення прозорих методів рекомендації хештегів.

Нарешті, отримані в розділі результати задають основу для подальших експериментів і розвитку методу. На їх основі можна досліджувати чутливість інтегральної оцінки до вибору вагового параметра, тестувати альтернативні схеми виділення ядра хештегу, порівнювати різні методи атрибуцій важливості та переносити підхід на багатомовні чи мультиавторські датасети. У сукупності розроблений у розділі метод створює теоретично обґрунтовану та практично придатну основу для автоматизованого, пояснюваного контролю релевантності хештегів у цифрових комунікаційних середовищах.

## **РОЗДІЛ 3 Проєктування інтелектуальної системи оцінювання відповідності хештегів коротким текстам**

### **3.1 Вибір засобів розробки інтелектуальної системи**

У межах проєктування інтелектуальної системи оцінювання відповідності хештегів коротким текстам ключовим етапом є обґрунтований вибір засобів розробки, що мають забезпечити сумісність із сучасними методами обробки природної мови, підтримку глибинних нейромережових моделей, можливість інтерактивних експериментів і побудову зручного для користувача інтерфейсу. З огляду на дослідницький характер роботи, обмеженість обчислювальних ресурсів у локальному середовищі та необхідність швидкого прототипування, було обрано технологічний стек, орієнтований на мову Python [50], хмарне інтерактивне середовище Google Colab [51] та засоби побудови вебінтерфейсів, що безпосередньо інтегруються з таким середовищем.

Мова програмування Python є базовим засобом розробки як для ядра інтелектуальної системи, так і для експериментальних скриптів та прототипів. Вибір саме Python зумовлений його домінуючою роллю в галузі машинного навчання та обробки природної мови, наявністю розвиненої екосистеми фреймворків глибинного навчання, засобів роботи з текстом і засобів пояснюваного ШІ [52]. Важливим є також те, що Python забезпечує низький поріг входу при достатній виразності для реалізації складних архітектур, зокрема трансформерних моделей типу BERT, і дає змогу поєднувати в одному середовищі етапи підготовки даних, навчання моделей, аналізу результатів та інтеграції з вебінтерфейсами.

Як середовище розробки та виконання обчислювальних експериментів використано Google Colab, що поєднує модель роботи інтерактивного блокнота Jupyter з доступом до хмарних графічних процесорів [53]. Такий вибір дозволяє реалізувати повний цикл розробки інтелектуальної системи без розгортання локальної інфраструктури, забезпечити відтворюваність експериментів за рахунок збереження ноутбуків і налаштувань середовища, а також зручно працювати з

зовнішніми даними на кшталт датасету Instagram Reach Data. Інтерактивний характер середовища дає змогу поєднувати в одному документі код, текстові пояснення, проміжні візуалізації метрик навчання та результати оцінювання релевантності хештегів, що є критично важливим для магістерського дослідження з акцентом на аналізі та інтерпретації.

Для реалізації власне нейромережевої частини системи обрано фреймворк глибинного навчання PyTorch [54] у поєднанні з екосистемою попередньо навчених мовних моделей. Такий вибір відповідає потребі в гнучкому керуванні навчальними циклами, можливості легко модифікувати архітектуру «голови» над трансформерним енкодером, реалізовувати власні функції втрат та процедури обчислення атрибутів. PyTorch добре інтегрується з апаратним прискоренням у середовищі Colab, що забезпечує прийнятний час донавчання моделі класифікації пар «текст-хештег» навіть за використання BERT-рівневих архітектур. Наявність зрілої підтримки трансформерних моделей у відповідних фреймворках дає змогу зосередитися на постановці задачі, формуванні пар позитивних і негативних прикладів та реалізації пояснюваної складової, не витрачаючи ресурси на низькорівневу реалізацію базової архітектури.

З огляду на те, що однією з цілей роботи є створення не лише дослідницького прототипу, а й інструменту, придатного для інтерактивного використання, додатково обрано засоби побудови веборієнтованого інтерфейсу, які можуть працювати безпосередньо всередині Google Colab. Для цього доцільно застосовувати легковагові фреймворки на кшталт Gradio [55], що дозволяють описати інтерфейс як набір вхідних полів (для тексту повідомлення та множини кандидатних хештегів) і вихідних компонентів (таблиці з оцінками релевантності, візуалізації важливих слів) без необхідності розгортання окремого серверного застосунку [56]. Такий підхід відповідає дослідницькому формату магістерської роботи: основна логіка системи залишається в межах ноутбука, а вебінтерфейс використовується як зручний інструмент демонстрації та тестування запропонованого методу.

У сукупності обрані засоби розробки формують узгоджене технологічне середовище, орієнтоване на експериментальну побудову інтелектуальної системи з

елементами пояснюваного III. Мова Python, фреймворк глибинного навчання на основі PyTorch, інтерактивне хмарне середовище Google Colab та веборієнтований інтерфейс, реалізований засобами, що інтегруються з блокнотами, забезпечують необхідний баланс між обчислювальними можливостями, гнучкістю налаштування, інтерпретованістю отриманих результатів та доступністю рішень для подальшого використання й демонстрації.

### **3.2 Проектування структури інтелектуальної системи**

Здійснено проектування внутрішньої структури інтелектуальної системи оцінювання відповідності хештегів коротким текстам. Логіка системи організована у вигляді багаторівневої клієнт-серверної архітектури, де вебінтерфейс Gradio забезпечує взаємодію користувача із серверною частиною, а серверна логіка, реалізована мовою Python з використанням PyTorch, виконує всі обчислювальні та аналітичні операції. На рисунку 3.1 подано загальну структурну схему системи, яка відображає основні підсистеми, сервіси й функціональні модулі, а також їх зв'язки із зовнішніми ресурсами.

Верхній рівень становить вебінтерфейс Gradio, що представлений набором тематичних вкладок. Вкладка «Огляд даних» призначена для первинного знайомства з датасетом, перегляду фрагментів таблиці, розподілів ознак і прикладів підписів та хештегів. Вкладка «Навчання моделі» надає доступ до параметрів донавчання мовної моделі, запуску навчального процесу та моніторингу його перебігу за допомогою метрик і графіків. Вкладка «Моделі» зосереджена на роботі з уже навченими екземплярами класифікатора, дозволяючи обирати, завантажувати й зберігати варіанти моделі. Вкладка «Оцінка хештегів» реалізує основний користувацький сценарій: введення короткого тексту, переліку кандидатних хештегів і отримання для них інтегральних оцінок релевантності разом із поясненнями.

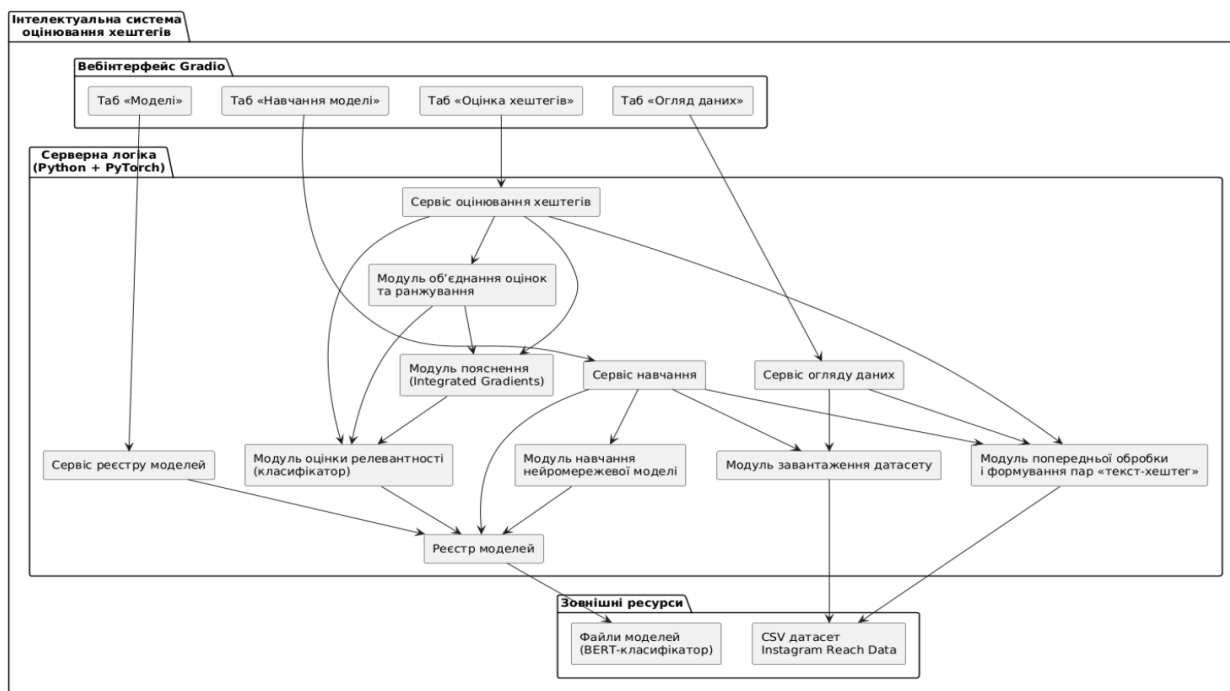


Рисунок 3.1 – Схема взаємодії складових інтелектуальної системи

Кожна з вкладок інтерфейсу не звертається безпосередньо до низькорівневих модулів, а взаємодіє з відповідним сервісом серверної логіки. Так, «Огляд даних» підключений до сервісу огляду даних, «Навчання моделі» – до сервісу навчання, «Моделі» – до сервісу реєстру моделей, а «Оцінка хештегів» – до сервісу оцінювання хештегів. Така організація забезпечує чітке розділення відповідальностей: інтерфейс відповідає лише за введення та візуалізацію, тоді як кожен сервіс інкапсулює сценарій використання і координує роботу внутрішніх модулів.

Серверна логіка поділена на чотири основні сервіси, що відображено на рисунку. Сервіс огляду даних відповідає за завантаження CSV-датасету Instagram Reach Data, його попереднє перетворення та підготовку фрагментів для відображення у вкладці «Огляд даних». Для цього він взаємодіє з модулем завантаження датасету, який здійснює читання даних з файлу, та з модулем попередньої обробки й формування пар «текст–хештег», який очищує текст, розбиває рядки з хештегами на окремі теги й формує структури даних, необхідні для подальшого навчання і оцінювання. Сервіс навчання використовує ті самі модулі роботи з даними, але доповнює їх модулем навчання нейромережевої моделі, що реалізує цикл донавчання

BERT-класифікатора на сформованих парах «текст–хештег» і, після завершення, зберігає параметри моделі через реєстр моделей.

Сервіс реєстру моделей слугує єдиною точкою доступу до збережених екземплярів класифікатора. Він взаємодіє з модулем «Реєстр моделей», який організовує структуру зберігання ваг моделі та пов'язаних з ними службових даних у файлової системі, а також забезпечує завантаження вибраного варіанта моделі до пам'яті. Вкладка «Моделі» через цей сервіс може відображати доступні екземпляри, ініціювати їх збереження після навчання або завантаження для подальшого використання на етапі оцінювання хештегів.

Центральне місце в архітектурі посідає сервіс оцінювання хештегів, який реалізує запропонований у другому розділі метод. Він взаємодіє з трьома спеціалізованими модулями: модулем оцінки релевантності (класифікатором), модулем пояснення та модулем об'єднання оцінок і ранжування. Модуль оцінки релевантності отримує від сервісу текст повідомлення та кандидатний хештег, завантажує відповідний екземпляр моделі з реєстру й повертає ймовірність класу «хештег відповідає тексту». Модуль пояснення, використовуючи метод інтегрованих градієнтів, для тієї ж пари «текст–хештег» обчислює атрибуції важливості слів у тексті та формує набір найважливіших лексем із відповідними вагами. Модуль об'єднання оцінок та ранжування на основі ймовірності класифікатора, пояснюваних ваг важливості та форми самого хештегу обчислює показник узгодженості й інтегральну оцінку релевантності, після чого формує впорядкований список кандидатних тегів для відображення у вебінтерфейсі.

На нижньому рівні розташовано зовнішні ресурси, з якими взаємодіють модулі серверної логіки. CSV-датасет Instagram Reach Data використовується модулем завантаження датасету та модулем попередньої обробки для формування навчальних і тестових вибірок. Файли моделей BERT-класифікатора обслуговуються модулем «Реєстр моделей», який зберігає результати навчання та забезпечує доступ до них для модуля оцінки релевантності. Таким чином, робота з файловою системою ізольована від рівня інтерфейсу й навіть від окремих сервісів, що підвищує мобільність і відтворюваність системи в різних середовищах виконання.

Запроектowana структура демонструє поєднання шаруватої й сервісно-орієнтованої архітектури. Шар вебінтерфейсу відповідає за взаємодію з користувачем та візуалізацію результатів; шар серверної логіки інкапсулює прикладну функціональність у вигляді сервісів і модулів; шар зовнішніх ресурсів містить дані та моделі, з якими працює система. Такий поділ спрощує супровід та розширення системи: за потреби можна замінити, наприклад, модуль пояснення на інший метод інтерпретації або додати нові вкладки в інтерфейсі, не змінюючи базових механізмів роботи з датасетом і моделями. В результаті сформовано гнучку та модульну структуру, що відображає логіку запропонованого методу оцінювання релевантності хештегів і забезпечує практичну реалізацію інтелектуальної системи в середовищі Python/Gradio.

### **3.3 Проекування елементів інтелектуальної системи**

Проекування елементів інтелектуальної системи є ключовим етапом, на якому загальна ідея та архітектура перетворюються на чітко окреслені функціональні можливості, орієнтовані на користувача [57, 58]. Саме на рівні елементів визначається, які дії система повинна підтримувати, які сценарії роботи будуть для неї типовими, як поєднуватимуться аналітичні модулі й інтерфейсні компоненти. Коректно спроековані елементи забезпечують логічну завершеність функцій, узгодженість з обраною архітектурою, а також зрозумілість і передбачуваність поведінки системи для кінцевого користувача. Це особливо важливо для інтелектуальних систем, де за простими на вигляд операціями стоять складні нейромережеві обчислення та процедури пояснюваного аналізу.

На рисунку 3.2 подано діаграму варіантів використання, яка відображає основні елементи функціональності інтелектуальної системи оцінювання хештегів і взаємодію користувача з ними. Єдиним актором виступає «Користувач системи», що узагальнює можливих аналітиків, дослідників або контент-менеджерів, які працюють із короткими текстами та хештегами. У середині прямокутника системи змодельовано п'ять ключових варіантів використання, які відповідають основним підсистемам

розробленого програмного комплексу. Варіант «Перегляд датасету та статистик» репрезентує можливість завантажити й візуально проаналізувати вихідний датасет Instagram Reach Data, ознайомитися з прикладами дописів, хештегів та базовими статистичними характеристиками. Варіант «Керування моделями (збереження/завантаження, вибір активної)» відображає роботу з реєстром нейромережових моделей: користувач може зберігати результати навчання, завантажувати наявні класифікатори та вибирати, яка модель буде використовуватися для подальшого оцінювання релевантності.

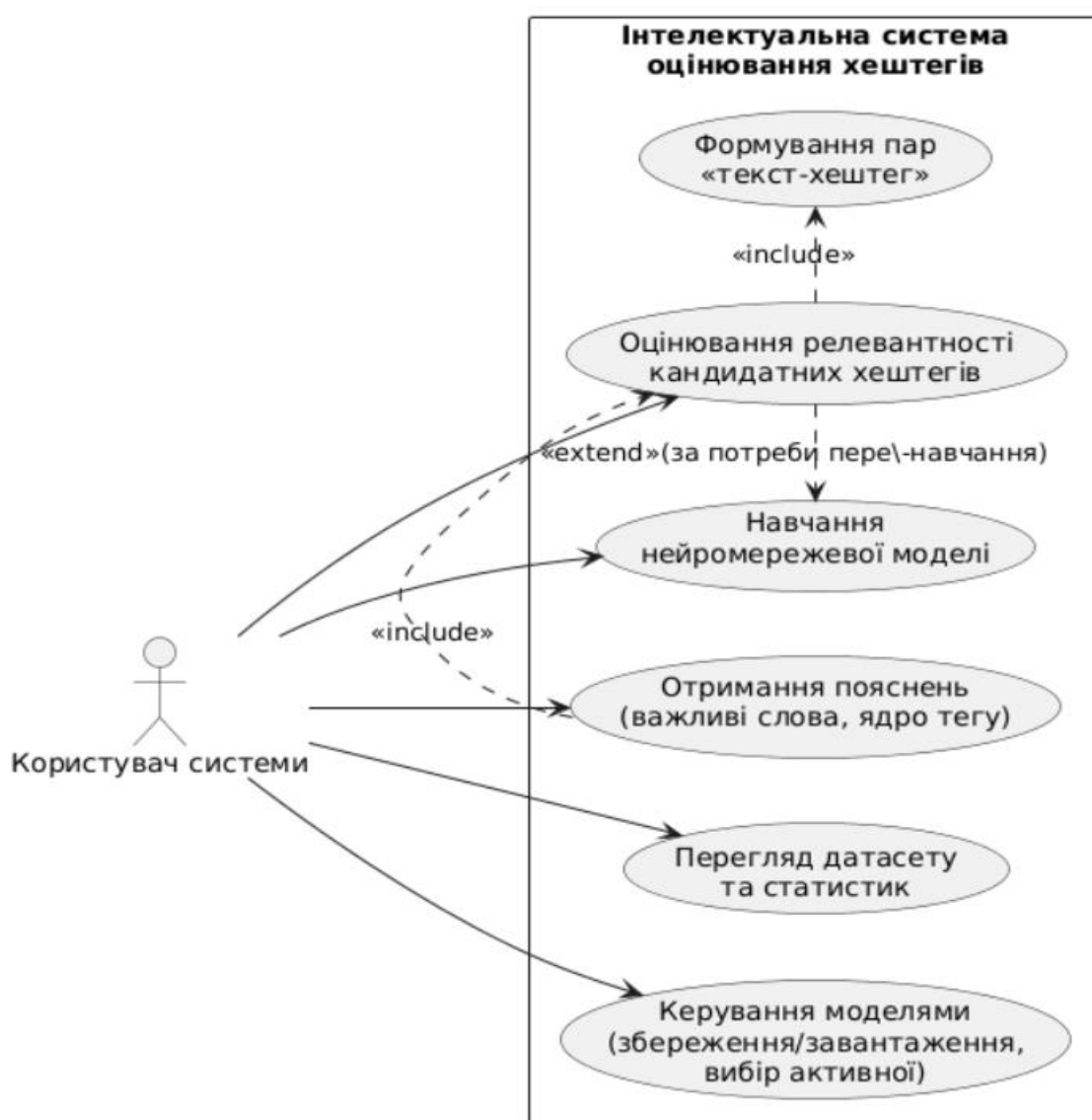


Рисунок 3.2 – Діаграма варіантів використання інтелектуальної системи

Центральне місце на діаграмі посідає варіант використання «Оцінювання релевантності кандидатних хештегів», який реалізує основну цінність системи. Саме

в межах цього сценарію користувач подає короткий текст допису і множину потенційних хештегів, а система повертає інтегральні оцінки релевантності для кожного з них. Даний варіант має зв'язок «include» з варіантом «Формування пар текст-хештег», що відображає обов'язкову для оцінювання службову операцію: перед тим як виконувати нейромеревий аналіз, система повинна перетворити вхідні дані на внутрішні пари «підпис-кандидатний тег» у форматі, придатному для подання трансформерній моделі. Крім того, варіант «Оцінювання релевантності кандидатних хештегів» пов'язаний відношенням «extend» з варіантом «Навчання нейромережевої моделі». Це означає, що в окремих сценаріях, коли поточна якість класифікації є недостатньою або з'явилися нові дані, користувач має можливість розширити базовий сценарій оцінювання запуском додаткового донавчання моделі без виходу з основного режиму роботи.

Варіант використання «Отримання пояснень (важливі слова, ядро тегу)» відображає пояснювальний компонент системи й логічно спирається на результати оцінювання релевантності. Його зв'язок «include» з варіантом «Оцінювання релевантності кандидатних хештегів» означає, що побудова пояснення можлива лише після того, як для відповідної пари «текст-хештег» уже обчислено ймовірність релевантності та проведено аналіз атрибутів важливості слів. У межах цього варіанта користувач отримує інтерпретовану інформацію: які слова тексту зробили найбільший внесок у рішення моделі, яке семантичне ядро має хештег і наскільки добре воно узгоджується з важливими фрагментами повідомлення. Таким чином діаграма варіантів використання наочно фіксує, що елементи системи не є ізольованими функціями, а утворюють узгоджений набір сценаріїв, де базове оцінювання хештегів, їх пояснення, навчання моделі, робота з датасетом і керування моделями логічно доповнюють одне одного та забезпечують повний цикл роботи з інтелектуальною системою.

На поданій діаграмі послідовності (рисунок 3.3) змодельовано динаміку роботи інтелектуальної системи в сценарії оцінювання релевантності хештегів. Вона демонструє, як користувач, вебінтерфейс Gradio та серверні сервіси взаємодіють між собою в часі під час аналізу однієї пари «текст-кандидатний хештег». Вертикальні

штрихові лінії відповідають «життєвим лініям» учасників, а горизонтальні стрілки відображають виклики методів і передавання даних між ними.

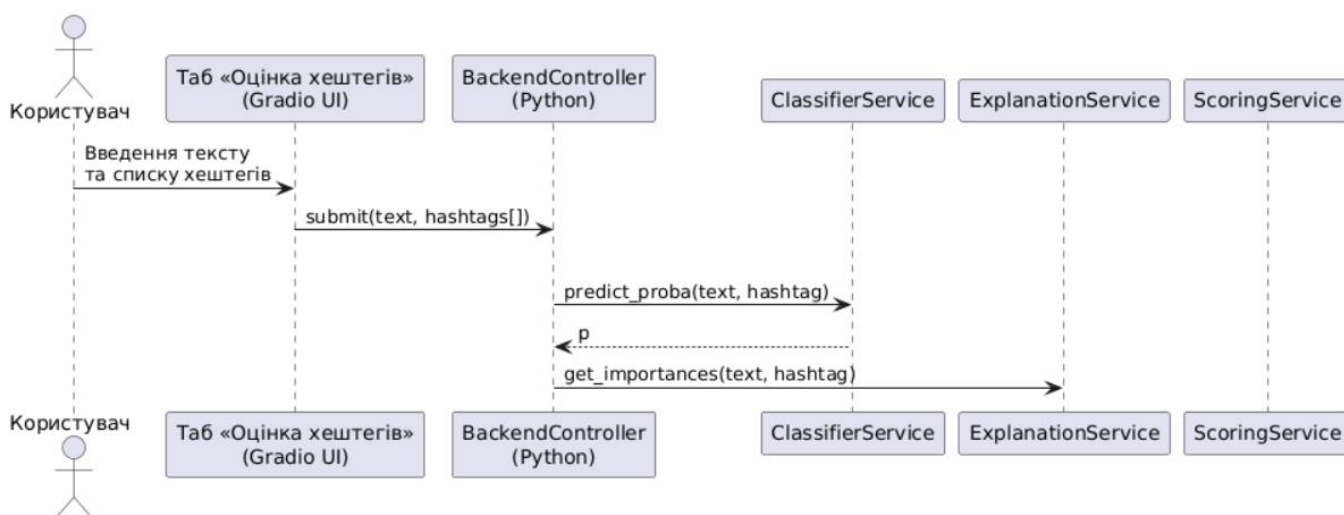


Рисунок 3.3 – Діаграма послідовності

Вихідною точкою сценарію є дії користувача у вкладці «Оцінка хештегів» вебінтерфейсу Gradio. Користувач вводить текст короткого повідомлення і перелік кандидатних хештегів, після чого ініціює відправлення форми. Ця дія відображена на діаграмі як повідомлення `submit(text, hashtags[])` від об'єкта «Таб Оцінка хештегів» до компонента `BackendController (Python)`. Таким чином інтерфейс виконує роль тонкого клієнта, який лише збирає вхідні дані та передає їх на серверний бік без власної аналітики.

Далі вся логіка зосереджена в серверному контролері, який координує роботу спеціалізованих сервісів. Для кожного окремого хештегу (на діаграмі показано типовий крок такого циклу) `BackendController` звертається до `ClassifierService` з викликом `predict_proba(text, hashtag)`. Класифікаційний сервіс використовує активну нейромережеву модель для обчислення ймовірності того, що даний хештег відповідає змісту тексту, і повертає це значення контролеру; на діаграмі зворотний рух даних позначено пунктирною стрілкою з результатом `p`. Таким чином фіксується чисто нейромережева оцінка релевантності без пояснювальної компоненти.

Після отримання ймовірності класифікації контролер ініціює побудову пояснення для тієї ж пари «текст-хештег». Для цього він викликає метод

`get_importances(text, hashtag)` service `ExplanationService`. Останній застосовує метод інтегрованих градієнтів до виходу моделі, визначає внесок окремих токенів тексту в рішення «хештег релевантний» і повертає контролеру структуру, що містить токени з відповідними вагами важливості. На діаграмі це відображено суцільною стрілкою від `BackendController` до `ExplanationService` і зворотним потоком даних із результатами атрибуцій.

Наявність на діаграмі «життєвої лінії» `ScoringService` без показаних повідомлень підкреслює, що в розширеній версії сценарію контролер, маючи значення ймовірності  $p$  та результати атрибуцій, передає їх до сервісу об'єднання оцінок для розрахунку показника узгодженості й інтегральної міри релевантності. У наведеному фрагменті основний акцент зміщено на взаємодію з класифікаційним та пояснювальним сервісами як ядром запропонованого методу. Таким чином діаграма послідовності деталізує часову структуру обробки запиту користувача: від введення тексту та списку хештегів у вебінтерфейсі до отримання нейромережевої оцінки та пояснень, які надалі використовуються для формування інтегральної оцінки релевантності кожного тегу.

На діаграмі станів (рисунок 3.4) подано життєвий цикл нейромережевої моделі, що використовується в інтелектуальній системі оцінювання хештегів. Діаграма описує, які стани може набувати модель у процесі роботи системи, які події спричиняють переходи між цими станами та як враховуються ситуації помилки, збереження кількох версій і вибору активної моделі. Такий опис дозволяє формально зафіксувати можливі сценарії роботи з моделлю та виключити некоректні конфігурації, наприклад використання неініціалізованої або неконтрольовано зміненої моделі в задачі оцінювання хештегів.

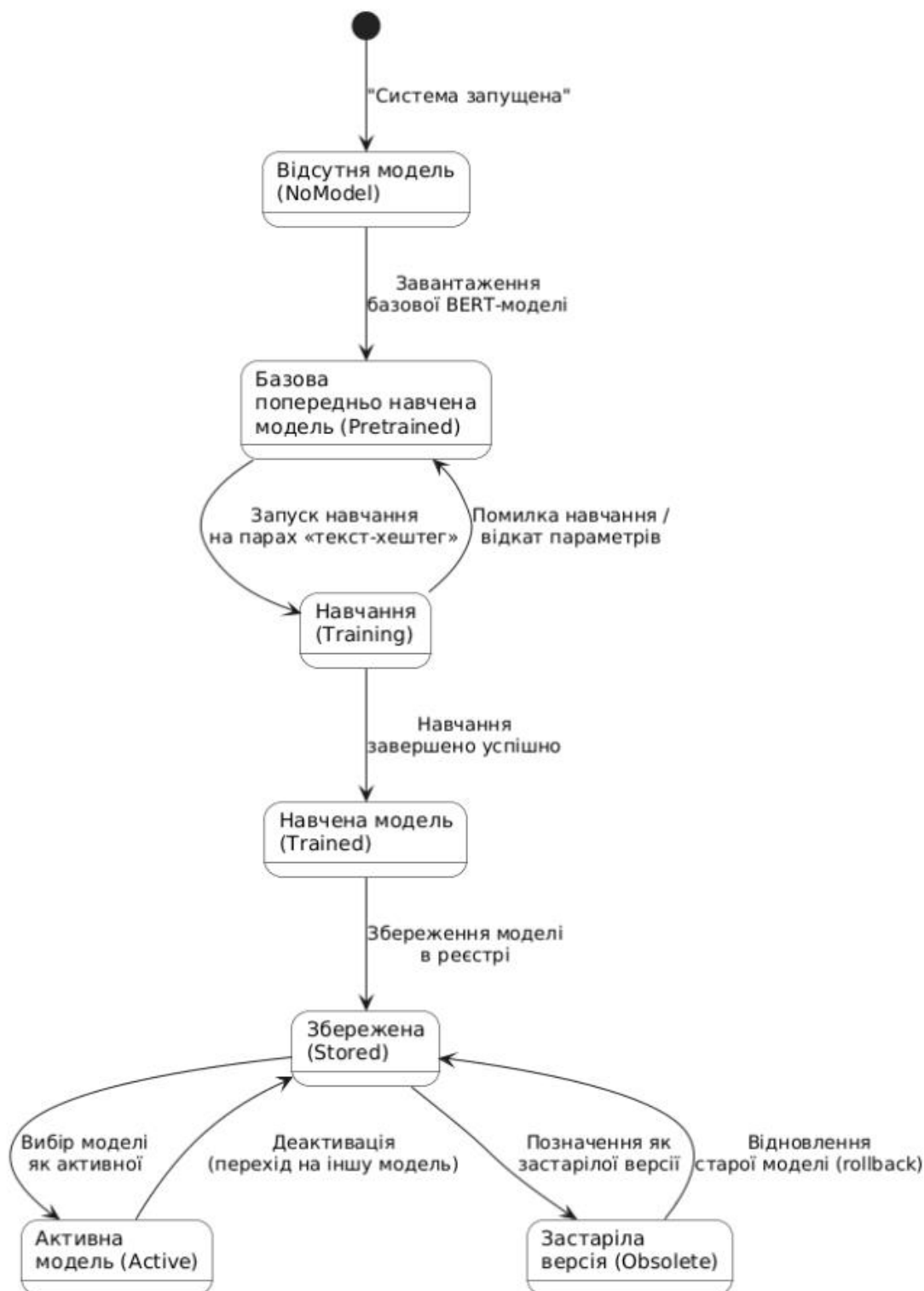


Рисунок 3.4 – Діаграма станів

Початковий маркер відповідає події «Система запущена» і веде до стану «Відсутня модель (NoModel)», який описує ситуацію, коли в оперативній пам'яті ще немає жодного екземпляра класифікатора. Перехід до стану «Базова попередньо навчена модель (Pretrained)» відбувається після успішного завантаження базової

BERT-моделі, що виступає відправною точкою для подальшого донавчання на предметно-орієнтованому датасеті пар «текст–хештег». Із цього стану можливий запуск процедури навчання: ініціюється перехід до стану «Навчання (Training)», у якому модель послідовно адаптується до задачі класифікації відповідності хештегів коротким текстам. Якщо під час цього процесу виникає помилка (наприклад, збій апаратних ресурсів чи некоректність даних), відбувається відкат параметрів і повернення до стану «Pretrained», що гарантує відсутність напівнавчених або нестабільних моделей.

У разі успішного завершення навчання система переходить до стану «Навчена модель (Trained)», де параметри моделі вже відображають статистичні закономірності предметної області, однак ще не зафіксовані в реєстрі. Наступним кроком є збереження моделі, що відповідає переходу до стану «Збережена (Stored)»: ваги мережі та супровідна інформація записуються у файлову структуру або спеціалізований сховище. Від збереженого стану можливі три подальші гілки. По-перше, модель може бути обрана як робоча, тоді вона переходить у стан «Активна модель (Active)» і використовується сервісами класифікації та пояснення в усіх поточних сеансах оцінювання хештегів; повернення з цього стану до «Stored» відображає деактивацію у разі вибору іншої моделі. По-друге, збережена модель може бути позначена як застаріла, що фіксується станом «Застаріла версія (Obsolete)» і означає, що вона більше не повинна використовуватися за замовчуванням, але зберігається для можливого аналізу чи відкату. По-третє, зі стану «Obsolete» можливе відновлення старої моделі (rollback), коли попередня версія знову стає збереженою й потенційно може бути активована. Така організація життєвого циклу моделі забезпечує контрольоване керування версіями, прозоре реагування на помилки навчання та чітке розмежування між моделями, що готуються, зберігаються, активно використовуються або виведені з експлуатації.

### **3.4 Метрики нейромережевого оцінювання релевантності хештегів**

У межах розробленої системи використовуються як класичні метрики якості бінарної класифікації, так і спеціалізовані показники, що відбивають пояснюваний характер оцінювання релевантності хештегів. Усі вони розраховуються для пар «текст-хештег», де клас 1 відповідає релевантному хештегу, а клас 0 – нерелевантному.

Основною оптимізаційною величиною при навчанні трансформерної моделі є функція втрат Cross-Entropy Loss. У реалізації використовується стандартний модуль CrossEntropyLoss бібліотеки PyTorch, який для кожного прикладу порівнює вектор логітів моделі з істинною бінарною міткою та обчислює середнє значення крос-ентропії по мінібатчу. Зменшення значення Cross-Entropy Loss на тренувальній та валідаційній вибірках інтерпретується як ознака того, що модель краще відокремлює релевантні хештеги від нерелевантних і коректніше оцінює їхні апостеріорні ймовірності.

Для кількісної оцінки класифікаційної якості на тренувальній та валідаційній вибірках використовується метрика Accuracy [59]. Вона визначається як відношення кількості правильно класифікованих пар «текст-хештег» (тобто таких, де передбачений клас збігається з істинною міткою) до загальної кількості пар у вибірці. У задачі оцінювання відповідності хештегів Accuracy показує, яку частку випадків модель правильно віднесла до «відповідає» або «не відповідає», і слугує інтуїтивно зрозумілим інтегральним показником якості.

Разом з Accuracy використовується метрика  $F_1$ -score, яка є гармонічним середнім між Precision та Recall для позитивного класу «релевантний хештег». Precision відображає, яка частка хештегів, позначених моделлю як релевантні, справді є коректними; Recall показує, яку частку всіх істинно релевантних хештегів модель змогла виявити.  $F_1$ -score поєднує обидва аспекти в одному значенні, штрафуючи випадки, коли одна з цих складових суттєво гірша за іншу. Це особливо важливо в умовах потенційної асиметрії помилок: хибно позитивні рішення призводять до появи нерелевантних тегів, а хибно негативні – до втрати корисних хештегів [60, 61].

На рівні окремої пари «текст-хештег» базовим виходом моделі є Probability of Relevance  $p$  – оцінка ймовірності релевантності, отримана після застосування softmax

до логітів класифікатора. Значення  $p \in [0;1]$  інтерпретується як міра впевненості нейромережі в тому, що конкретний хештег відповідає змісту заданого тексту. Цей показник використовується як для прийняття бінарного рішення (через поріг), так і як безперервна величина для подальшого об'єднання з пояснюваною складовою.

Спеціалізованою пояснюваною метрикою є Explanation Overlap Score  $S$ . Вона обчислюється на основі результатів методу Integrated Gradients: для пари «текст–хештег» визначається набір найважливіших токенів підпису, які дали найбільший внесок у рішення «хештег релевантний», та семантичне «ядро» самого хештегу, виділене шляхом очищення та розкладання на змістові компоненти. Показник  $S$  визначається як частка важливих слів тексту, що лексично або словотвірно пов'язані з ядром тегу. Таким чином він відображає, наскільки рішення моделі спирається саме на ті фрагменти тексту, які з погляду людини є смислово узгодженими з хештегом.

Фінальним інтегральним показником, який використовується для ранжування кандидатних хештегів, є Combined Relevance Score  $R$ . Він поєднує нейромережеву ймовірність релевантності  $p$  та пояснюваний показник узгодженості  $S$  з урахуванням вагового параметра  $\alpha$ : при більших значеннях  $\alpha$  домінує внесок власне класифікатора, при менших – більша вага надається структурі пояснення.

Таким чином  $R$  відображає одночасно і впевненість моделі, і якість її аргументації, що дозволяє формувати ранжовані списки хештегів з урахуванням як прогнозної точності, так і інтерпретованості рішень.

### **Висновки до розділу 3**

У третьому розділі здійснено цілісне проектування інтелектуальної системи оцінювання відповідності хештегів коротким текстам, яке поєднує методи обробки природної мови, глибинного навчання та пояснюваного ШІ з практичною реалізацією у вигляді інтерактивного програмного комплексу. Обґрунтовано вибір технологічного середовища, орієнтованого на мову Python, фреймворк PyTorch, інтерактивне хмарне середовище Google Colab та вебінтерфейс Gradio, що забезпечує повний цикл роботи з даними й моделями без розгортання окремої інфраструктури.

Такий стек дозволяє поєднати експериментальну гнучкість, характерну для дослідницьких проєктів, із можливістю створення зручного для користувача інструменту.

Розроблено структурну архітектуру системи у вигляді багаторівневої клієнт-серверної моделі. На рівні інтерфейсу проєкт системи організовано через тематичні вкладки, що відповідають основним сценаріям роботи: огляд датасету та статистик, навчання нейромережевої моделі, керування збереженими моделями, оцінювання релевантності хештегів. Кожна вкладка пов'язана з окремим серверним сервісом, який інкапсулює логіку відповідного сценарію та координує взаємодію з низькорівневими модулями. На рівні серверної логіки виділено модулі завантаження й попередньої обробки даних, формування пар «текст-хештег», навчання трансформерної моделі, оцінки релевантності, побудови пояснень та об'єднання оцінок у інтегральну міру. Робота з зовнішніми ресурсами ізольована через модулі доступу до CSV-датасету Instagram Reach Data та реєстру збережених моделей, що спрощує розгортання системи у різних середовищах.

На основі UML-діаграм виконано деталізацію елементів та поведінки системи. Діаграма варіантів використання зафіксувала узгоджений набір користувацьких сценаріїв і показала, що центральним елементом є оцінювання релевантності кандидатних хештегів, яке логічно пов'язане з формуванням пар «текст-хештег», навчанням моделі, переглядом даних та отриманням пояснень. Діаграма послідовності продемонструвала часову структуру обробки запиту: від введення тексту та списку тегів у вебінтерфейсі до виклику сервісів класифікації та пояснення, що реалізують обчислення ймовірності релевантності та атрибутів важливості слів. Діаграма станів описала життєвий цикл нейромережевої моделі – від базової попередньо навченої версії через етапи донавчання, збереження, активації та можливого позначення як застарілої з подальшим відновленням. Це дозволило формалізувати процедури керування моделями та забезпечити контрольовану роботу з декількома версіями класифікатора.

Окремо сформульовано систему метрик, що використовується для оцінювання якості нейромережевого компонента і пояснюваної частини методу. Для процесу

навчання та базової бінарної класифікації використано Cross-Entropy Loss як оптимізаційну функцію, Accuracy та  $F_1$ -score як основні інтегральні показники якості на тренувальній і валідаційній вибірках. На рівні окремих пар «текст–хештег» введено Probability of Relevance як ймовірнісну оцінку моделі. Для пояснюваної складової визначено метрику Explanation Overlap Score, що відображає узгодженість найважливіших з погляду моделі слів тексту із семантичним ядром хештегу. На їхній основі побудовано Combined Relevance Score, який поєднує прогностичну точність класифікатора та якість пояснення з урахуванням вагового параметра. Така система метрик забезпечує можливість не лише вимірювати формальну точність моделі, а й оцінювати, наскільки обґрунтованими є її рішення з точки зору інтерпретованості.

У підсумку третій розділ формує завершений проєкт інтелектуальної системи оцінювання відповідності хештегів коротким текстам: від обґрунтування вибору засобів розробки та побудови архітектури до детального опису елементів, сценаріїв взаємодії та системи метрик. Запропоновані структурні та поведінкові моделі створюють основу для подальшої програмної реалізації системи, її експериментальної апробації та аналізу ефективності запропонованого методу пояснюваного оцінювання релевантності хештегів.

## РОЗДІЛ 4 Експериментальне дослідження методу оцінювання відповідності хештегів коротким текстам засобами NLP

### 4.1 Структура компонентів інтелектуальної системи оцінювання відповідності хештегів коротким текстам

На рисунку 4.1 подано UML-діаграму класів, яка формалізує структуру компонентів інтелектуальної системи оцінювання відповідності хештегів коротким текстам. Вона відображає логічні сутності, їх атрибути, методи та взаємозв'язки між основними елементами ядра системи. Такий опис дає змогу узгодити архітектуру програмного коду з концептуальною моделлю даних, визначити залежності між модулями та забезпечити розширюваність системи у процесі подальшої експлуатації.

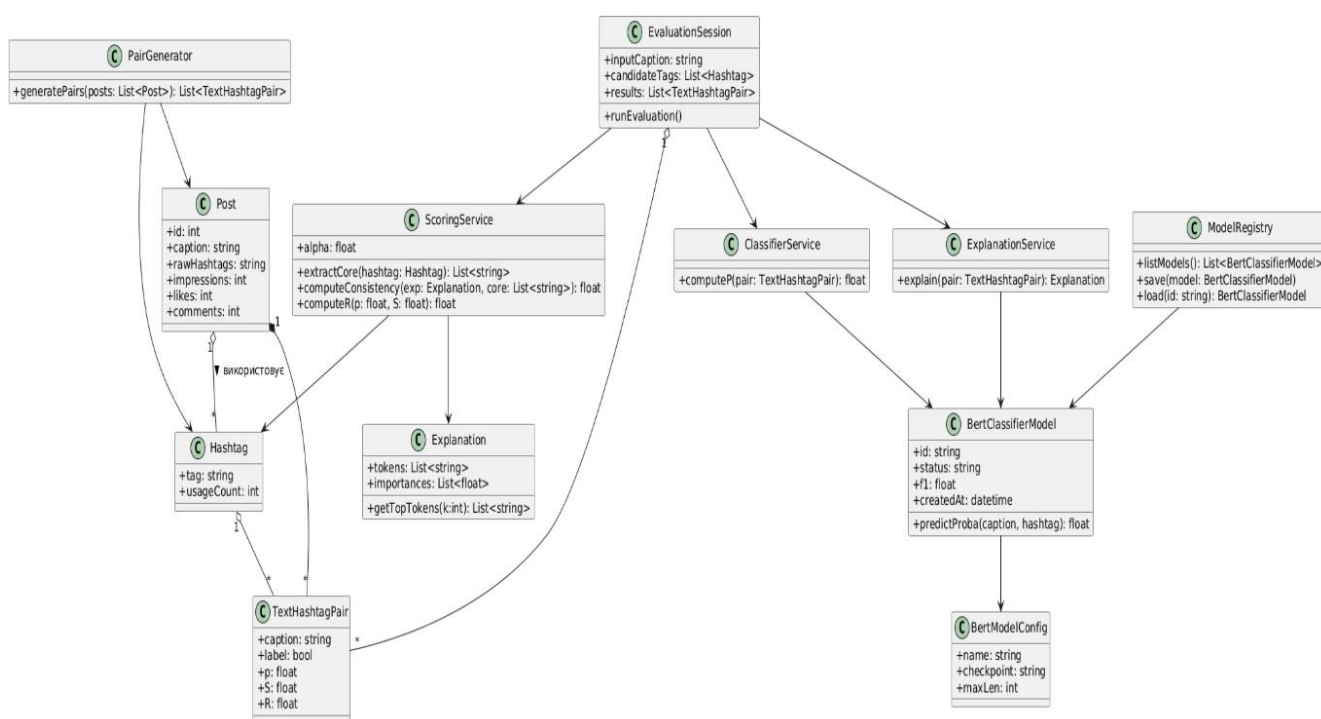


Рисунок 3.1 – Діаграма класів інтелектуальної системи

Центральним елементом є клас `EvaluationSession`, який координує виконання повного циклу оцінювання. Його атрибути `inputCaption`, `candidateTags` і `results` описують, відповідно, вхідні короткі тексти, множину кандидатних хештегів і сформовані результати у вигляді списку пар `TextHashtagPair`. Метод `runEvaluation()`

ініціює основний сценарій: для кожної пари викликаються сервіси класифікації, пояснення та обчислення інтегрального показника релевантності. Таким чином, цей клас забезпечує оркестрацію взаємодії між усіма підсистемами.

Групу даних, на основі яких виконується аналіз, становлять класи Post, Hashtag і TextHashtagPair. Клас Post описує допис у соціальній мережі, містить атрибути id, caption, rawHashtags, impressions, likes та comments і реалізує зв'язок «використовує» з класом Hashtag. Клас Hashtag містить атрибути tag і usageCount, що характеризують сам тег і частоту його вживання. Клас TextHashtagPair формує єдиний навчальний та оцінювальний елемент, який поєднує текст публікації (caption), відповідний хештег (tag), а також вихідні значення: label (істинна релевантність),  $p$  (ймовірність релевантності),  $S$  (пояснювальний показник узгодженості) і  $R$  (підсумковий інтегральний бал). Підготовкою таких пар займається клас PairGenerator, метод generatePairs(posts: List<Post>) якого генерує необхідні поєднання текстів і хештегів для подальшого навчання або тестування.

Обчислювальна частина системи реалізована через сервіси. Клас ClassifierService викликає метод computeP(pair: TextHashtagPair), який використовує активну модель для прогнозування ймовірності релевантності. Клас ExplanationService відповідає за формування пояснення рішення за допомогою методу explain(pair), що повертає об'єкт Explanation, у якому зберігаються списки токенів (tokens) та їхні ваги важливості (importances). Клас ScoringService виконує інтеграцію результатів нейромережевого оцінювання та пояснювальної складової. Він містить параметр alpha – ваговий коефіцієнт поєднання, методи extractCore(), computeConsistency() та computeR(), які реалізують виділення семантичного ядра хештегу, оцінювання узгодженості та розрахунок підсумкового інтегрального показника релевантності відповідно.

Серце класифікаційної частини становить клас BertClassifierModel, який містить параметри id, status, fl, createdAt та метод predictProba(caption, hashtag), що повертає оцінку ймовірності класу «релевантний». Модель має асоціацію з класом BertModelConfig, який визначає конфігураційні параметри – name, checkpoint та maxLen, що відповідають ідентифікатору моделі, контрольній точці збереження та

максимальній довжині вхідної послідовності токенів. Завдяки цій структурі реалізується можливість зберігати, відновлювати й перенавчати різні варіанти моделі без зміни коду основного сервісу.

Для централізованого керування всіма екземплярами моделей передбачено клас `ModelRegistry`. Його методи `listModels()`, `save()` і `load()` реалізують, відповідно, перелік доступних моделей, збереження поточного стану та завантаження потрібного класифікатора. Саме цей клас забезпечує взаємодію між етапами навчання, тестування й використання моделі у виробничому режимі, гарантуючи узгодженість станів, описаних на діаграмі станів у попередньому розділі.

Таким чином, наведена діаграма класів відображає повну структуру компонентів інтелектуальної системи: від рівня даних і навчальних сутностей до рівня сервісів та керування моделями. Вона наводить логічну зв'язність між основними об'єктами, дописами, хештегами, парами «текст-тег», поясненнями, моделями та сервісами і забезпечує узгоджене функціонування всіх елементів системи під час експериментального оцінювання релевантності хештегів коротким текстом.

Подана схема (рисунок 4.2) відображає варіант розгортання інтелектуальної системи оцінювання відповідності хештегів у типовому для дослідження середовищі `Google Colab`. У верхній частині розміщено «Браузер користувача», у якому завантажується згенерований `Gradio` інтерфейс. З технічного погляду цей інтерфейс являє собою `HTML`-сторінку з компонентами на `JavaScript`, що забезпечують відображення вкладок, текстових полів, таблиць із результатами та обробку дій користувача. Усі введені дані – текст допису, список кандидатних хештегів, параметри навчання, шлях до датасету – залишаються в межах браузера доти, доки не буде сформовано `HTTP`-запит до серверної частини застосунку.

Середній блок позначає середовище «`Google Colab / Python runtime`», де фактично виконується вся прикладна логіка системи. Безпосередньо з браузером взаємодіє `Gradio backend`, який виступає легковаговим вебсервером. Він приймає запити від інтерфейсу, десеріалізує параметри, викликає відповідні `Python`-функції та повертає структуровані відповіді у форматі, придатному для відображення на стороні

клієнта. Gradio backend не містить власної аналітики, а лише маршрутизує виклики до внутрішніх модулів логіки системи.

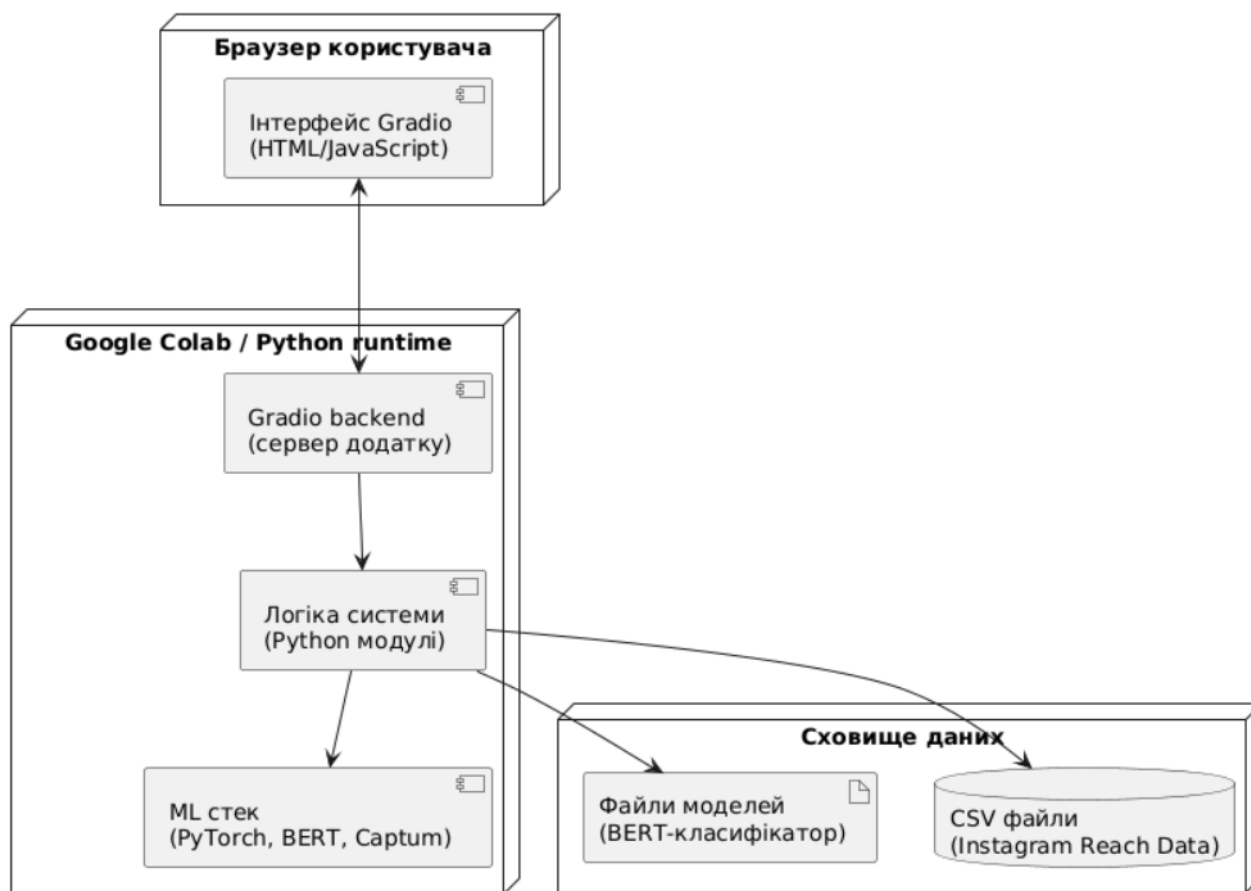


Рисунок 4.2 – Схема розгортання інтелектуальної системи

У блоці «Логіка системи (Python модулі)» зосереджені основні компоненти, реалізовані мовою Python: модулі завантаження та попередньої обробки датасету, формування пар «текст-хештег», донавчання BERT-класифікатора, обчислення метрик якості, побудови атрибуцій Integrated Gradients, розрахунку міри узгодженості та інтегрального скору релевантності. Саме тут реалізовано алгоритми, подані у вигляді псевдокоду в попередньому підрозділі, а також логіку підтримки кількох моделей та їхнього вибору для оцінювання. Ці модулі взаємодіють з ML-стеком, винесеним на окремий рівень.

Нижче розміщено блок «ML стек (PyTorch, BERT, Captum)», який містить бібліотеки глибокого навчання та пояснюваного ШІ. Під час виклику відповідних функцій логіки системи PyTorch забезпечує виконання операцій тензорної алгебри на

CPU або GPU, модулі трансформерів реалізують архітектуру BERT і пов'язані з нею процедури токенізації та інференсу, а Cartum відповідає за обчислення інтегрованих градієнтів. Таким чином Python-модулі вищого рівня використовують ML-стек як спеціалізований обчислювальний шар, який інкапсулює всі низькорівневі деталі роботи нейромережових моделей.

Праворуч відображено окремий контур «Сховище даних», який реалізується на рівні файлової системи середовища Colab. У ньому зберігаються файли моделей (зокрема ваги BERT-класифікатора й конфігураційні файли, що підтримуються відповідними бібліотеками) та CSV-файли з даними Instagram Reach Data. Логіка системи звертається до цього сховища під час початкового завантаження датасету, формування навчальних вибірок, а також при збереженні й подальшому відтворенні навчених моделей через реєстр моделей. Стрілки на діаграмі від «Логіки системи» до обох підблоків сховища фіксують двосторонній характер взаємодії: читання вхідних даних та запис результатів навчання.

Таким чином схема демонструє три основні шари розгортання: клієнтський рівень у браузері, серверний обчислювальний рівень у середовищі Google Colab та рівень зберігання даних і моделей. Такий поділ забезпечує чітке розмежування відповідальностей: користувач працює лише з інтерфейсом Gradio, обчислювально складні завдання виконуються у Python-runtime з доступом до ML-стеку, а всі дані та результати навчання зберігаються у файловому сховищі, що підвищує відтворюваність експериментів і спрощує перенесення системи на інші середовища виконання.

## **4.2 Особливості створення інтелектуальної системи**

Доцільно виділити кілька ключових алгоритмів, на яких ґрунтується створення інтелектуальної системи: формування навчальних пар «текст-хештег», донавчання мовної моделі-класифікатора та пояснюване оцінювання кандидатних хештегів. Усі вони реалізовані в одному програмному модулі, що працює в середовищі Google Colab і безпосередньо інтегрується з вебінтерфейсом Gradio, але

далі подані в абстрактному псевдокоді, незалежному від конкретної мови програмування.

Алгоритм формування навчальних пар «текст-хештег»:

```

Algorithm GeneratePairs(DataFrame DF, integer NEG_PER_POS):
  Set ALL_TAGS := порожня множина
  For кожен рядок r у DF:
    tags_r := множина хештегів, отриманих з r.Hashtags
    Додати всі елементи tags_r до множини ALL_TAGS
  Список ALL_TAGS_LIST := впорядкований список елементів ALL_TAGS
  Список PAIRS := порожній список
  For кожен рядок r у DF:
    caption := r.Caption
    tags_r := унікальні хештеги з r.Hashtags
    Якщо tags_r порожня:
      перейти до наступного рядка
    For кожен тег t у tags_r:
      Додати до PAIRS потрійку (caption, t, 1)
      Повторити NEG_PER_POS разів:
        Випадково вибрати тег neg з ALL_TAGS_LIST
        Поки neg належить tags_r:
          випадково вибрати інший тег neg
        Додати до PAIRS потрійку (caption, neg, 0)
  Повернути PAIRS та ALL_TAGS_LIST

```

Зміст цього алгоритму полягає у перетворенні початкового табличного датасету на навчальну вибірку для моделі бінарної класифікації. Спочатку зі стовпця з хештегами будується глобальна множина усіх унікальних тегів, що забезпечує достатню різноманітність негативних прикладів. Для кожного допису беруться всі вказані автором хештеги; кожен із них формує позитивну пару «Caption-Hashtag» з міткою «1». Далі до кожного позитивного прикладу додається задана кількість негативів, що отримуються випадковою вибіркою інших тегів із глобальної множини з перевіркою, що вони не належать до поточного допису. У результаті формується збалансована або контрольовано розбалансована вибірка пар, у якій модель навчається відрізнити семантично пов'язані хештеги від випадкових.

Алгоритм донавчання мовної моделі-класифікатора наведено нижче:

```

Algorithm TrainClassifier(DF, BaseModelName, EPOCHS, LR, BATCH_SIZE, NEG_PER_POS):
  (PAIRS, ALL_TAGS) := GeneratePairs(DF, NEG_PER_POS)
  labels := список міток з PAIRS
  (TRAIN_PAIRS, VAL_PAIRS) := стратифікований поділ PAIRS на train/val
  Ініціалізувати токенизатор Tokenizer відповідно до BaseModelName
  Побудувати об'єкти TrainDataset та ValDataset на основі TRAIN_PAIRS, VAL_PAIRS
  Побудувати TrainLoader і ValLoader з розміром пакета BATCH_SIZE
  Модель Model := завантажити попередньо навчену мовну модель BaseModelName
                  з двома вихідними класами
  Перенести Model на пристрій DEVICE
  Оптимізатор Opt := AdamW(Model.parameters, LR)
  Функція втрат Crit := CrossEntropyLoss
  Історія History := порожній список
  For epoch від 1 до EPOCHS:
    Model у режимі train
    Ініціалізувати накопичувачі TrainLoss, TrainPreds, TrainLabels
    For кожен batch у TrainLoader:
      Перенести batch на DEVICE
      Labels := batch.labels
      Обнулити градієнти Opt
      Outputs      :=      Model(batch.input_ids,      batch.attention_mask,
batch.token_type_ids, labels)
      Loss := Outputs.loss
      Logits := Outputs.logits
      Виконати зворотне поширення градієнта Loss.backward()
      Оновити ваги Opt.step()
      Оновити TrainLoss та додати предикти й мітки до TrainPreds, TrainLabels
      Обчислити TrainAccuracy та TrainF1 за TrainPreds, TrainLabels

```

Для валідації мовної моделі-класифікатора оцінка виконується алгоритмом, наведеним нижче:

```

// Фаза валідації
  Model у режимі eval
  Ініціалізувати ValLoss, ValPreds, ValLabels

  Без градієнтів:
    For кожен batch у ValLoader:

```

```

Перенести batch на DEVICE
Outputs := Model(...)
Loss := Outputs.loss
Logits := Outputs.logits
Оновити ValLoss та додати предикти й мітки до ValPreds, ValLabels
Обчислити ValAccuracy та ValF1 за ValPreds, ValLabels
Додати запис до History з усіма метриками за поточну епоху
Зберегти ваги Model, токенизатор та таблицю History у директорію models/run_name
Повернути посилання на збережену модель та історію навчання

```

Цей алгоритм реалізує стандартний цикл донавчання трансформерної моделі у постановці бінарної класифікації, але з урахуванням специфіки сформованих пар «текст-хештег». На вхід подається табличний датасет; на першому кроці викликається описаний вище алгоритм GeneratePairs, після чого пари стратифіковано розбиваються на тренувальну та валідаційну вибірки. Кожна пара подається токенизатору у вигляді конкатенації підпису й хештегу, що забезпечує спільне кодування контексту і тегу в одному векторному просторі. Далі запускається класичний епохальний цикл: у режимі навчання обчислюється функція втрат Cross-Entropy Loss, виконується зворотне поширення та оновлення ваг; у режимі валідації вимірюються Accuracy та F<sub>1</sub>-score. Історія значень метрик зберігається для подальшого аналізу, а сама модель разом із токенизатором записується в окрему директорію, що дозволяє керувати кількома версіями класифікатора. На рисунку 4.3 наведено виведення метрик після завершення процесу навчання нейромережі.

Запустити навчання

Модель збережено в models/run\_bert\_1

Історія навчання

epoch	train_loss	train_acc	train_f1	val_loss	val_acc	val_f1
1	0.36033633345518185	0.8498515219005197	0.7715334651228467	0.2057104459573895	0.9257609502598366	0.88558352402746
2	0.17577298179631135	0.9367112100965108	0.9044014578076816	0.19812704709309423	0.9287305122494433	0.8942731277533039
3	0.11197437156225487	0.96362286562732	0.9454342984409799	0.12011082829429084	0.9561989606533037	0.9346622369878184

Рисунок 4.3 – Приклад виведення метрик навчання

Алгоритм пояснюваного оцінювання релевантності хештегів:

```

Algorithm ScoreHashtagsForCaption(
  caption: string,
  hashtags: List<string>,
  Tokenizer,
  Model,
  alpha: float,
  top_k: integer
) -> Table:
  Результати := порожній список
  For кожен тег h у списку hashtags:
    p := ClassifierScore(caption, h, Tokenizer, Model)
    S := ExplanationOverlapScore(caption, h, Tokenizer, Model, top_k)
    R := alpha * p + (1 - alpha) * S
    Додати до Результати запис:
      (hashtag = h, p_classifier = p, overlap = S, combined = R)
  Впорядкувати Результати за полем combined у спадаючому порядку
  Повернути впорядковану таблицю

```

У цьому блоці реалізовано власне пояснюваний механізм оцінювання. Зовнішній алгоритм `ScoreHashtagsForCaption` приймає один текст і список хештегів-кандидатів. Для кожного з них усередині відбуваються дві операції: обчислення нейромережевої ймовірності релевантності та розрахунок пояснювального показника узгодженості. Функція `ClassifierScore` використовує треновану модель для отримання ймовірності класу «відповідає» за допомогою стандартної процедури токенизації й проходження через мережу. Таким чином отримується числова міра впевненості моделі.

Також допоміжні функції можна подати у вигляді алгоритму:

```

Function ClassifierScore(caption, hashtag, Tokenizer, Model) -> float:
  Побудувати вхідну послідовність із caption та hashtag
  Перетворити послідовність на токени за допомогою Tokenizer
  Передати токени до Model у режимі eval
  Отримати вектор логітів, застосувати softmax
  Повернути ймовірність класу "релевантний"
Function ExplanationOverlapScore(caption, hashtag, Tokenizer, Model, top_k) -> float:
  (tokens, segment_ids, scores) := GetTokenImportances(caption, hashtag, Tokenizer,
Model)

```

```

Вибрати токени caption, що належать до сегмента тексту й не є службовими
Для кожного такого токена взяти абсолютне значення його важливості
Визначити top_k токенів з максимальною важливістю
CanonTokens := нормалізовані форми цих токенів
Core := ExtractHashtagCore(hashtag) // множина базових частин тегу
Якщо Core порожня або CanonTokens порожні:
    повернути  $\emptyset$ 
matches := кількість токенів із CanonTokens,
            для яких існує елемент w у Core,
            що містить t або міститься в t
Повернути matches / |CanonTokens|

```

Функція `ExplanationOverlapScore` спирається на метод інтегрованих градієнтів. Спершу викликається допоміжна процедура `GetTokenImportances`, яка повертає токени вхідної послідовності, сегментні індекси та нормовані значення важливості для кожного токена. Далі відбираються лише токени, що належать до тексту підпису, і з них вибирається `top_k` найбільш важливих. Ці токени нормуються до «канонічної» форми (без службових символів і суфіксів). Паралельно з хештегу виділяється його семантичне ядро: тег очищується від символу «#», розбивається за символом підкреслення, фрагменти переводяться до нижнього регістру. Показник  $S$  визначається як частка важливих слів тексту, для яких спостерігається лексичний або словотвірний збіг із елементами ядра тегу. Ця величина відображає, наскільки рішення моделі «підтримується» тими фрагментами тексту, які інтуїтивно пов'язані з хештегом.

У підсумку інтегральний показник  $R$  комбінує два джерела інформації: ймовірність  $p$ , що характеризує прогностичну здатність нейромережі, та  $S$ , який описує якість аргументації з погляду узгодженості пояснення з семантикою тегу. Ваговий коефіцієнт  $\alpha$  дозволяє налаштовувати баланс між ними залежно від пріоритету точності або інтерпретованості.

### 4.3 Особливості тестування розробленої інтелектуальної системи

Першим тестовим випадком є перевірка на справність пунктів головного меню розробленої інтелектуальної системи (рисунок 4.4).

### Лабораторія оцінки релевантності хештегів (explanation-based)

[Дані](#)   [Навчання](#)   [Моделі](#)   [Оцінка релевантності](#)

Шлях до датасету

Рисунок 4.4 – Головне меню

Кроки тест-кейса наведено у таблиці 4.1.

Таблиця 4.1 – Тест-кейс AI0001

<b>Тест-кейс ID: AI0001</b>	<b>Пріоритет: 1</b>	<b>Створено: 9.10.2025,</b> Ольга ТРОХИМЧУК
<b>Назва:</b> Перевірка працездатності головного меню		
<b>Кроки</b>	<b>Очікуваний результат</b>	
1. Запустити вебінтерфейс в середовищі «Google Colab».	Запущено вебінтерфейс.	
2. Натиснути на пункт меню «Навчання».	Відбувся перехід на сторінку «Навчання».	
3. Перейти на сторінку навчання нейромережових моделей.	Відбувся перехід на сторінку «Навчання».	
4. Натиснути на кнопку «Моделі».	Відбувся перехід на сторінку «Моделі».	
5. Перейти на сторінку «Моделі».	Відбувся перехід на сторінку «Моделі».	
6. Натиснути на кнопку «Оцінка релевантності».	Відбувся перехід на сторінку «Оцінка релевантності».	
7. Перейти на сторінку «Оцінка релевантності».	Відбувся перехід на сторінку «Оцінка релевантності».	
<b>Результат виконання тест-кейсу:</b> пройдено успішно		

Результат успішного виконання тест-кейсу AI0001 наведено на рисунку 4.5.

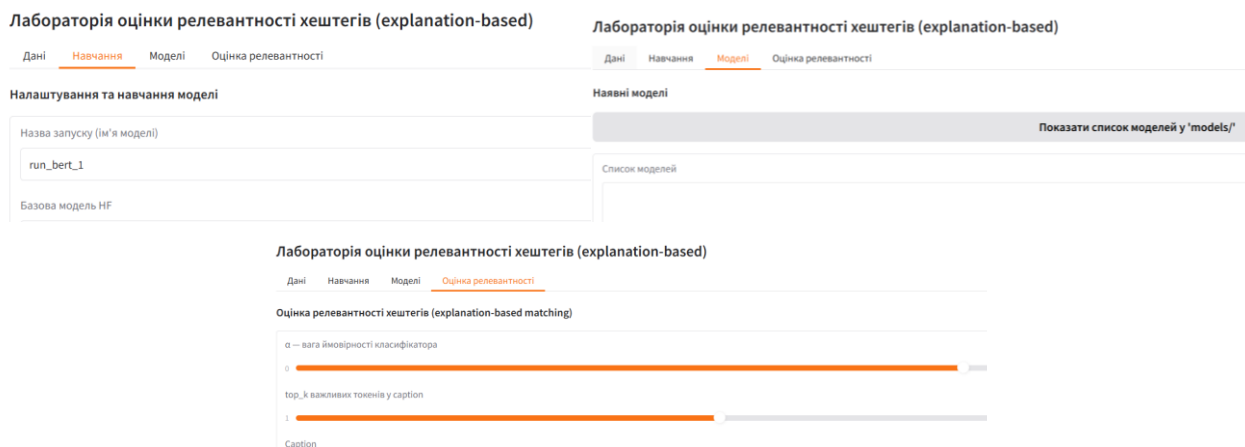


Рисунок 4.5 – Ілюстрація переходів на сторінки

Тест-кейс AI0002 (таблиця 4.2) перевіряє базову працездатність підсистеми завантаження даних на вкладці «Дані» та коректність обробки валідного CSV-файлу.

Таблиця 4.2 – Тест-кейс AI0002

<b>Тест-кейс ID: AI0002</b>	<b>Пріоритет: 1</b>	<b>Створено: 15.10.2025, Ольга ТРОХИМЧУК</b>
<b>Назва:</b> Завантаження коректного датасету		
<b>Кроки</b>	<b>Очікуваний результат</b>	
<ol style="list-style-type: none"> <li>Запустити вебінтерфейс у середовищі Google Colab.</li> <li>Перейти на вкладку «Дані».</li> <li>У полі «Шлях до датасету» вказати шлях до файлу Instagram data.csv.</li> <li>Натиснути кнопку «Завантажити дані».</li> <li>Дочекатися відображення інформаційного повідомлення та таблиці перших 20 рядків.</li> </ol>	<p>Вебінтерфейс успішно запущено. У блоці «Розмір датасету...» з'являється текст із кількістю рядків і колонок без повідомлень про помилки.</p> <p>У таблиці «Перші 20 рядків» відображаються дані (не порожня, містить стовпці Caption, Hashtags тощо).</p>	
<b>Результат виконання тест-кейсу:</b> пройдено успішно		

Мета даного тест-кейсу підтвердити, що датасет Instagram Reach Data коректно читається, фільтрується від порожніх значень і відображається у вигляді

інформаційного повідомлення та таблиці перших записів. Особливу увагу приділяють наявності колонок Caption і Hashtags та відсутності помилок при читанні файлу. Результати проходження тесту, кроки якого наведені в таблиці 4.2 наведено на рисунку 4.6.

Завантажити дані

Розмір датасету: 119 рядків, 13 колонок

Перші 20 рядків

Impressio...	From No...	From Hashtags	From Explore	From Other	Saves	Comments	Shar...	Likes	Profile Visi...	Follows	Caption
3884	2046	1214	329	43	74	7	10	144	9	2	Here are some of the most important soft skills t
2621	1543	599	333	25	22	5	1	76	26	0	Learn how to analyze a candlestick chart as a dat
3541	2071	628	500	60	135	4	9	124	12	6	Here are some of the best books that you can foll
3749	2384	857	248	49	155	6	8	159	36	4	Here are some of the best data analysis project i
4115	2609	1104	178	46	122	6	3	191	31	6	Here are two best ways to count the number of let
2218	1597	411	162	15	28	6	3	81	29	4	Learn the implementation of AlexNet Convolutional
3234	2414	476	185	75	122	8	14	151	15	0	HereDs how to get the live stock price data of an
4344	2168	1274	673	40	119	7	11	162	8	2	Here are some of the most important Python librar
3216	2524	212	201	223	121	5	5	142	20	4	HereDs how to visualize an interactive candlestic
9453	2525	5799	208	794	100	6	10	294	181	42	Omicron Variant Sentiment Analysis using Python:
5055	2017	2351	298	108	101	7	11	159	17	6	In Data Science, Time Series Analysis is a method
4002	3401	278	128	73	111	17	18	205	16	2	Here are some of the highest paying skills in 202

Індекс рядка для перегляду

Рисунок 4.6 – Результат успішного проходження тест-кейсу AI0002

Наступний тест-кейс (таблиця 4.3) перевіряє повний цикл навчання трансформерної моделі на основі завантаженого датасету з використанням валідних гіперпараметрів. Мета – підтвердити, що модель успішно проходить етапи формування пар «текст-хештег», навчання, валідації та збереження результатів у директорії models, а історія метрик відображається в інтерфейсі. Особливість тесту – одночасна перевірка як обчислювальної частини (PyTorch, BERT), так і правильної фіксації метрик у файлі metrics.csv.

Таблиця 4.3 – Тест-кейс AI0003

<b>Тест-кейс ID: AI0003</b>	<b>Пріоритет: 1</b>	<b>Створено:5.11.2025,</b>
-----------------------------	---------------------	----------------------------

		Ольга ТРОХИМЧУК
<b>Назва:</b> Навчання моделі з коректними параметрами		
	<b>Кроки</b>	<b>Очікуваний результат</b>
	<ol style="list-style-type: none"> <li>1. Виконати тест-кейс AI0002 (успішно завантажити датасет).</li> <li>2. Перейти на вкладку «Навчання».</li> <li>3. У полі «Назва запуску (ім'я моделі)» вказати, наприклад, run_bert_test.</li> <li>4. Переконавшись, що поле «Базова модель HF» містить bert-base-uncased (за замовчуванням).</li> <li>5. Задати 1–2 епохи, learning rate 2e-5, batch size 8.</li> <li>6. Натиснути «Запустити навчання».</li> <li>7. Дочекатися завершення процесу.</li> </ol>	<p>У полі статусу з'являється повідомлення виду: Модель збережено в models/run_bert_test.</p> <p>У таблиці «Історія навчання» відображаються щонайменше один-два рядки з колонками epoch, train_loss, train_acc, train_f1, val_loss, val_acc, val_f1.</p> <p>В директорії models/ створюється підкаталог run_bert_test з файлами моделі та metrics.csv.</p>
<b>Результат виконання тест-кейсу:</b> пройдено успішно		

Результат виконання кроків таблиці 4.3 наведено на рисунку 4.7

Модель збережено в models/run\_bert\_1

Історія навчання

epoch	train_loss	train_acc	train_f1	val_loss	val_acc	val_f1
1	0.36033633345518185	0.8498515219005197	0.7715334651228467	0.2057104459573095	0.9257609502598366	0.88558352402746
2	0.17577298179631135	0.9367112100965188	0.9044014578076816	0.19812704709309423	0.9287305122494433	0.8942731277533039

Рисунок 4.7 – Результат успішного виконання тест-кейсу

Наступний тест-кейс (таблиця 4.4) перевіряє працездатність механізму реєстру моделей, зокрема відображення списку доступних моделей та завантаження вибраної моделі в оперативну пам'ять. Мета – підтвердити, що користувач може обрати потрібний екземпляр класифікатора для подальшого використання без

повторного навчання. Особливу увагу звертають на коректність повідомлень про успішне завантаження та відсутність помилок при ініціалізації токенизатора і моделі.

Таблиця 4.4 – Тест-кейс AI0004

<b>Тест-кейс ID: AI0004</b>	<b>Пріоритет: 1</b>	<b>Створено: 6.11.2025,</b> Ольга ТРОХИМЧУК
<b>Назва:</b> Завантаження збереженої моделі з реєстру		
<b>Кроки</b>		<b>Очікуваний результат</b>
<ol style="list-style-type: none"> <li>1. Виконати тест-кейс AI0004 (створити модель run_bert_test).</li> <li>2. Перейти на вкладку «Моделі».</li> <li>3. Натиснути кнопку «Показати список моделей у 'models/'».</li> <li>4. Перевірити, що у текстовому полі «Список моделей» присутній шлях models/run_bert_test.</li> <li>5. У полі «Шлях до моделі для завантаження» вказати models/run_bert_test (якщо ще не вказано).</li> <li>6. Натиснути кнопку «Завантажити модель».</li> </ol>		<p>У полі статусу відображається повідомлення: Модель завантажено з models/run_bert_test.</p> <p>Помилки виконання не виникає.</p> <p>Подальші операції (перехід на вкладку «Оцінка релевантності») показують, що модель доступна для використання (відсутні попередження про необхідність завантаження моделі).</p>
<b>Результат виконання тест-кейсу:</b> пройдено успішно		

Вкладка «Моделі» для виконання тест кейсу AI0004 наведена на рисунку 4.8.

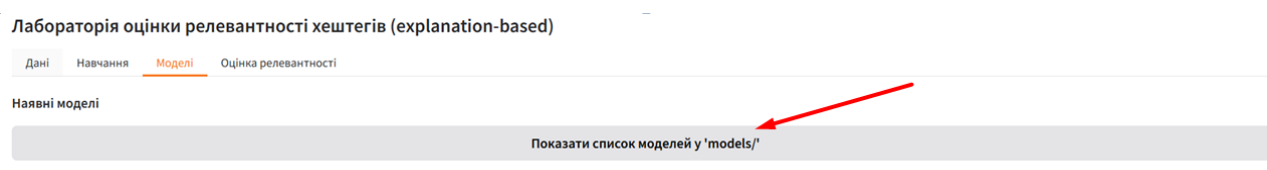


Рисунок 4.8 – Вкладка «Моделі»

Наступний тест-кейс (таблиця 4.5) перевіряє ключовий функціональний сценарій системи – оцінювання релевантності множини кандидатних хештегів для заданого тексту з використанням завантаженої нейромережевої моделі. Мета – переконатися, що система обчислює ймовірності  $p$ , показник `overlap` та інтегральний `combined score` для кожного хештегу, формує таблицю результатів і коректно ранжує теги за спаданням `combined`. Особливістю є контроль того, що розрахунки виконуються без помилок для кількох тегів одночасно, а інтерфейс наочно відображає всі три показники.

Таблиця 4.5 – Тест-кейс AI0005

<b>Тест-кейс ID: AI0005</b>	<b>Пріоритет: 1</b>	<b>Створено: 6.11.2025,</b> Ольга ТРОХИМЧУК
<b>Назва:</b> Оцінка релевантності хештегів при завантаженій моделі		
<b>Кроки</b>		<b>Очікуваний результат</b>
1. Виконати тест-кейс AI0004 (модель успішно завантажена).		У полі статусу з'являється повідомлення: Оцінка виконана.
2. Перейти на вкладку «Оцінка релевантності».		У таблиці «Результати» відображається не менше одного рядка, кожний рядок містить стовпці <code>hashtag</code> , <code>p_classifier</code> , <code>overlap</code> , <code>combined</code> .
3. У полі «Caption» залишити текст за замовчуванням або ввести власний короткий підпис англійською.		
4. У полі «Хештеги через пробіл» ввести кілька тегів, наприклад: <code>#finance #datavisualization #cooking</code> .		Сортування здійснено за спаданням стовпця <code>combined</code> (першим відображається хештег з найбільшою інтегральною оцінкою).
5. За потреби змінити параметри $\alpha$ та <code>top_k</code> або залишити значення за замовчуванням.		
6. Натиснути кнопку «Оцінити».		
<b>Результат виконання тест-кейсу:</b> пройдено успішно		

Результат виконання тест-кейсу наведено на рисунку 4.9.

Caption

Here are some of the most important data visualizations that every Financial Data Analyst/Scientist should know.

Хештеги через пробіл

#finance #datavisualization #cooking

Оцінити

Оцінка виконана.

Результати

hashtag	p_classifier	overlap	combined
#finance	0.9869983922088994	0	0.6988932745456695
#datavisualization	0.7626652717598332	0.5	0.6838656982313233
#cooking	0.084998677814938677	0	0.083493474470451474

Рисунок 4.9 – Результат виконання тест-кейсу AI0005

Узагальнюючи результати проведеного тестування, можна стверджувати, що розроблена інтелектуальна система оцінювання відповідності хештегів коротким текстам продемонструвала коректну роботу основних функціональних підсистем. Послідовне виконання тест-кейсів AI0001-AI0005 підтвердило працездатність навігації головного меню, стабільність механізму завантаження та первинної обробки датасету, коректність циклу навчання нейромережевої моделі й збереження її параметрів, а також ефективність реєстру моделей, що дозволяє безперервно переходити від етапу тренування до практичного використання класифікатора. Окремо продемонстровано, що ключовий сценарій застосування, а саме оцінка релевантності множини кандидатних хештегів, реалізується відповідно до проєктних вимог: система коректно обчислює й візуалізує нейромережеву ймовірність, показник перекриття пояснення з ядром теґу та інтегральну метрику, забезпечуючи семантично обґрунтоване ранжування теґів.

Проведені випробування охоплюють повний логічний ланцюг роботи комплексу від взаємодії користувача з вебінтерфейсом у середовищі Google Colab до звернення до файлового сховища й виклику глибинної моделі. Отримані результати свідчать про узгодженість між архітектурним проєктуванням і програмною реалізацією, відсутність критичних помилок у типових сценаріях і достатній рівень готовності системи до подальших експериментальних досліджень. Разом із тим сформований набір тест-кейсів може бути розширений стрес-тестами, перевіркою

граничних випадків та негативних сценаріїв, що стане наступним кроком у напрямі підвищення надійності та масштабованості інтелектуальної системи.

#### 4.4 Експериментальна установка

Експериментальне дослідження методу оцінювання відповідності хештегів коротким текстам виконувалося в середовищі Google Colab, яке поєднує модель інтерактивного блокнота з можливістю використання апаратного прискорення на графічному процесорі. Такий вибір дає змогу відтворювати експерименти в уніфікованому хмарному середовищі, не прив'язуючись до конкретної локальної конфігурації, а також забезпечує достатню обчислювальну потужність для донавчання трансформерної моделі BERT на сформованому наборі пар «текст-хештег». Для забезпечення відтворюваності результатів фіксувалися початкові значення генераторів випадкових чисел для Python, NumPy [62] і PyTorch [63], а всі операції виконувалися в одному ноутбучі, що містив як код моделі, так і інтерфейсні компоненти.

Базовим програмним засобом для реалізації нейромережевої частини експериментальної установки є фреймворк PyTorch. Він використовується для опису архітектури моделі, організації навчального циклу, оптимізації параметрів і виконання обчислень на GPU. Оскільки PyTorch реалізує об'єктно-орієнтовану модель представлення нейромереж, це дозволило налаштувати класифікаційну голову над трансформерним енкодером, визначати функцію втрат, організувати мінібатчі даних і відстежувати значення метрик на тренувальній і валідаційній вибірках. Механізми автоматичного диференціювання PyTorch додатково використовуються на етапі пояснюваного аналізу, де потрібен доступ до градієнтів моделі відносно вхідних ембедінгів.

Для роботи з попередньо навченими мовними моделями та токенизаторами використовується бібліотека transformers, яка надає засоби завантаження моделі BERT-base-uncased та її інтеграції в задачу бінарної класифікації. У рамках експериментальної установки transformers забезпечує побудову спільного векторного

подання пари «caption-hashtag», автоматичну обробку спеціальних токенів, масок уваги й сегментів, а також узгодження формату даних із вимогами трансформерної архітектури. Завдяки цьому донавчання моделі зводиться до адаптації останніх шарів під конкретну предметну область, тоді як мовні представлення залишаються успадкованими з великих корпусів загального призначення.

Пояснювана складова методу реалізована за допомогою бібліотеки Captum, яка спеціалізується на інтерпретації моделей PyTorch. У рамках експерименту використовується метод Integrated Gradients, що дозволяє обчислювати атрибуції важливості для окремих токенів тексту відносно виходу моделі для класу «релевантний хештег». Captum забезпечує чисельно стійку реалізацію інтегрованих градієнтів, у тому числі для вхідних ембедінгів трансформера, що уможливорює коректне виділення найважливіших слів підпису й подальше обчислення показника узгодженості із семантичним ядром хештегу. Таким чином, ця бібліотека виступає ключовим інструментом для реалізації пояснюваної частини запропонованого методу.

Для підготовки даних і обчислення класичних метрик класифікації використовується бібліотека scikit-learn [64]. Вона забезпечує розбиття вибірки пар «текст-хештег» на тренувальну та валідаційну підвибірки зі збереженням розподілу класів, а також обчислення показників accuracy та F1-score. Ці метрики використовуються для кількісної оцінки роботи моделі на різних етапах навчання та валідації, що дозволяє відстежувати динаміку якості та вибирати найкращі налаштування гіперпараметрів. Таким чином scikit-learn виступає інструментом статистичного контролю й аналізу класифікаційних властивостей моделі.

Опрацювання табличних даних і текстових полів датасету здійснюється засобами бібліотек pandas [65] і NumPy. Pandas використовується для читання CSV-файлу Instagram Reach Data, фільтрації записів без підписів і хештегів, формування структур даних для навчання й візуалізації перших рядків у вебінтерфейсі. NumPy забезпечує векторизовані операції над масивами чисел, зокрема нормування атрибуцій важливості, сортування значень та обчислення допоміжних статистичних характеристик. У сукупності ці бібліотеки формують стандартний аналітичний шар,

який поєднує роботу з даними та підготовку їх до подальшого подання нейромережевій моделі.

Користувацький доступ до експериментальної установки організовано за допомогою фреймворка Gradio, який дозволяє безпосередньо в середовищі Google Colab побудувати вебінтерфейс із кількома вкладками. У рамках експерименту Gradio використовується для реалізації вкладок «Дані», «Навчання», «Моделі» та «Оцінка релевантності», де користувач може завантажувати датасет, налаштовувати параметри навчання, працювати з реєстром моделей і запускати оцінювання релевантності та пояснюваний аналіз. Gradio забезпечує двосторонню інтеграцію з Python-кодом: дані з інтерфейсу передаються в бекенд-функції, а результати (таблиці, повідомлення, історія метрик) повертаються в інтерактивному вигляді, що робить експериментальну установку зручною для демонстрації й тестування.

Таким чином, експериментальна установка являє собою узгоджене поєднання хмарного середовища Google Colab, фреймворку глибокого навчання PyTorch, бібліотек transformers і Captum для роботи з мовними моделями та пояснюваними атрибутами, інструментів scikit-learn, pandas і NumPy для аналітики даних і метрик, а також вебфреймворку Gradio для організації інтерфейсної взаємодії. Такий набір засобів забезпечує повний цикл експерименту – від завантаження та попередньої обробки даних до навчання моделі, її оцінювання, інтерпретації результатів і візуалізації роботи інтелектуальної системи.

#### **4.5 Дослідження методу**

Основною частиною методу є нейронна мережа для оцінки відповідності хештегів коротким текстам. Нейромережевою архітектурою було обрано BERT, оскільки, оскільки ця модель забезпечує двобічний аналіз контексту й дає змогу однаково повно враховувати як попередні, так і наступні слова в реченні. Для задачі встановлення відповідності «підпис-хештег» така властивість є критичною: семантика тегу нерідко визначається не окремими ключовими словами, а їх поєднанням і оточенням у короткому тексті, тому однобічні або спрощені моделі

гірше фіксують тонкі змістові відтінки. Додатковою перевагою є попереднє навчання BERT на великих текстових корпусах, що дозволяє використовувати вже сформовані мовні представлення і донавчати модель на відносно невеликій кількості пар «текст-хештег» без втрати здатності узагальнювати.

Важливою причиною вибору є природна підтримка форматів «питання-відповідь» та «речення-речення», коли на вхід подається пара текстових фрагментів. У роботі це використовується для спільного кодування підпису та хештегу в одну послідовність із маркуванням сегментів, що дає можливість моделі явно бачити границю між текстом і тегом та будувати приховані представлення їх взаємодії, а не обробляти їх окремо. Модель навчалась протягом 3-х епох, графік динаміки значень метрики Ассурасу наведено на рисунку 4.10.

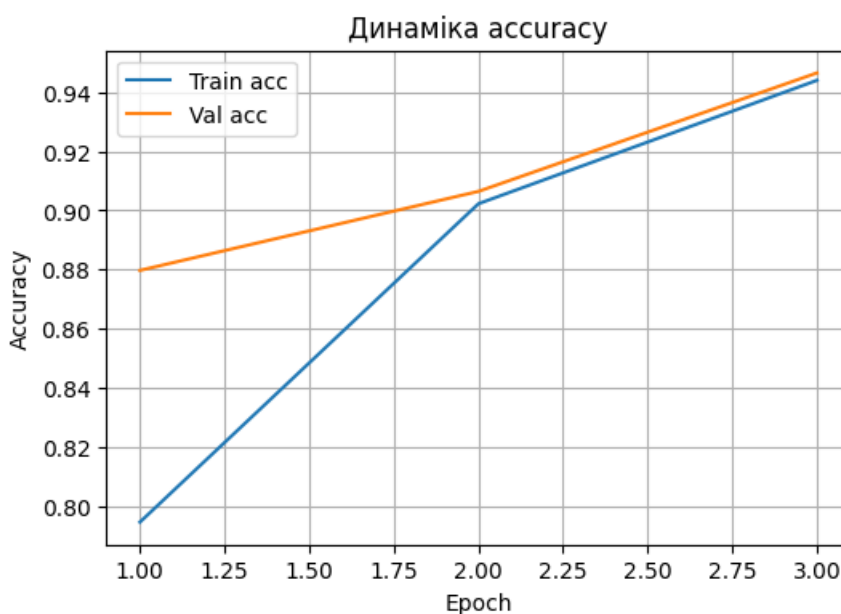


Рисунок 4.10 – Динаміка метрики Ассурасу

Наведений графік демонструє послідовне зростання точності як на тренувальній, так і на валідаційній вибірках протягом трьох епох навчання: з приблизно 0.79 до понад 0.94 для тренувальної й з 0.88 до близько 0.95 для валідаційної. Така динаміка показує, що модель поступово опановує структуру даних без різкого розходження кривих, характерного для перенавчання. Незначна перевага валідаційної точності над тренувальною може бути зумовлена відносно невеликим

обсягом вибірки та стохастичними ефектами розбиття даних, однак загалом крива валідації залишається гладкою та зростаючою, що вказує на стабільну здатність моделі до узагальнення й доцільність обраних гіперпараметрів у заданому діапазоні епох. Динаміка значень функції втрат наведена на рисунку 4.11.

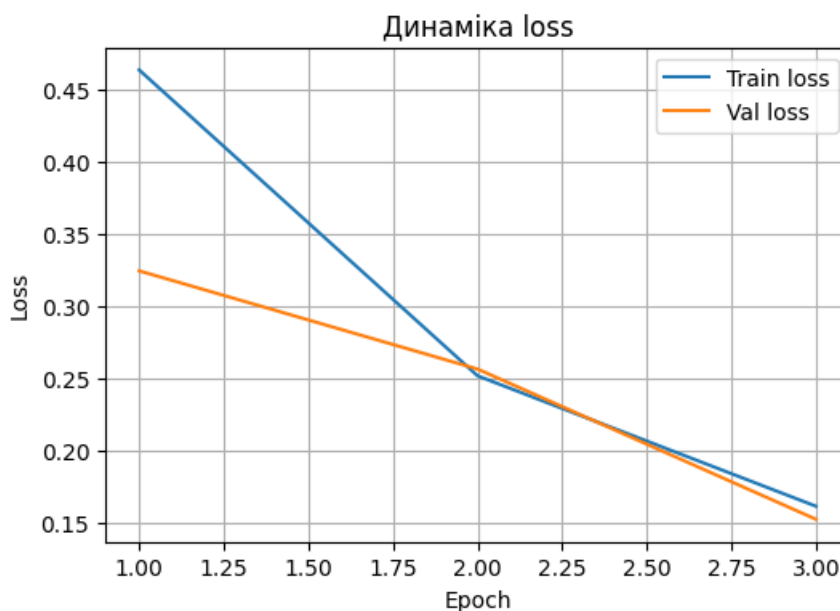


Рисунок 4.11 – Динаміка метрики Loss

Графік динаміки значення функції втрат показує монотонне зменшення як тренувального, так і валідаційного loss упродовж трьох епох: тренувальний показник спадає приблизно з 0.46 до 0.16, а валідаційний з близько 0.32 до 0.15. Уже на другій епосі криві майже збігаються, а надалі зберігається незначний зазор без перетину, що свідчить про стабільне покращення узгодженості моделі з даними й відсутність ознак перенавчання в розглянутому діапазоні епох. Паралельний спад обох кривих вказує на коректно підібраний темп навчання та достатню якість сформованих пар «текст-хештег», а досягнення низьких значень loss на валідаційній вибірці підтверджує здатність моделі узагальнювати знання поза межами тренувального підмножини. Динаміка значень метрики  $F_1$  наведена на рисунку 4.12.

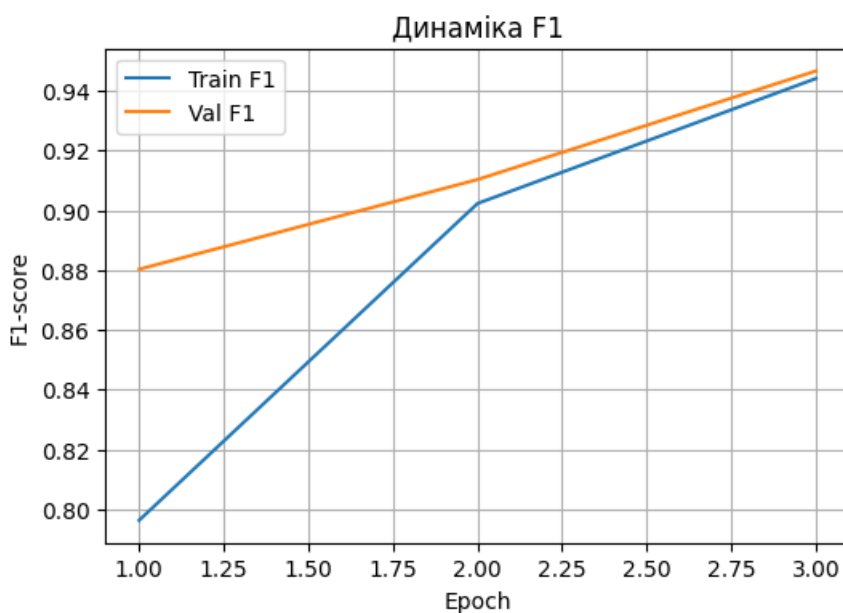


Рисунок 4.12 – Динаміка метрики  $F_1$

Графік динаміки  $F_1$ -міри наводить стає зростання як на тренувальній, так і на валідаційній вибірках: тренувальний  $F_1$  підвищується орієнтовно з 0.79 до 0.945, тоді як валідаційний із близько 0.88 до 0.95. Усі три епохи характеризуються тим, що крива валідаційного  $F_1$  розташована трохи вище за тренувальну, що свідчить про відсутність перенавчання та про те, що модель не лише добре відтворює патерни тренувальних даних, а й демонструє дещо кращий баланс між повнотою та точністю на валідаційній вибірці. Така поведінка узгоджується зі спаданням функції втрат і зростанням Ассигасу та підтверджує, що сформовані пари «текст-хештег» і обрана конфігурація BERT-класифікатора забезпечують високоякісне розпізнавання релевантних хештегів без суттєвої деградації узагальнювальної здатності.

Після завершення експерименту в блоці статусу (рисунок 4.13) відображається повідомлення «Модель збережено в models/run\_bert\_1», що свідчить про успішне збереження ваг мережі й супровідних файлів у вказаній директорії. Нижче наведено таблицю «Історія навчання», у якій для кожної епохи фіксуються значення функції втрат на тренувальній та валідаційній вибірках (стовпці `train_loss`, `val_loss`), а також показники `train_acc`, `val_acc`, `train_f1` та `val_f1`. Представлені числові результати показують поступове зменшення втрат та зростання як точності, так і  $F_1$ -міри від

першої до третьої епохи, що підтверджує коректну роботу інтелектуальної системи за обраною конфігурацією гіперпараметрів.

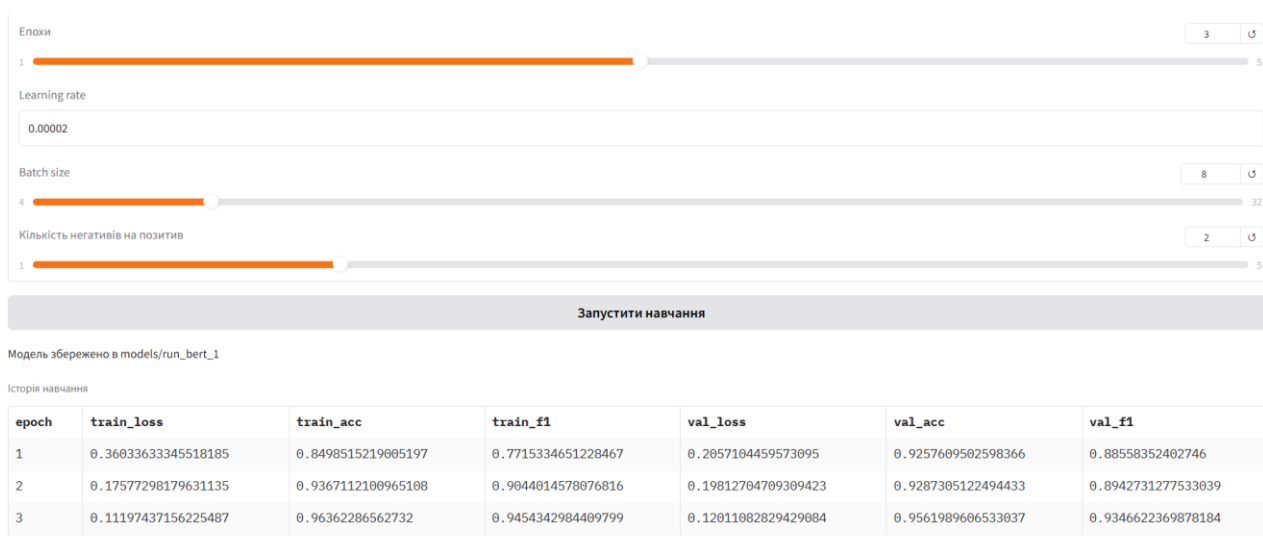


Рисунок 4.13 – Звіт з навчання нейромережі із вказаними гіперпараметрами

Наступною частиною дослідження методу є його емпірична перевірка. Користувачський текст «*Here are some of the most important data visualizations that every Financial Data Analyst/Scientist should know.*» був проаналізований на відповідність хештегам: #finance, #datavisualization, #cooking. Результат аналізу наведено на рисунку 4.14.

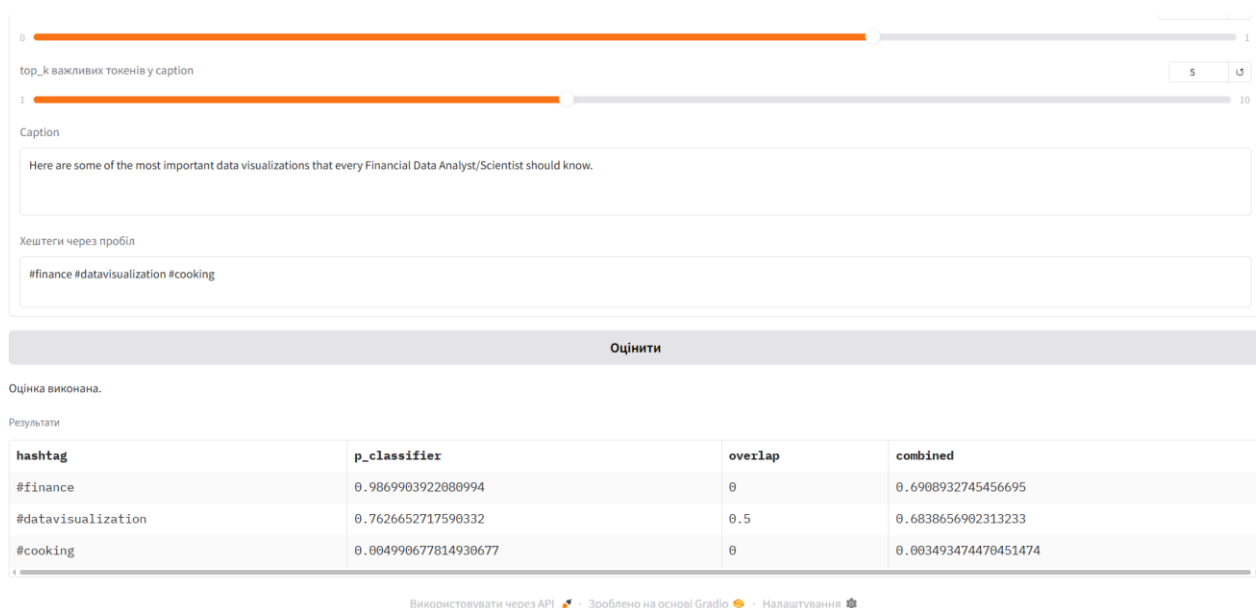


Рисунок 4.14 – Результат аналізу повідомлення на відповідність хештегам

Розроблений метод повертає для хештегів такі значення: для #finance ймовірність релевантності  $p\_classifier$  близько 0.99, показник перекриття  $overlap = 0$ , інтегральна оцінка  $combined$  близько 0.69; для #datavisualization  $p\_classifier$  близько 0.76,  $overlap = 0.5$ ,  $combined$  близько 0.68; для #cooking  $p\_classifier$  близько 0.005,  $overlap = 0$ ,  $combined$  близько 0.0035. За значенням інтегральної міри система ранжує хештеги в порядку #finance, #datavisualization, #cooking, що узгоджується зі змістом підпису: повідомлення стосується фінансової аналітики та візуалізації даних, а хештег #cooking є семантично чужорідним. При цьому для #finance модель показує майже максимальну нейромережеву впевненість, але нульове перекриття пояснення з ядром тегу (топові токени підпису не містять морфологічних збігів із словом finance), тоді як для #datavisualization імовірність нижча, зате половина найважливіших tokenів тексту лексично пов'язана з ядром тегу, що підвищує  $overlap$  і компенсує меншу  $p\_classifier$ . У підсумку інтегральний показник ( $combined$ ) відображає компроміс між прогностичною здатністю нейромережі та змістовою обґрунтованістю її рішення, а отримане ранжування демонструє коректну роботу як класифікаційної, так і пояснюваної складових методу.

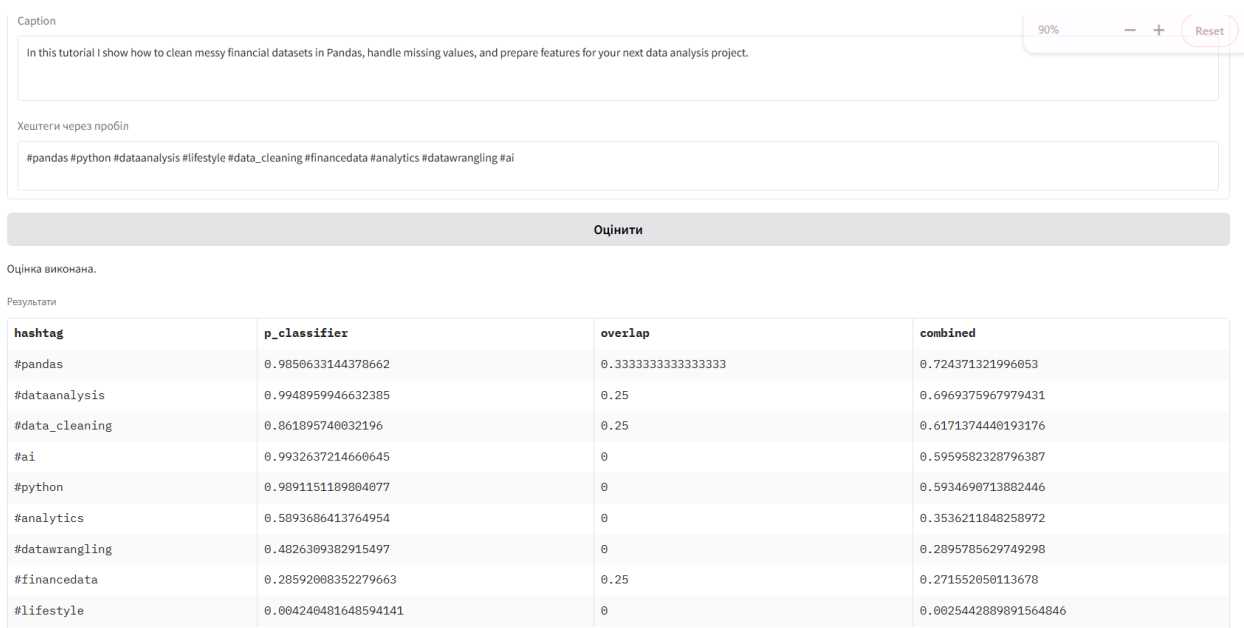
Наступним було проаналізовано повідомлення «*Today I built an interactive Python dashboard with Plotly to track real-time stock prices and visualize portfolio performance.*» щодо відповідності хештегам: #python #plotly #stockmarket #trading #portfolio #realtime #datascience #machinelearning #travel #fashion #foodblogger. Результати аналізу наведені на рисунку 4.15.

hashtag	p_classifier	overlap	combined
#plotly	0.9908876419067383	0.6666666666666666	0.8611992518107897
#portfolio	0.9844406843185425	0.5	0.7906644105911256
#realtime	0.8503953814506531	0.5	0.7102372288703918
#python	0.9940563440322876	0.25	0.6964338064193725
#trading	0.9589259028434753	0.25	0.6753555417060851
#stockmarket	0.9152764081954956	0.25	0.6491658449172973
#datascience	0.884280800819397	0	0.5305684804916382
#machinelearning	0.3919723331928253	0	0.23518339991569517
#travel	0.007838129065930843	0	0.004702877439558506
#foodblogger	0.004789276979863644	0	0.002873566187918186
#fashion	0.004343916196376085	0	0.002606349717825651

Рисунок 4.15 – Подання важливості хештегів

Таблиця, наведена на скріншоті, показує, що метод коректно виділяє найбільш релевантні теги до змісту повідомлення. Найвищі інтегральні оцінки `combined` мають `#plotly`, `#portfolio`, `#realtime`, `#python`, `#trading` та `#stockmarket`. Для них одночасно висока ймовірність класифікатора (`p_classifier` близько 0.95-0.99) і ненульовий показник `overlap` (0.25-0.67), тобто важливі токени підпису – `Plotly`, `portfolio`, `real-time`, `Python`, `trading`, `stock` – явно перетинаються з ядром відповідних хештегів. Теги `#datascience` і `#machinelearning` мають дуже високу нейромережеву оцінку, але `overlap` = 0, оскільки ці слова прямо не згадані в тексті; у результаті інтегральний бал нижчий, але все ще помітно вищий, ніж у відверто нерелевантних тегів. Для `#travel`, `#fashion` та `#foodblogger` модель дає дуже малі `p_classifier` і нульовий `overlap`, тому їхні інтегральні значення `combined` майже нульові та опиняються внизу рейтингу. Таким чином, пояснюваний інтегральний показник зміщує пріоритет до тих тегів, у яких і семантика, і важливі слова підпису узгоджені зі змістом повідомлення, та ефективно відсікає другорядні й випадкові хештеги.

Аналіз короткого тексту «*In this tutorial I show how to clean messy financial datasets in Pandas, handle missing values, and prepare features for your next data analysis project.*» на відповідність хештегам `#pandas` `#python` `#dataanalysis` `#lifestyle` `#data_cleaning` `#financedata` `#analytics` `#datawrangling` `#ai` наведено на рисунку 4.16.



Caption

In this tutorial I show how to clean messy financial datasets in Pandas, handle missing values, and prepare features for your next data analysis project.

Хештеги через пробіл

#pandas #python #dataanalysis #lifestyle #data\_cleaning #financedata #analytics #datawrangling #ai

Оцінити

Оцінка виконана.

Результати

hashtag	p_classifier	overlap	combined
#pandas	0.9850633144378662	0.3333333333333333	0.724371321996053
#dataanalysis	0.9948959946632385	0.25	0.6969375967979431
#data_cleaning	0.861895740032196	0.25	0.6171374440193176
#ai	0.9932637214660645	0	0.5959582328796387
#python	0.9891151189804077	0	0.5934690713882446
#analytics	0.5893686413764954	0	0.35362110482508972
#datawrangling	0.4826309382915497	0	0.2895785629749298
#financedata	0.28592008352279663	0.25	0.271552050113678
#lifestyle	0.004240481648594141	0	0.0025442889891564846

Рисунок 4.16 – Оцінка релевантних та нерелевантних хештегів

За результатами роботи методу найвищі інтегральні бали отримали хештеги #pandas, #dataanalysis та #data\_cleaning. Для них одночасно висока нейромережева оцінка p\_classifier (близько 0.86-0.99) і ненульовий показник overlap (0.25-0.33), тобто важливі токени на кшталт Pandas, data analysis, cleaning явно збігаються з ядром тегів, тому метод вважає їх найбільш релевантними до туторіалу про очищення фінансових датасетів. Теги #ai та #python мають дуже високі p\_classifier, але overlap = 0, оскільки у тексті немає прямих згадок цих слів, тож інтегральний бал нижчий: модель показує, що вони загалом доречні до технічного контенту, але пояснення не підтверджує явної лексичної відповідності. #analytics, #datawrangling і #financedata отримують середні значення p\_classifier без перекриття з важливими токенами, тому опиняються в середині рейтингу як помірно релевантні. #lifestyle має майже нульову ймовірність і combined, що очікувано для контенту про обробку даних.

Загалом метод коректно підсилює теги, де поєднуються тематика даних/очищення та явна згадка в тексті, і відсікає випадкові або надто загальні хештеги.

#### **Висновки до розділу 4**

У цьому розділі здійснено повний цикл експериментального дослідження запропонованого методу оцінювання відповідності хештегів коротким текстам та його програмної реалізації у вигляді інтелектуальної системи. На основі UML-діаграм сформовано цілісну структуру компонентів, що охоплює сутності даних, сервіси класифікації, пояснення та інтегрального скорингу, а також реєстр моделей і механізм керування версіями BERT-класифікатора. Схема розгортання в середовищі Google Colab показала, що обраний технологічний стек (PyTorch, transformers, Captum, scikit-learn, pandas, NumPy, Gradio) забезпечує взаємодію між вебінтерфейсом, Python-модулями та файловим сховищем і дозволяє виконувати експерименти в хмарному оточенні з використанням GPU-прискорення.

Подані алгоритми формування навчальних пар «текст-хештег», донавчання BERT-моделі та пояснюваного оцінювання хештегів продемонстрували коректну

роботу в реальних умовах. Експерименти з навчанням моделі засвідчили стабільне покращення якості: протягом трьох епох спостерігається зростання Accuracy та F1-міри на тренувальній і валідаційній вибірках до рівня близько 0.95 при одночасному монотонному спаданні функції втрат, що вказує на відсутність перенавчання та адекватність обраних гіперпараметрів. Збереження історії метрик і ваг моделі в реєстрі дозволяє відтворювати результати та порівнювати альтернативні конфігурації класифікатора.

Розроблений пояснюваний механізм, який поєднує нейромережеву ймовірність релевантності з показником перекриття атрибутів Integrated Gradients із семантичним ядром хештегу, підтвердив свою практичну ефективність на низці репрезентативних прикладів. У всіх розглянутих сценаріях інтегральний показник combined коректно ранжує хештеги, підсилюючи ті з них, що одночасно мають високий прогнозний бал і лексико-семантичну відповідність ключовим словам тексту, та знижуючи рейтинг випадкових або тематично далеких тегів. Це демонструє, що запропонований метод не лише досягає високої класифікаційної якості, а й надає інтерпретовані, змістовно обґрунтовані оцінки, придатні для використання аналітиками та контент-менеджерами.

Проведене функціональне тестування вебінтерфейсу й серверної логіки показало працездатність усіх ключових підсистем: навігації головного меню, завантаження та попередньої обробки датасету, навчання та збереження моделі, роботи реєстру моделей і сценарію оцінювання релевантності множини хештегів. У сукупності результати розділу 4 підтверджують, що запропонований метод і реалізована на його основі інтелектуальна система є функціонально завершеними, здатними до узагальнення на реальних даних і можуть слугувати базою для подальших досліджень та впровадження в практику роботи з хештегами в соціальних медіа.

## **Загальні висновки**

Метою кваліфікаційної роботи було підвищення достовірності та інтерпретованості оцінювання релевантності хештегів коротким текстовим повідомленням шляхом використання пояснюваного нейромережевого підходу, який поєднує трансформерну класифікацію з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

Для досягнення мети було поставлено та виконано наступні задачі:

- виконано аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах, окреслено типові помилки класифікації й обмеження традиційних евристичних підходів;

- здійснено огляд і порівняльний аналіз сучасних нейромережевих та лінгвістичних підходів до класифікації коротких текстів, рекомендації тегів і пояснюваного аналізу рішень моделей обробки природної мови, визначено їх сильні та слабкі сторони в контексті задачі роботи;

- проаналізовано наявні набори даних соціальних мереж, обґрунтовано вибір конкретного датасету для формування пар «текст – хештег – мітка» та сформульовано систему показників якості оцінювання релевантності, придатну для подальших експериментів;

- розроблено нейромережевий метод пояснюваного оцінювання релевантності хештегів, у якому трансформерна класифікація поєднана з метриками узгодженості важливих фрагментів тексту зі змістовим ядром хештегу, а також визначено схему інтеграції цих метрик в єдину інтегральну оцінку;

- виконано підготовку та анотування вибірки коротких текстів і хештегів на основі даних соціальних мереж із формуванням позитивних і негативних прикладів для донавчання моделі, включаючи очищення, нормалізацію та побудову збалансованих підвибірок;

- здійснено програмну реалізацію запропонованого методу з використанням сучасних засобів глибинного навчання та пояснюваного штучного інтелекту, зокрема трансформерних мовних моделей і методів атрибуцій важливості;

- проведено експериментальні дослідження ефективності розробленого методу порівняно з базовими підходами, що ґрунтувалися лише на класифікаційних оцінках або векторних поданнях, та проаналізовано отримані числові результати;

- виконано поглиблений аналіз результатів, оцінено вплив пояснюваних метрик на якість відсікання нерелевантних хештегів і виявлення семантично

коректних, але рідкісних тегів, а також сформульовано практичні рекомендації щодо застосування методу в системах підтримки роботи з хештегами та контент-менеджменту соціальних медіа.

У роботі було здійснено подальший розвиток методів оцінювання відповідності хештегів коротким текстам за рахунок поєднання нейромережевої класифікації з показниками пояснюваності. В результаті навчання нейромережевої архітектури BERT було досягнуто точності 0.956 за метрикою Ассигасу. Запропонований підхід передбачав, що релевантність хештегу визначається не лише за ймовірністю віднесення пари «текст-хештег» до класу «відповідає» / «не відповідає», отриманою трансформерною мовною моделлю, але й за мірою узгодженості внутрішніх пояснень моделі із семантичним ядром самого хештегу. На цій основі було введено інтегральний показник релевантності, який синтезує класифікаційну оцінку та частку найбільш інформативних для моделі фрагментів тексту, лексично чи словотвірно пов'язаних із тегом, що відрізняє запропоноване рішення від традиційних підходів, орієнтованих лише на векторні відстані або бінарне рішення класифікатора.

Додатковим результатом стала методика автоматизованого формування корпусу пар «підпис-хештег-мітка» на основі реальних дописів соціальної мережі з використанням негативної вибірки та подальший аналіз режимів роботи моделі за поєднанням високої/низької впевненості класифікатора з високою/низькою узгодженістю пояснень. Показано, що введена пояснювана метрика дала змогу не лише надійніше відсікати нерелевантні хештеги, а й виявляти семантично коректні, але рідкісні або відсутні в навчальній вибірці теги, які базовий класифікатор схилився б відхиляти. У підсумку було сформовано новий підхід до оцінювання якості хештегів, у якому вона трактується через прозорість та інтерпретованість рішень нейромережевої моделі поряд із її прогностичною точністю.

Основні наукові й практичні результати роботи доповідались у доповіді «Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP» на XVII Всеукраїнській науково-практичній конференції «Актуальні проблеми комп'ютерних наук АПКН-2025» (м.Хмельницький) 14-15 листопада 2025 року [66].

## Перелік посилань

1. Kumar A., R. R. K. A Comprehensive Review of Image Denoising Techniques: From Traditional Filters to Deep Learning. *International Journal For Multidisciplinary Research*. 2025. Vol. 7, no. 4. URL: <https://doi.org/10.36948/ijfmr.2025.v07i04.46673> (date of access: 04.12.2025).
2. Що таке хештег і як його використовувати? – Підвищення цифрової компетентності. *Підвищення цифрової компетентності – інструменти для онлайн-навчання*. URL: <https://cikt.kubg.edu.ua/що-таке-хештег-і-як-його-використовува/> (дата звернення: 04.12.2025).
3. Haried P., Wu S., Han Y. The Hidden Impact of Hashtags on Instagram: Navigational Heuristics on Source Trustworthiness. *International Journal of Web Based Communities*. 2024. Vol. 1, no. 1. URL: <https://doi.org/10.1504/ijwbc.2024.10062541> (date of access: 04.12.2025).
4. Навіщо потрібні хештеги і яку користь приносять. *Агенція комплексного діджитал-маркетингу Content.UA*. URL: <https://content.ua/ukr/blog/zachem-nuzhny-heshtegi> (дата звернення: 04.12.2025).
5. Lin B., Lee W., Choe Y. Social media engagement of hashtag users in the context of local events: mixed method approach. *Journal of Hospitality and Tourism Technology*. 2023. URL: <https://doi.org/10.1108/jhtt-03-2023-0074> (date of access: 04.12.2025).
6. Що таке хештег та репост? Як саме вони працюють в соцмережах ? - Press Association. *Press Association*. URL: <https://pressassociation.org.ua/ua/shho-take-heshteg-ta-repost-yak-same-voni-praczuuyut-v-soczmerzah/> (дата звернення: 04.12.2025).
7. Що таке хештег і як їх використовувати їх в SMM. *Promodo*. URL: <https://www.promodo.ua/blog/chi-potribno-staviti-heshtegi-tak-rozpovidaiemo-pro-strategiyu-vikoristannya-heshtegiv-u-2024-roci> (date of access: 04.12.2025).
8. Хештеги Instagram: як правильно їх використовувати. *Блог Digital агенції UAMASTER*. URL: <https://blog.uamaster.com/heshtegy-instagram-yak-pravylnovykorystovuvaty-heshtegy> (дата звернення: 04.12.2025).

9. Bauer M. J., Jestaedt D. Hashtags im Markenmanagement. *Claims, Slogans und Hashtags als Instrumente der strategischen Markenführung*. Wiesbaden, 2024. P. 65–78. URL: [https://doi.org/10.1007/978-3-658-44879-0\\_4](https://doi.org/10.1007/978-3-658-44879-0_4) (date of access: 04.12.2025).
10. Lin B., Lee W., Choe Y. Social media engagement of hashtag users in the context of local events: mixed method approach. *Journal of Hospitality and Tourism Technology*. 2023. URL: <https://doi.org/10.1108/jhtt-03-2023-0074> (date of access: 04.12.2025).
11. Hashtag Re-Appropriation for Audience Control on Recommendation-Driven Social Media Xiaohongshu (rednote) / R. Wan et al. *CHI 2025: CHI Conference on Human Factors in Computing Systems*, Yokohama Japan. New York, NY, USA, 2025. P. 1–25. URL: <https://doi.org/10.1145/3706598.3713379> (date of access: 04.12.2025).
12. Gerber A. A content analysis: analyzing topics of conversation under the #sustainability hashtag on Twitter. *Environmental Data Science*. 2024. Vol. 3. URL: <https://doi.org/10.1017/eds.2024.4> (date of access: 04.12.2025).
13. Gupta D., Chakraverty S. HaRNA<sub>T</sub> - A dynamic hashtag recommendation system using news. *Online Social Networks and Media*. 2025. Vol. 45. P. 100294. URL: <https://doi.org/10.1016/j.osnem.2024.100294> (date of access: 04.12.2025).
14. Dynamic Hashtag Recommendation in Social Media With Trend Shift Detection and Adaptation / R. Cantini et al. *IEEE Transactions on Computational Social Systems*. 2025. P. 1–13. URL: <https://doi.org/10.1109/tcss.2025.3621724> (date of access: 04.12.2025).
15. Hashtag recommendation for enhancing the popularity of social media posts / P. Chakrabarti et al. *Social Network Analysis and Mining*. 2023. Vol. 13, no. 1. URL: <https://doi.org/10.1007/s13278-023-01024-9> (date of access: 04.12.2025).
16. La Rocca G., Boccia Artieri G. Interpreting the changeable meaning of hashtags: Toward the theorization of a model. *Frontiers in Sociology*. 2023. Vol. 7. URL: <https://doi.org/10.3389/fsoc.2022.1104686> (date of access: 04.12.2025).
17. Prakash V J., Vijay S A. A. A Comprehensive Multimodal Framework for Optimizing Social Media Hashtag Recommendations. *IEEE Transactions on*

- Computational Social Systems*. 2024. P. 1–12.  
URL: <https://doi.org/10.1109/tcss.2024.3508733> (date of access: 04.12.2025).
18. Unlocking Deeper Data Insights on Social Media: Removing Hashtag and Tweets Spam for Improved Content Analysis / G. V. Nikhil Sai et al. *2024 5th International Conference for Emerging Technology (INCET)*, Belgaum, India, 24–26 May 2024. 2024. URL: <https://doi.org/10.1109/incet61516.2024.10593042> (date of access: 04.12.2025).
19. Bansal S., Gowda K., Kumar N. Multilingual personalised hashtag recommendation for low resource Indic languages using graph-based deep neural network. *Expert Systems with Applications*. 2023. P. 121188. URL: <https://doi.org/10.1016/j.eswa.2023.121188> (date of access: 04.12.2025).
20. Emotion classification for short texts: an improved multi-label method / X. Liu et al. *Humanities and Social Sciences Communications*. 2023. Vol. 10, no. 1. URL: <https://doi.org/10.1057/s41599-023-01816-6> (date of access: 04.12.2025).
21. A Survey on Text Classification: From Traditional to Deep Learning / Q. Li et al. *ACM Transactions on Intelligent Systems and Technology*. 2022. Vol. 13, no. 2. P. 1–41. URL: <https://doi.org/10.1145/3495162> (date of access: 04.12.2025).
22. Deep Learning--based Text Classification / S. Minaee et al. *ACM Computing Surveys*. 2021. Vol. 54, no. 3. P. 1–40. URL: <https://doi.org/10.1145/3439726> (date of access: 04.12.2025).
23. A Method of Short Text Representation Based on the Feature Probability Embedded Vector / Zhou et al. *Sensors*. 2019. Vol. 19, no. 17. P. 3728. URL: <https://doi.org/10.3390/s19173728> (date of access: 04.12.2025).
24. Recipe Recommendation System Using TF-IDF / S. Chhipa et al. *ITM Web of Conferences*. 2022. Vol. 44. P. 02006. URL: <https://doi.org/10.1051/itmconf/20224402006> (date of access: 04.12.2025).
25. GeeksforGeeks. What is Text Classification? - *GeeksforGeeks*. *GeeksforGeeks*. URL: <https://www.geeksforgeeks.org/nlp/what-is-text-classification/> (date of access: 04.12.2025).

26. L'Homme M.-C. The processing of terms in dictionaries: New models and techniques. *Terminology*. 2006. Vol. 12, no. 2. P. 181–188. URL: <https://doi.org/10.1075/term.12.2.02hom> (date of access: 04.12.2025).
27. Rakshit P., Sarkar A. A supervised deep learning-based sentiment analysis by the implementation of Word2Vec and GloVe Embedding techniques. *Multimedia Tools and Applications*. 2024. URL: <https://doi.org/10.1007/s11042-024-19045-7> (date of access: 04.12.2025).
28. Bag of Tricks for Efficient Text Classification / A. Joulin et al. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, Valencia, Spain. Stroudsburg, PA, USA, 2017. URL: <https://doi.org/10.18653/v1/e17-2068> (date of access: 04.12.2025).
29. BERT-based combination of convolutional and recurrent neural network for Indonesian sentiment analysis / H. Murfi et al. *Applied Soft Computing*. 2024. Vol. 151. P. 111112. URL: <https://doi.org/10.1016/j.asoc.2023.111112> (date of access: 04.12.2025).
30. KIM, Yoon. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014. URL: <https://arxiv.org/abs/1408.5882> (date of access: 04.12.2025).
31. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin et al. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, Minnesota. Stroudsburg, PA, USA, 2019. URL: <https://doi.org/10.18653/v1/n19-1423> (date of access: 04.12.2025).
32. Li C., Xie Z., Wang H. Short Text Classification Based on Enhanced Word Embedding and Hybrid Neural Networks. *Applied Sciences*. 2025. Vol. 15, no. 9. P. 5102. URL: <https://doi.org/10.3390/app15095102> (date of access: 04.12.2025).
33. Zhao W. Deep Active Learning for Short-Text Classification : thesis. 2017. URL: <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-212577> (date of access: 04.12.2025).
34. Alsini A., Huynh D. Q., Datta A. Hashtag Recommendation Methods for Twitter and Sina Weibo: A Review. *Future Internet*. 2021. Vol. 13, no. 5. P. 129. URL: <https://doi.org/10.3390/fi13050129> (date of access: 04.12.2025).

35. Jain K., Jindal R. NLP-enabled Recommendation of Hashtags for Covid based Tweets using Hybrid BERT-LSTM Model. *ACM Transactions on Asian and Low-Resource Language Information Processing*. 2024. URL: <https://doi.org/10.1145/3640812> (date of access: 04.12.2025).
36. Połap D. Hybrid image analysis model for hashtag recommendation through the use of deep learning methods. *Expert Systems with Applications*. 2023. P. 120566. URL: <https://doi.org/10.1016/j.eswa.2023.120566> (date of access: 04.12.2025).
37. Cross-modality representation learning from transformer for hashtag prediction / M. M. Y. Khalil et al. *Journal of Big Data*. 2023. Vol. 10, no. 1. URL: <https://doi.org/10.1186/s40537-023-00824-2> (date of access: 04.12.2025).
38. Dadgar S., Neshat M. A Novel Hybrid Multi-Modal Deep Learning for Detecting Hashtag Incongruity on Social Media. *Sensors*. 2022. Vol. 22, no. 24. P. 9870. URL: <https://doi.org/10.3390/s22249870> (date of access: 04.12.2025).
39. Giannoulakis S., Tsapatsoulis N., Djouvas C. Evaluating the use of Instagram images color histograms and hashtags sets for automatic image annotation. *Frontiers in Big Data*. 2023. Vol. 6. URL: <https://doi.org/10.3389/fdata.2023.1149523> (date of access: 04.12.2025).
40. Bansal S., Gowda K., Kumar N. Multilingual personalised hashtag recommendation for low resource Indic languages using graph-based deep neural network. *Expert Systems with Applications*. 2023. P. 121188. URL: <https://doi.org/10.1016/j.eswa.2023.121188> (date of access: 04.12.2025).
41. Alsini A., Huynh D. Q., Datta A. #REval: A Semantic Evaluation Framework for Hashtag Recommendation. *IEEE Transactions on Knowledge and Data Engineering*. 2025. P. 1–13. URL: <https://doi.org/10.1109/tkde.2025.3553683> (date of access: 04.12.2025).
42. BERT applications in natural language processing: a review / N. M. Gardazi et al. *Artificial Intelligence Review*. 2025. Vol. 58, no. 6. URL: <https://doi.org/10.1007/s10462-025-11162-5> (date of access: 04.12.2025).
43. Reddy G. P., Kumar Y. V. P. Explainable AI (XAI): Explained. *2023 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*, Vilnius,

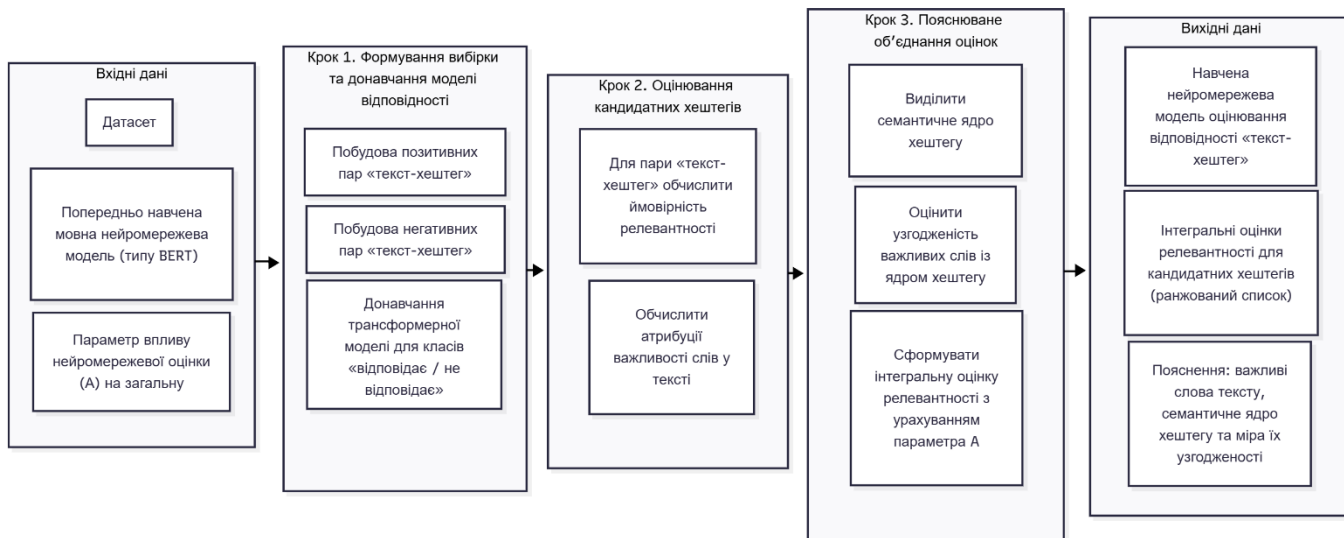
- Lithuania, 27 April 2023. 2023.  
 URL: <https://doi.org/10.1109/esteam59056.2023.10134984> (date of access: 04.12.2025).
44. How to explain AI systems to end users: a systematic literature review and research agenda / S. Laato et al. *Internet Research*. 2022. Vol. 32, no. 7. P. 1–31.  
 URL: <https://doi.org/10.1108/intr-08-2021-0600> (date of access: 04.12.2025).
45. Wegmann A., Nguyen D., Jurgens D. Tokenization is Sensitive to Language Variation. *Findings of the Association for Computational Linguistics: ACL 2025*, Vienna, Austria. Stroudsburg, PA, USA, 2025. P. 10958–10983.  
 URL: <https://doi.org/10.18653/v1/2025.findings-acl.572> (date of access: 04.12.2025).
46. What we mean when we say semantic: Toward a multidisciplinary semantic glossary / J. Reilly et al. *Psychonomic Bulletin & Review*. 2024.  
 URL: <https://doi.org/10.3758/s13423-024-02556-7> (date of access: 04.12.2025).
47. Boselli R., D’Amico S., Nobani N. eXplainable AI for Word Embeddings: A Survey. *Cognitive Computation*. 2024. Vol. 17, no. 1.  
 URL: <https://doi.org/10.1007/s12559-024-10373-2> (date of access: 04.12.2025).
48. Instagram Reach Data. *Kaggle*.  
 URL: <https://www.kaggle.com/datasets/amankumarkharwal/instagram-reach-data> (date of access: 04.12.2025).
49. Kaggle. *Kaggle: Your Machine Learning and Data Science Community*.  
 URL: <http://www.kaggle.com> (date of access: 04.12.2025).
50. Welcome to Python.org. *Python.org*. URL: <https://www.python.org/> (date of access: 04.12.2025).
51. colab.google. *colab.google*. URL: <https://colab.google/> (date of access: 04.12.2025).
52. How to use Google Colab - GeeksforGeeks. *GeeksforGeeks*.  
 URL: <https://www.geeksforgeeks.org/websites-apps/how-to-use-google-colab/> (date of access: 04.12.2025).
53. W3Schools.com. *W3Schools Online Web Tutorials*.  
 URL: [https://www.w3schools.com/python/python\\_intro.asp](https://www.w3schools.com/python/python_intro.asp) (date of access: 04.12.2025).
54. PyTorch. *PyTorch*. URL: <https://pytorch.org/> (date of access: 04.12.2025).

55. Gradio. *Gradio*. URL: <https://gradio.app/> (date of access: 04.12.2025).
56. Introduction to Gradio - Hugging Face LLM Course. *Hugging Face – The AI community building the future*. URL: <https://huggingface.co/learn/llm-course/chapter9/1> (date of access: 04.12.2025).
57. Jedličková A. Ethical approaches in designing autonomous and intelligent systems: a comprehensive survey towards responsible development. *AI & SOCIETY*. 2024. URL: <https://doi.org/10.1007/s00146-024-02040-9> (date of access: 04.12.2025).
58. Intelligent System Design Based on Soft Computing Models / ed. by O. Castillo, P. Melin. Cham : Springer Nature Switzerland, 2025. URL: <https://doi.org/10.1007/978-3-031-97309-3> (date of access: 04.12.2025).
59. Шевченко К. Довіряй, але перевіряй. Як тестувати ML-моделі. *High Bar Journal*. URL: <https://journal.gen.tech/post/yak-testuvaty-ml-modeli> (дата звернення: 04.12.2025).
60. Довідник по Machine Learning – Machine Learning Model Evaluation | База знань IT. *База знань IT технологій*. URL: <https://itwiki.dev/data-science/ml-reference/ml-glossary/machine-learning-model-evaluation> (дата звернення: 04.12.2025).
61. A Review of Large Language Model Evaluation Metrics in Education / N. Paoprasert et al. *Communications in Computer and Information Science*. Singapore, 2025. P. 402–415. URL: [https://doi.org/10.1007/978-981-95-4499-8\\_33](https://doi.org/10.1007/978-981-95-4499-8_33) (date of access: 04.12.2025).
62. NumPy. *NumPy*. URL: <https://numpy.org/> (date of access: 04.12.2025).
63. PyTorch. *PyTorch*. URL: <https://pytorch.org/> (date of access: 04.12.2025).
64. scikit-learn: machine learning in Python – scikit-learn 0.16.1 documentation. *scikit-learn: machine learning in Python – scikit-learn 0.16.1 documentation*. URL: <http://scikit-learn.org/> (date of access: 04.12.2025).
65. pandas – Python Data Analysis Library. *pandas - Python Data Analysis Library*. URL: <https://pandas.pydata.org/> (date of access: 04.12.2025).
66. Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP / Трохимчук О.В., Пасічник О.А., Поплавська О.А , Міхалевський В.Ц. // *Актуальні проблеми комп'ютерних наук : зб. наук. пр. за матеріалами XVII Всеукр.*

*наук.-практ. конф. (АПКН-2025).* – Хмельницький, 14–15 листоп. 2025 р. –  
Хмельницький, 2025. – С. 409–412.

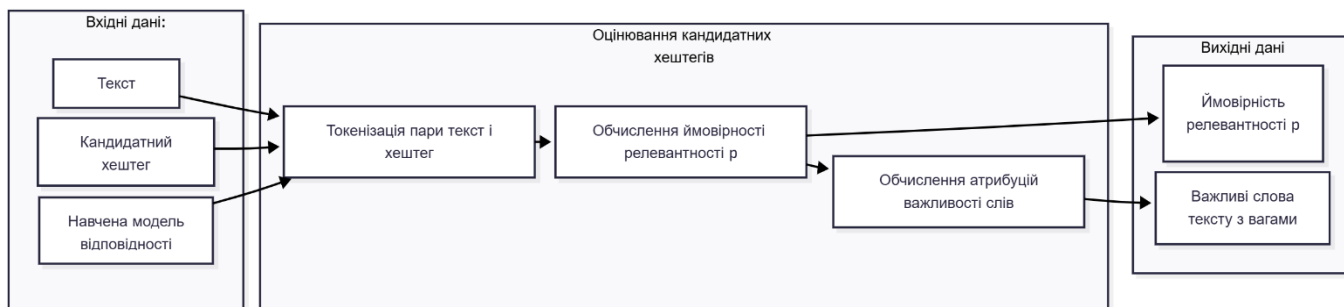
# ДОДАТКИ

## Додаток А

Схема та етапи підходу до оцінювання відповідності хештегів коротким  
текстам засобами NLP

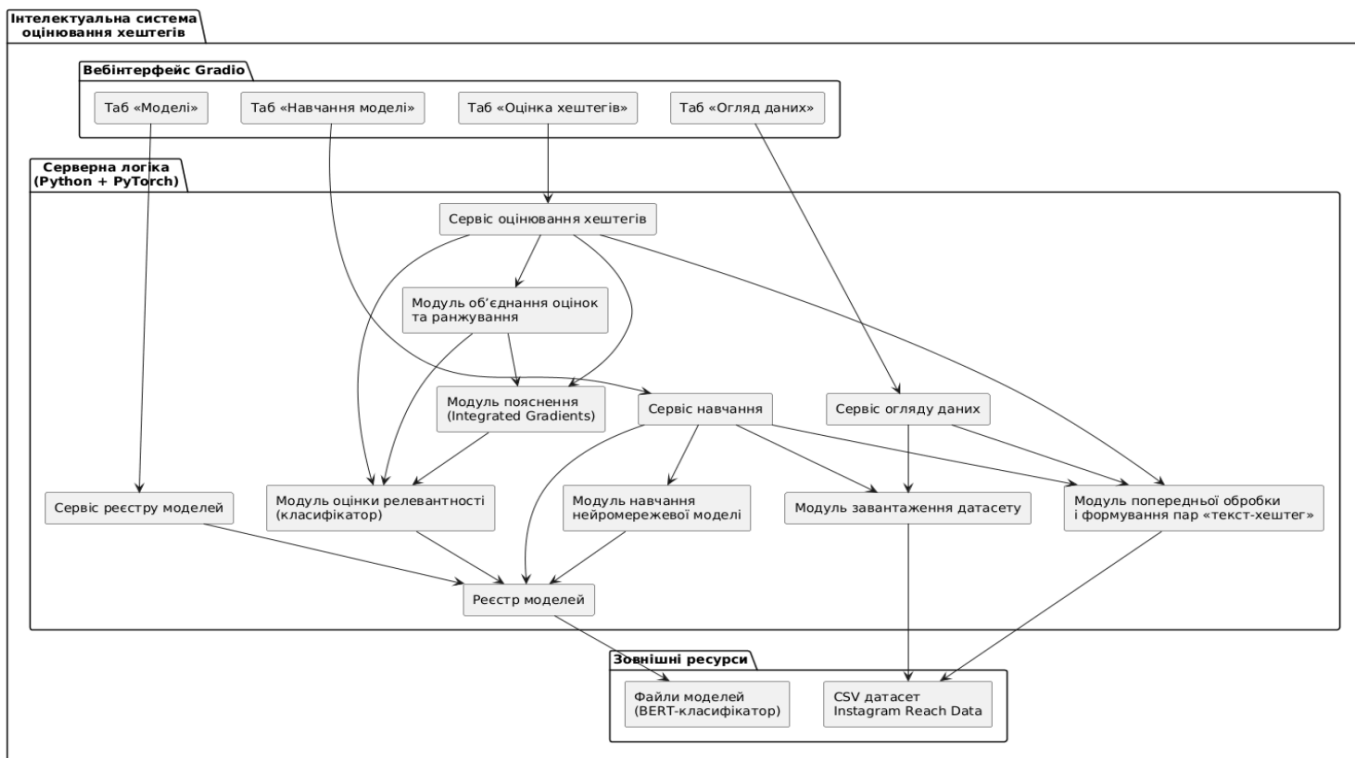
## Додаток Б

### Схема процесу оцінювання кандидатних хештегів



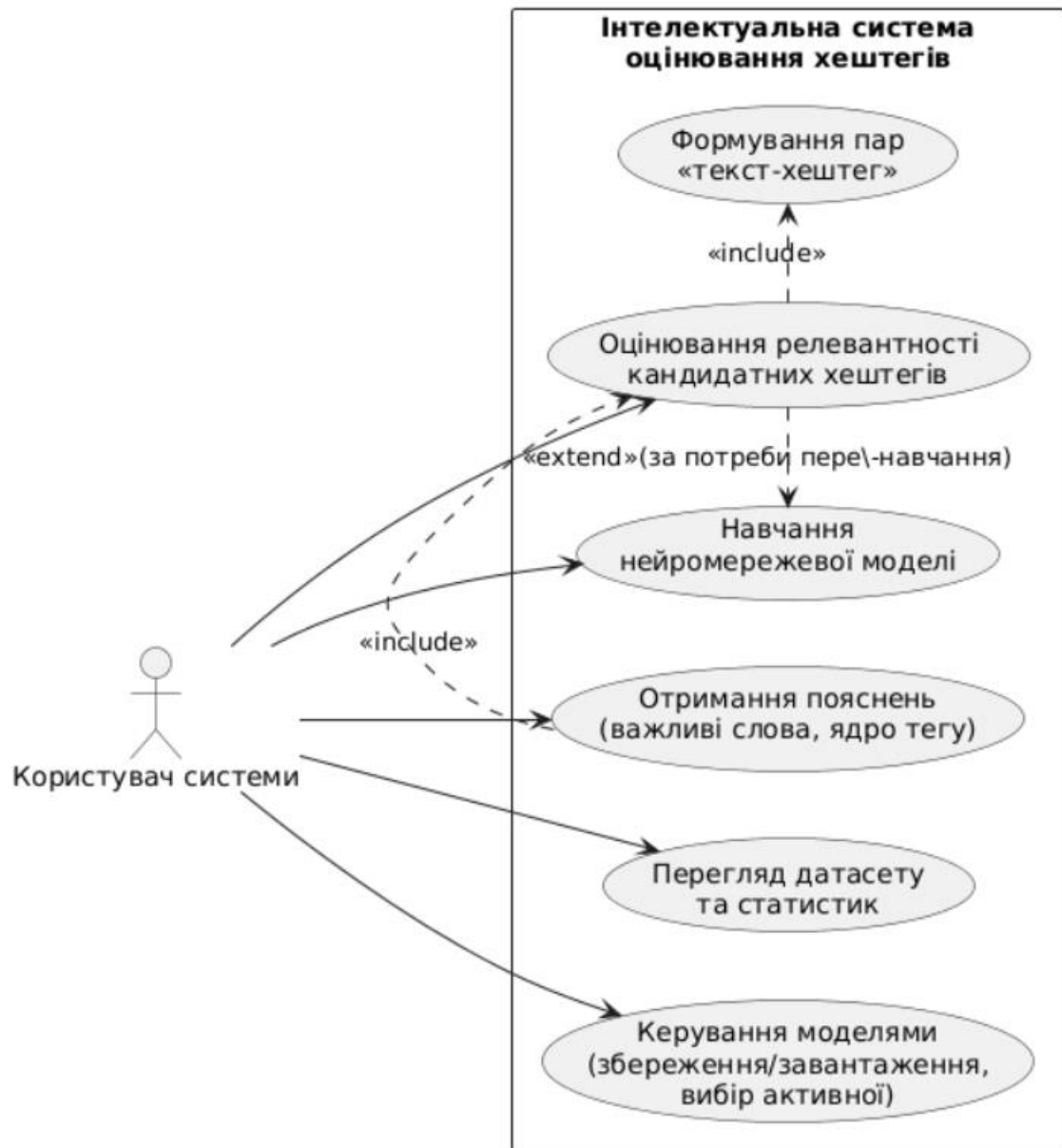
## Додаток В

### Схема взаємодії складових інтелектуальної системи ідсистем інтелектуальної системи автоматизованого визначення архітектурних стилів



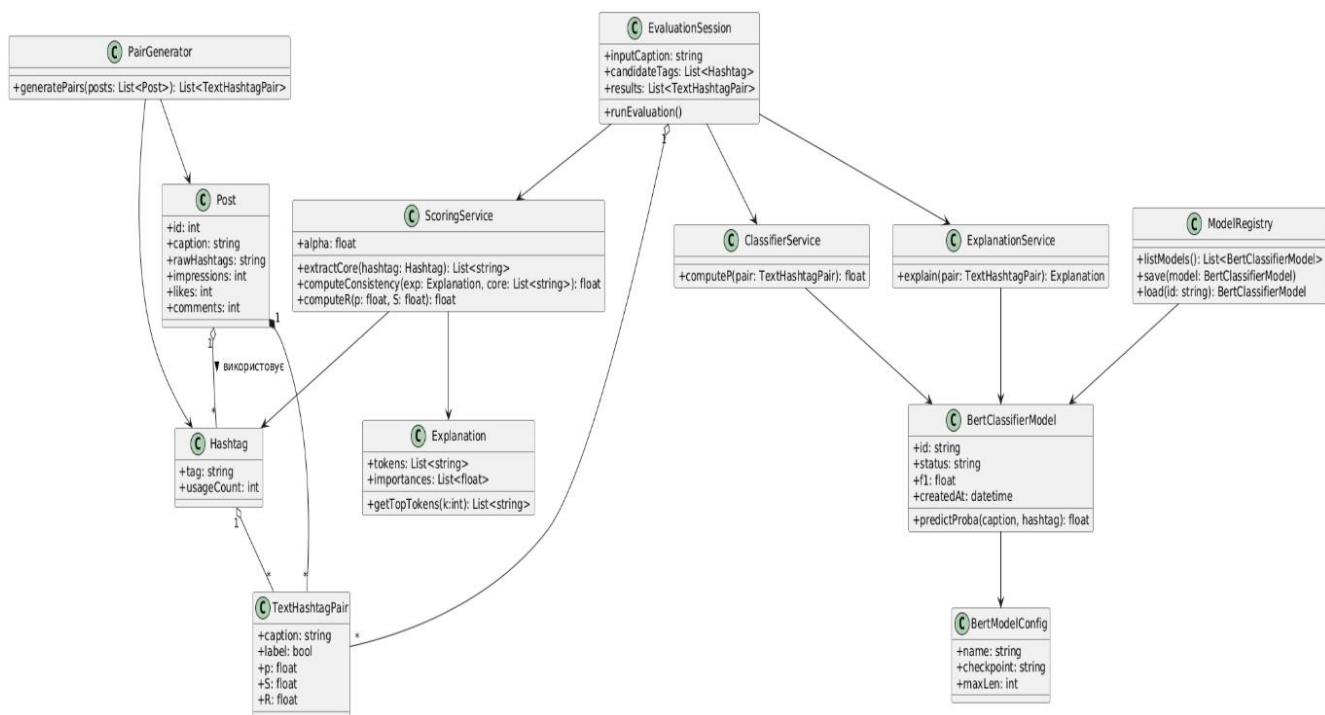
## Додаток Г

## Діаграма варіантів використання інтелектуальної системи



## Додаток Д

## Діаграма класів інтелектуальної системи



## Додаток Е

### Світлини екрану інтелектуальної системи

Лабораторія оцінки релевантності хештегів (explanation-based)

[Дані](#) [Навчання](#) [Моделі](#) [Оцінка релевантності](#)

Шлях до датасету

Instagram data.csv

Завантажити дані

Перші 20 рядків

1	2	3

Індекс рядка для перегляду

0

Caption

Hashtags

Лабораторія оцінки релевантності хештегів (explanation-based)

[Дані](#) [Навчання](#) [Моделі](#) [Оцінка релевантності](#)

Налаштування та навчання моделі

Назва запуску (ім'я моделі)

run\_bert\_1

Базова модель HF

bert-base-uncased

Епохи

1 5

Learning rate

0.00002

Batch size

4 8 32

Кількість негативів на позитив

1 5

Запустити навчання

### Лабораторія оцінки релевантності хештегів (explanation-based)

Дані   Навчання   **Моделі**   Оцінка релевантності

**Наявні моделі**

Показати список моделей у 'models/'

Список моделей

Шлях до моделі для завантаження

models/run\_bert\_1

Завантажити модель

### Лабораторія оцінки релевантності хештегів (explanation-based)

Дані   Навчання   Моделі   **Оцінка релевантності**

**Оцінка релевантності хештегів (explanation-based matching)**

$\alpha$  — вага ймовірності класифікатора 0.7 ↕

0 ————— 1

top\_k важливих токенів у caption 5 ↕

1 ————— 10

Caption

Here are some of the most important data visualizations that every Financial Data Analyst/Scientist should know.

Хештеги через пробіл

#finance #datavisualization #cooking

Оцінити

Результати

1	2	3
---	---	---

Caption

Here are some of the most important data visualizations that every Financial Data Analyst/Scientist should know.

Хештеги через пробіл

#finance #datavisualization #cooking

Оцінити

Спочатку завантажте модель (вкладка 'Моделі').

Результати

1	2	3
---	---	---

Індекс рядка для перегляду

25

Рядок #25


Caption

Deep learning is a subset of machine learning in which we use artificial neural networks to train models to solve complex problems where we need a large amount of data compared to other machine learning algorithms. Here are some of the best deep learning project ideas that you must try.

Hashtags

#deeplearning #machinelearning #artificialintelligence #datascience #ai #python #coding #technology #programming #bigdata #dataanalytics #datascientist #data #computerscience #tech #neuralnetwork #amankharwal #thecleverprogrammer

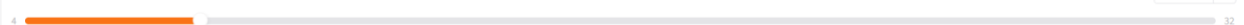
Епоки 3

1  5


Learning rate

0.00002

Batch size 8

4  32

Кількість негативів на позитив 2

1  5

**Запустити навчання**

Модель збережено в models/run\_bert\_1

Історія навчання

epoch	train_loss	train_acc	train_f1	val_loss	val_acc	val_f1
1	0.36033633345518185	0.8498515219005197	0.7715334651228467	0.2057104459573095	0.9257609502598366	0.88558352402746
2	0.17577298179631135	0.9367112100965108	0.9044014578076816	0.19812704709309423	0.9287305122494433	0.8942731277533039
3	0.11197437156225487	0.96362286562732	0.9454342984409799	0.12011082829429084	0.9561989606533037	0.9346622369878184

## Додаток Ж

### Програмні коди

Вихідний код інтелектуальної системи доступний у репозиторії GitHub:  
<https://github.com/trokhymchukolha/hashtag-relevance-assessment> (дата звернення:  
10.12.2025).

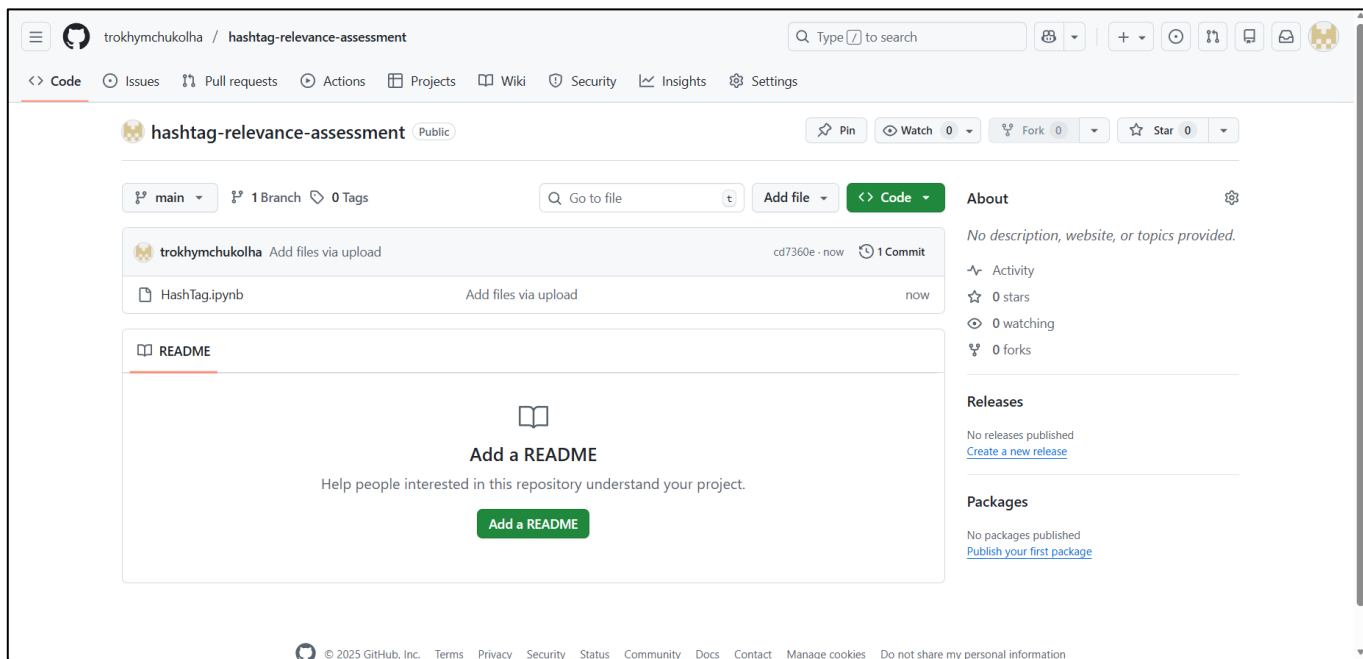


Рисунок Д.1 – Сторінка репозиторію

## Додаток К

### Світлини наукових публікацій, виконаних при роботі над кваліфікаційною роботою магістра

1. Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP/ Трохимчук О.В., Пасічник О.А., Поплавська О.А., Міхалевський В.Ц. // *Актуальні проблеми комп'ютерних наук : зб. наук. пр. за матеріалами XVII Всеукр. наук.-практ. конф. (АПКН-2025).* – Хмельницький, 14–15 листоп. 2025 р. – Хмельницький, 2025. – С. 409–412.

Міністерство освіти і науки України  
Хмельницький національний університет



ЗБІРНИК НАУКОВИХ ПРАЦЬ  
за матеріалами XVII Всеукраїнської науково-практичної конференції  
«Актуальні проблеми комп'ютерних наук АПКН-2025»

*14-15 листопада 2025*

Хмельницький 2025

<b>Сиротенко Д.А., Троц В.В., Анікін В.А.</b> Ризики використання штучного інтелекту з перспективи кібербезпеки та захисту інформації.....	377
<b>Скрипнюк О.Ю., Манзюк Е.А., Базрій Р.О., Петровський С.С.</b> Метод виявлення трасувальних зв'язків між вимогами та програмним кодом із використанням великих мовних моделей.....	380
<b>Соколовський В.С., Манзюк Е.А.</b> Метод класифікації патологій листя рослин на основі згорткових нейронних мереж.....	385
<b>Старостенко К.В.</b> Аналіз ефективності методів машинного навчання для виявлення мережових атак типу DDoS.....	390
<b>Стецюк П.П., Форкун Ю.В.</b> Метод удосконаленого модульного проєктування агентно-орієнтованих програмних систем з підтримкою розширюваності та повторного використання.....	393
<b>Тимофієв І.А., Мазурець О.В.</b> Нейромережовий підхід до виявлення депресивних патернів за аналізом текстового контенту цифрових сервісів у закладах освіти.....	395
<b>Тросянецький Н.О., Кльоц Ю.П., Калій К.В., Откидач В.В.</b> Виявлення атак типу «блокування IP через NAT» в публічних мережах.....	405
<b>Трохимчук О.В., Пасічник О.А., Поплавська О.А., Міхалевський В.Ц.</b> Підхід до оцінювання відповідності хештегів коротким текстам засобами NLP.....	409
<b>Філюк Є.В., Джулій В.М.</b> Метод безпеки криптоактивів на основі технології багатосторонніх обчислень.....	413
<b>Футорний Р.В., Медведчук Н.К.</b> Дослідження виявлення кібератак в Агро-ІОТ з використанням аналізу енергоспоживання.....	417
<b>Ціцьвіра І.О., Радюк П.М., Скрипник Т. К.</b> Метод агентно-орієнтованого аналізу ринку криптовалют з використанням великих мовних моделей.....	421
<b>Червончук І.С.</b> Аналіз методів та засобів виявлення логів у програмному забезпеченні.....	425

УДК 004.8

Трохимчук О.В., Пасічник О.А., Поплавська О.А., Міхалевський В.Ц.

*Хмельницький національний університет***ПІДХІД ДО ОЦІНЮВАННЯ ВІДПОВІДНОСТІ ХЕШТЕГІВ КОРОТКИМ  
ТЕКСТАМ ЗАСОБАМИ NLP**

*Розглянуто задачу автоматизованого оцінювання відповідності хештегів коротким текстам соціальних мереж в умовах змішаних мов, неформальних конструкцій та маніпулятивного маркування контенту. Запропоновано підхід, що базується на контекстних мовних моделях типу трансформерів і побудові спільного семантичного простору для текстів і хештегів, попередньо сегментованих і перетворених на фрази природної мови. Показано можливість інтеграції підходу в модулі модерації, рекомендаційні сервіси й аналітичні системи для виявлення нерелевантних і маніпулятивних тегів та підвищення якості тегування.*

*The paper addresses the task of automated assessment of hashtag relevance to short social media texts under conditions of mixed languages, informal style and manipulative tagging. The proposed approach relies on transformer-based contextual language models and a joint semantic space for texts and hashtags, which are first segmented and converted into natural-language phrases. The approach can be integrated into moderation modules, recommender systems and analytical platforms to detect irrelevant or manipulative hashtags and improve the overall quality of tagging.*

Поширення соціальних мереж та мікроблогів сформувало окрему «мову» коротких повідомлень, у якій хештеги виконують роль компактних маркерів тематики, настрою, події та цільової аудиторії [1]. Від коректності їх використання залежать якість пошуку, релевантність рекомендацій, прозорість аналітики та можливості автоматизованого моніторингу інформаційних потоків. Натомість практика свідчить про масове зловживання хештегами: від випадкового добору «модних» тегів до цілеспрямованого маніпулятивного маркування контенту. Це актуалізує задачу автоматизованого оцінювання відповідності хештегів коротким текстам, яка природно розв'язується засобами обробки природної мови.

Проблема є нетривіальною вже на рівні формалізації. Короткі тексти соціальних мереж містять розмовні конструкції, емодзі, скорочення, змішані мови, орфографічні відхилення, а хештеги поєднують словосполучення, сленг та елементи брендваної лексики [2]. Семантика хештегу не завжди збігається із буквальним змістом його компонентів: теги можуть позначати кампанію, мем, локальний

інформаційний контекст. Тому прості евристичні алгоритми підрахунку перетину слів чи збігу окремих ключових слів дають обмежені результати. Потрібна модель, яка працює в спільному семантичному просторі для тексту й хештегу, враховує контекст уживання та здатна відрізнити тематичну релевантність від поверхневої схожості.

Мета роботи полягає у розробленні підходу до оцінювання відповідності хештегів коротким текстам на основі сучасних моделей NLP із використанням контекстних подань, що дає змогу формувати кількісні показники релевантності, придатні для інтеграції в системи модерації, рекомендаційні сервіси та аналітичні панелі. Для досягнення мети пропонується узгоджена послідовність етапів: формування корпусу, лінгвістичний препроцесинг, побудова семантичних подань текстів і хештегів, навчання моделей оцінювання відповідності та валідація результатів на експертно розмічених даних.

На рівні даних базовою сутністю є пара «короткий текст - множина хештегів». Для дослідження формується корпус постів із соціальних платформ, де хештеги використовуються як основний механізм індексування. Для частини записів здійснюється експертне маркування, коли кожен хештег віднесено до одного з рівнів: релевантний змісту, частково релевантний (дуже широкий або вторинний щодо основної теми) та нерелевантний або маніпулятивний. Така градація важлива для навчання моделей, які мають відрізнити цілком доречні теги від тегів, що лише формально не суперечать тексту, але не відображають його змістового ядра. Додатково фіксуються метадані, зокрема час публікації, домінуюча мова, тип акаунта, що дозволяє у подальшому враховувати платформні та часові зсуви.

Лінгвістичний препроцесинг коротких текстів передбачає очищення від надлишкової технічної інформації, нормалізацію скорочень та приведення токенів до лематизованої форми із збереженням емотивних маркерів і ключових емодзі, які впливають на інтерпретацію. Особливу увагу приділено обробці змішаних текстів із поєднанням української, англійської та транслітераційних форм. Хештеги проходять окрему фазу сегментації: розбиття композитних тегів на складники, усунення декоративних символів, перетворення на фразу природної мови.

Ключовим компонентом підходу є побудова спільного семантичного простору для текстів і хештегів. Для цього застосовуються контекстні мовні моделі типу трансформерів, які навчені на багатомовних або спеціалізованих соціально-медійних корпусах. Короткий текст пропускається через модель для отримання векторного подання, що агрегує інформацію на рівні висловлювання, а хештег, попередньо перетворений на фразу, обробляється тим самим механізмом. У

результаті і текст, і хештег репрезентовано у спільному латентному просторі, в якому семантично узгоджені об'єкти розташовані близько один до одного.

Оцінювання відповідності реалізується двома взаємодоповнювальними способами. Перший базується на безпосередньому вимірюванні подібності між векторами тексту та хештегу за допомогою косинусної метрики та подальшій калібровці порогових значень на валідаційній вибірці. Цей підхід є обчислювально економічним і підходить для масової первинної фільтрації. Другий спосіб розглядає пару «текст-хештег» як єдину послідовність і подає її на вхід моделі типу cross-encoder, яка навчається класифікувати ступінь відповідності на основі повного контексту, включно з лексичними, синтаксичними та прагматичними ознаками. Така модель краще захоплює складні випадки, зокрема іронію, метафоричне вживання та тематичні зсуви.

Навчання моделей здійснюється у напівконтрольованій постановці. Експертно розмічена частина корпусу використовується для супервізованого навчання класифікатора ступеня релевантності, тоді як невеликі вбудовані у корпус блоки даних без явних анотацій залучаються через контрастивні схеми: правильним пари «текст-хештег» протиставляються штучно згенерованим негативним прикладам, утвореним шляхом перестановки тегів між текстами схожої тематики. Це підсилює здатність моделі розрізняти не лише довільні нерелевантні поєднання, а й тонкі відмінності між близькими за тематикою, але семантично відмінними тегами.

Для оцінювання якості запропонованого підходу використовуються стандартні метрики класифікації, а також ранжувальні показники. На рівні окремих пар аналізуються точність, повнота та  $F_1$ -міра для кожного рівня релевантності, а також зважені усереднення за класами. На рівні множини хештегів для одного тексту розглядаються метрики, що характеризують порядок, у якому система пропонує або підтверджує теги, зокрема середня позиція релевантних хештегів та нормалізована кумулятивна оцінка приросту. Порівняння з базовими підходами, що використовують лише статистичні співзвуччя слів або поверхневі подібності, демонструє переваги контекстних моделей у розв'язанні задачі, особливо для коротких, емоційно насичених і семантично неоднозначних текстів.

Практичне значення розробленого підходу полягає у його здатності працювати як модуль у складі ширших інформаційних систем. У середовищі модератції контенту він дає змогу автоматично виявляти маніпулятивні чи «шумові» хештеги, які не відбивають зміст поста, але використовуються для привернення уваги або паразитування на трендових темах. У рекомендаційних сервісах оцінка відповідності може застосовуватись для очищення вхідних сигналів, що формують профіль користувача, зменшуючи вплив некоректних тегів на моделі персоналізації.

У системах аналітики підхід забезпечує побудову більш надійних карт тематичних кластерів, де хештеги виступають не довільними маркерами, а семантично верифікованими індикаторами дискусійних полів.

Окремим напрямом розвитку є поєднання оцінки відповідності з виявленням специфічних типів ризикового контенту. Нерелевантні або надто загальні хештеги часто супроводжують маніпулятивні повідомлення, приховану рекламу чи спроби обійти прості фільтри модерації. Інтеграція семантичних оцінок релевантності з іншими модулями аналізу, наприклад класифікаторами токсичності, дезінформації чи цифрової втоми, відкриває можливості для побудови багатовимірних профілів постів, де невідповідність хештегів є одним із сигналів ризику.

Підсумовуючи, запропонований підхід до оцінювання відповідності хештегів коротким текстам базується на використанні контекстних мовних моделей, спільного семантичного простору для текстів і тегів, а також поєднання подібнісних та класифікаційних процедур. Він враховує лінгвістичні особливості коротких, неформальних повідомлень, підтримує багатомовність, а також допускає гнучке налаштування порогів прийняття рішень залежно від завдань системи-споживача. Отримані результати свідчать про перспективність такого підходу для побудови інструментів контролю якості тегування, підвищення прозорості інформаційних потоків та зменшення впливу маніпулятивних практик у цифрових середовищах.

#### **Перелік посилань**

1. DialogueGCN: A Graph Convolutional Neural Network for Emotion Recognition in Conversation / D. Ghosal et al. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China. Stroudsburg, PA, USA, 2019. URL: <https://doi.org/10.18653/v1/d19-1015>
2. Wang L., Zhang L. Hawkes processes for understanding heterogeneity in information propagation on Twitter. *Frontiers in Physics*. 2022. Vol. 10. URL: <https://doi.org/10.3389/fphy.2022.1019380>

## Додаток Л

### Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

# МЕТОД ОЦІНЮВАННЯ ВІДПОВІДНОСТІ ХЕШТЕГІВ КОРОТКИМ ТЕКСТАМ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ



**Виконала:**  
*студентка групи КНм-24-1*  
**Ольга ТРОХИМЧУК**



**Керівник:**  
*к.т.н., доцент кафедри КН*  
**Олександр ПАСІЧНИК**

## Актуальність

Актуальність теми зумовлена різким зростанням обсягів коротких повідомлень у соціальних медіа, де хештеги виконують функції індексації, навігації та тематичного структурування контенту. Коректність їх добору визначає видимість публікацій, достовірність аналітики залученості та якість досліджень інформаційних потоків, тоді як нерелевантні або маніпулятивні теги спотворюють статистику та ускладнюють пошук. В умовах інформаційного перенасичення ручний контроль за хештегами є неможливим, що формує потребу в автоматизованих методах їх оцінювання.

Попри успіхи глибоких моделей у класифікації текстів, більшість підходів орієнтована на підвищення точності без належної прозорості та інтерпретованості. Оскільки моделі часто спираються на кореляції у вибірці або подібність векторних подань без гарантії семантичної обґрунтованості, зростає ризик формування хибних чи надмірних хештегів, що не відображають зміст повідомлення. Це знижує довіру до нейромережових інструментів аналізу соціальних медіа та обмежує їх використання у науковій аналітиці, маркетингу й системах модерації контенту.

## Мета і задачі роботи

**Об'єкт дослідження** – процес автоматизованого оцінювання релевантності хештегів для коротких текстових повідомлень у цифрових комунікаційних середовищах.

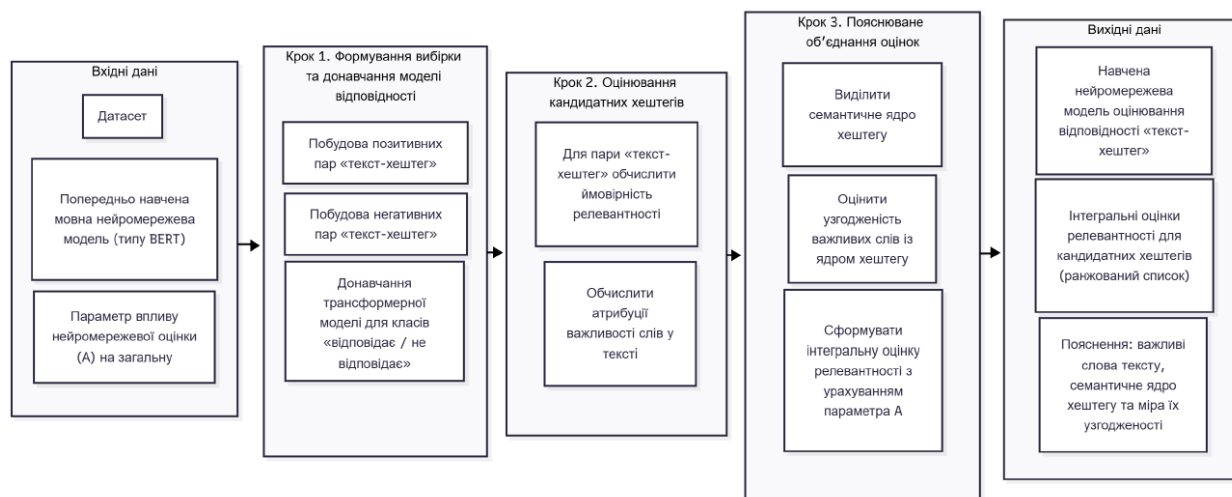
**Предмет дослідження** – Нейромережеві моделі та методи пояснюваного оцінювання релевантності хештегів коротким текстом, що поєднує класифікаційні моделі обробки природної мови з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

**Метою кваліфікаційної роботи магістра є підвищення достовірності та інтерпретованості оцінювання релевантності хештегів коротким текстовим повідомленням шляхом використання пояснюваного нейромережевого підходу, який поєднує трансформерну класифікацію з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.**

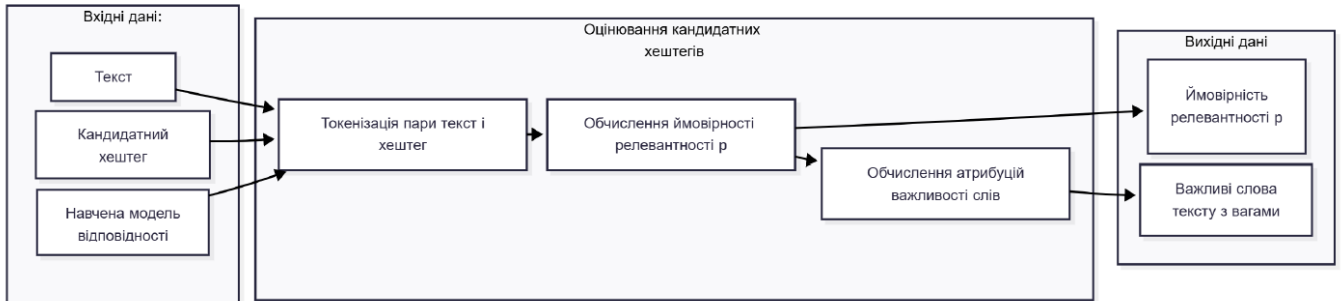
Для досягнення поставленої мети слід вирішити такі **завдання**:

- виконати аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах;
- здійснити аналіз сучасних нейромережевих та лінгвістичних підходів до класифікації коротких текстів, рекомендації тегів і пояснюваного аналізу рішень моделей обробки природної мови;
- проаналізувати наявні набори даних соціальних мереж, обґрунтувати вибір датасету для формування пар «текст – хештег – мітка» та визначити показники якості оцінювання релевантності;
- розробити нейромережевий метод пояснюваного оцінювання релевантності хештегів, що поєднує трансформерну класифікацію з метриками узгодженості важливих фрагментів тексту зі змістовим ядром хештегу;
- виконати підготовку та анотування вибірки коротких текстів і хештегів на основі даних соціальних мереж з формуванням позитивних і негативних прикладів для донавчання моделі;
- здійснити програмну реалізацію запропонованого методу з використанням сучасних засобів глибокого навчання та пояснюваного штучного інтелекту;
- провести експериментальні дослідження ефективності запропонованого методу порівняно з базовими підходами, що ґрунтуються лише на класифікаційних оцінках або векторних поданнях;
- виконати аналіз отриманих результатів, оцінити вплив пояснюваних метрик на якість відсікання нерелевантних хештегів і виявлення семантично коректних, але рідкісних тегів, та сформулювати практичні рекомендації щодо застосування методу.

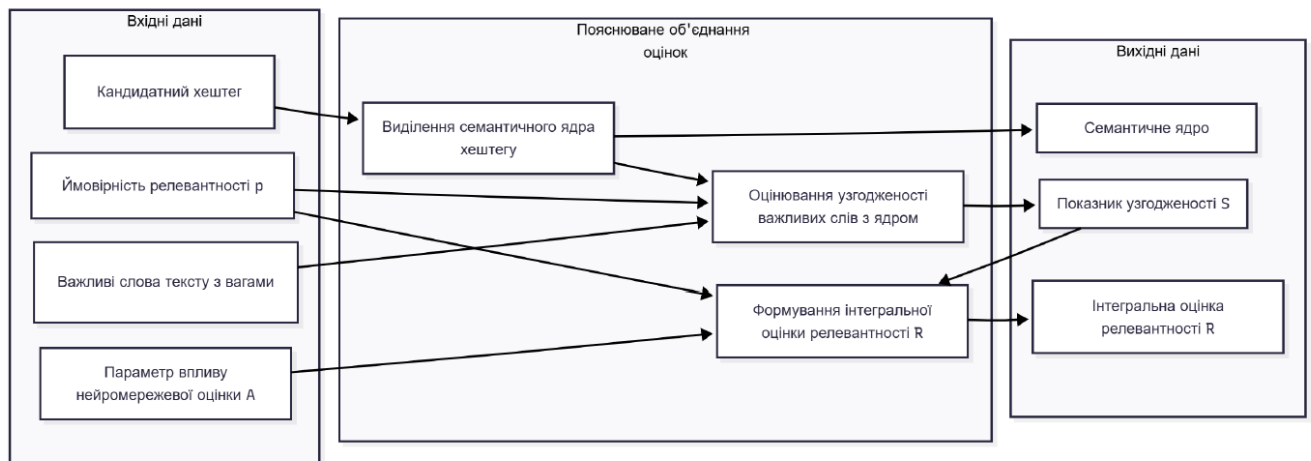
## Схема та етапи підходу до оцінювання відповідності хештегів коротким текстам засобами NLP



## Схема процесу оцінювання кандидатних хештегів



## Схема способу пояснюваного об'єднання оцінок та формування інтегральної міри релевантності



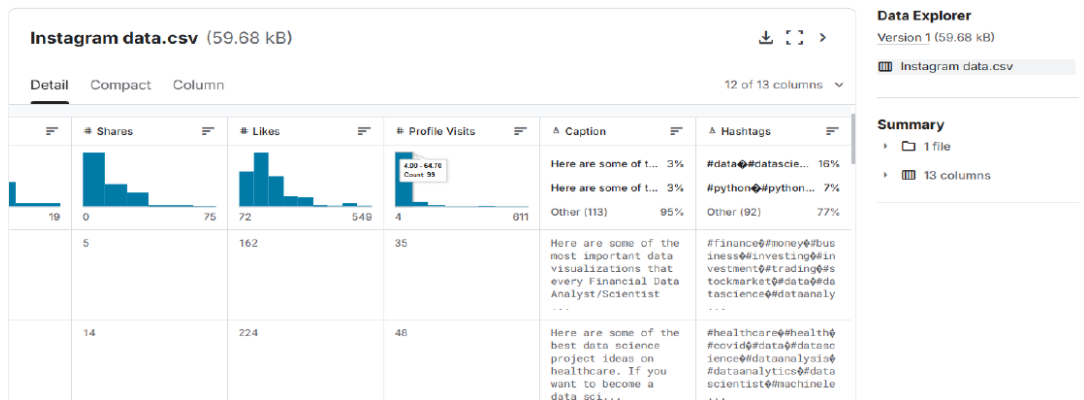
# Датасет

## Instagram Reach Data

6 <> Code Download

Data Card Code (2) Discussion (0) Suggestions (0)

13. Hashtags: Hashtags used in the post



Приклад записів датасету «Instagram Reach Data»

## Проектування структури інтелектуальної системи

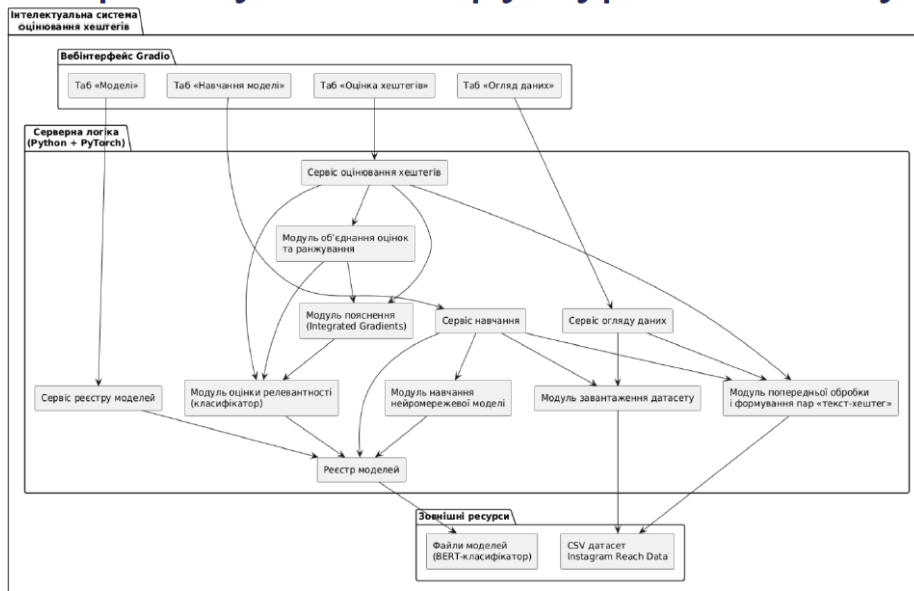
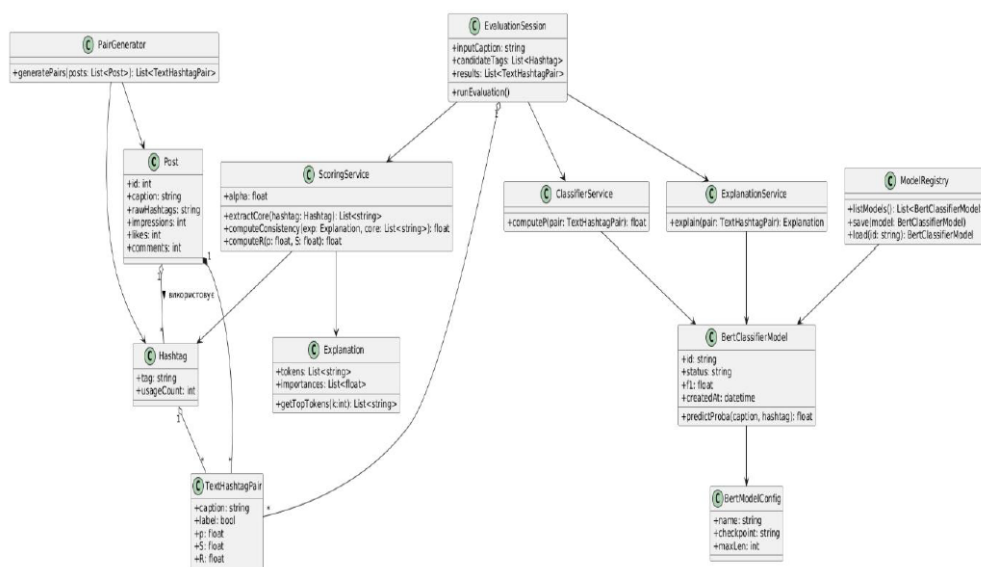


Схема взаємодії складових інтелектуальної системи

## Структура компонентів інтелектуальної системи оцінювання відповідності хештегів коротким текстам



Діаграма класів інтелектуальної системи

## Структура компонентів інтелектуальної системи оцінювання відповідності хештегів коротким текстам

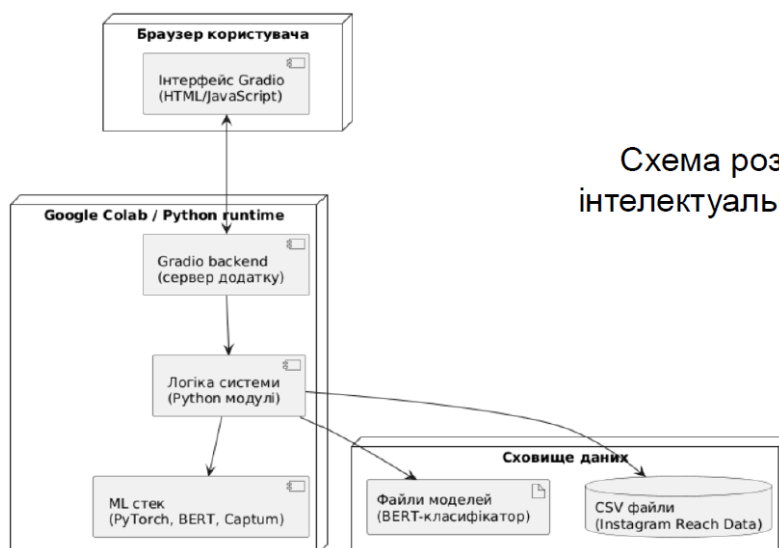


Схема розгортання інтелектуальної системи

## Інтелектуальна система

Лабораторія оцінки релевантності хештегів (explanation-based)

Дані Рівняння Відбиті Оцінка релевантності

Шлях до даних:

Вихідні дані:

Завантажити дані

Точка 20 років

1	2	3

Індекс рідності для перекладу:

Caption:

Назва:

Лабораторія оцінки релевантності хештегів (explanation-based)

Дані Рівняння Відбиті Оцінка релевантності

Оцінка релевантності хештегів (explanation-based matching)

ε - рівень важливості ключових слів:

top\_k важливих термінів у caption:

Caption

Here are some of the most important data visualizations that every Financial Data Analyst/Scientist should know

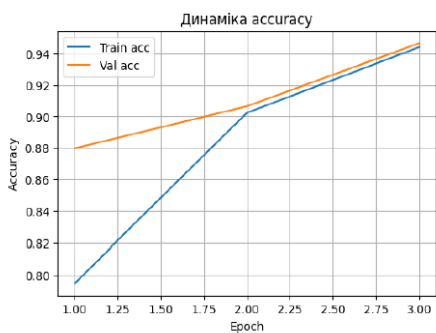
Ключові терміни зображення:

Finance visualization booking

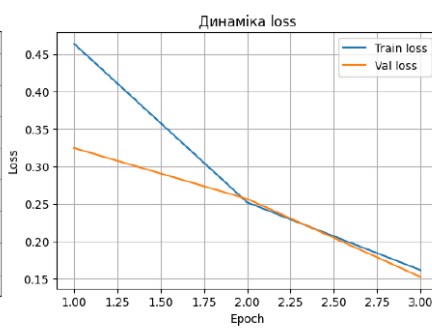
Оцінки

1	2	3

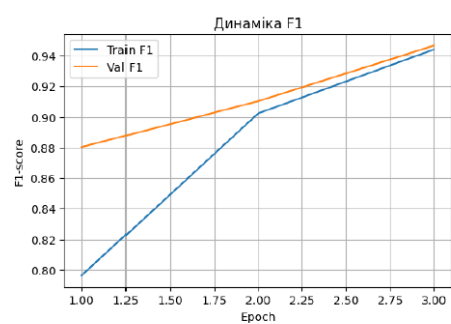
## Дослідження методу



Динаміка метрики Accuracy

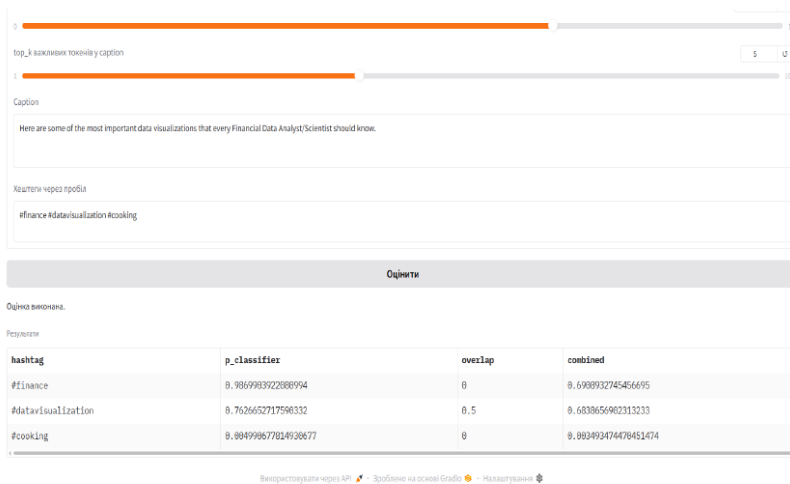


Динаміка метрики Loss



Динаміка метрики  $F_1$

## Дослідження методу



«Here are some of the most important data visualizations that every Financial Data Analyst/Scientist should know»

Розроблений метод повертає для хештегів такі значення: для #finance ймовірність релевантності p\_classifier близько 0.99, показник перекриття overlap = 0, інтегральна оцінка combined близько 0.69; для #datavisualization p\_classifier близько 0.76, overlap = 0.5, combined близько 0.68; для #cooking p\_classifier близько 0.005, overlap = 0, combined близько 0.0035.

## Висновки

Було досягнуто мету кваліфікаційної роботи магістра, а саме було підвищено достовірність та інтерпретованість оцінювання релевантності хештегів коротким текстовим повідомленням шляхом використання пояснюваного нейромережевого підходу, який поєднує трансформерну класифікацію з метриками узгодженості внутрішніх пояснень моделі зі змістовим ядром хештегу.

Для досягнення поставленої мети було поставлено та вирішено такі завдання:

- виконано аналіз проблеми автоматизованого оцінювання релевантності хештегів коротким текстовим повідомленням у цифрових комунікаційних середовищах, окреслено типові помилки класифікації й обмеження традиційних евристичних підходів;
- здійснено огляд і порівняльний аналіз сучасних нейромережевих та лінгвістичних підходів до класифікації коротких текстів, рекомендації тегів і пояснюваного аналізу рішень моделей обробки природної мови, визначено їх сильні та слабкі сторони в контексті задачі роботи;
- проаналізовано наявні набори даних соціальних мереж, обґрунтовано вибір конкретного датасету для формування пар «текст – хештег – мітка» та сформульовано систему показників якості оцінювання релевантності, придатну для подальших експериментів;
- розроблено нейромережевий метод пояснюваного оцінювання релевантності хештегів, у якому трансформерна класифікація поєднана з метриками узгодженості важливих фрагментів тексту зі змістовим ядром хештегу, а також визначено схему інтеграції цих метрик в єдину інтегральну оцінку;
- виконано підготовку та анотування вибірки коротких текстів і хештегів на основі даних соціальних мереж із формуванням позитивних і негативних прикладів для донавчання моделі, включаючи очищення, нормалізацію та побудову збалансованих підвибірок;
- здійснено програмну реалізацію запропонованого методу з використанням сучасних засобів глибинного навчання та пояснюваного штучного інтелекту, зокрема трансформерних мовних моделей і методів атрибуцій важливості;
- проведено експериментальні дослідження ефективності розробленого методу порівняно з базовими підходами, що ґрунтувалися лише на класифікаційних оцінках або векторних поданнях, та проаналізовано отримані числові результати;
- виконано поглиблений аналіз результатів, оцінено вплив пояснюваних метрик на якість відсікання нерелевантних хештегів і виявлення семантично коректних, але рідкісних тегів, а також сформульовано практичні рекомендації щодо застосування методу в системах підтримки роботи з хештегами та контент-менеджменту соціальних медіа.

## Anti-Plagiarism (UA) v-15.284 Educational

The maximum coincidence with one document 1.0%

Dictionaries check: en\_US, ru\_RU, ua\_UA. Errors in the documents: 19%

ID: 253286 Title: КВАЛІФІКАЦІЙНА РОБОТА на тему Метод оцінювання відповідності хештегів коротким текстам засобами обробки природної мови Added in a DB: 2025-12-16 Authors: Ольга ТРОХИМЧУК Heads: Олександр ПАСІЧНИК Consultants: Opponents:	Document		Sum coincidence on the DB	
	Symbols	Lexemes	Symbols	Lexemes
	135330	954	3310 (2%)	49 (5%)

### Plagiarism sources

ID	Description	Plagiarism presence in the document	
		Symbols	Lexemes

## Протокол аналізу звіту подібності науковим керівником

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

**Автор:** Ольга ТРОХИМЧУК

**Співавтор:**

**Назва:** КВАЛІФІКАЦІЙНА РОБОТА на тему Метод оцінювання відповідності хештегів коротким текстам засобами обробки природної мови

**Науковий керівник:** Олександр ПАСІЧНИК, к.т.н., доцент

**Підрозділ:** Кафедра комп'ютерних наук

**Коефіцієнт подібності 1:** 2.1%

**Коефіцієнт подібності 2:** 0.6%

**Мікропробіли:** 0

**Заміна букв:** 0

**Інтервали:** 0

**Білі знаки:** 1

**Дата створення звіту:** 2025-12-16 13:08:30.0

Після аналізу Звіту подібності констатую наступне:

Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.

Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.

Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедур. Таким чином робота не приймається.

Обґрунтування:

Дата 16.12.25

експерт

Петровський Р.Р. *Р.Р. Петровський*

**РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ КАФЕДРИ КОМП'ЮТЕРНИХ НАУК**

**ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ**

Назва кваліфікаційної роботи Метод оцінювання відповідності хештегів коротким  
текстам засобами обробки природної мови  
 Автор студент групи КНм-24-1 Ольга ТРОХИМЧУК  
 Освітня програма Комп'ютерні науки  
 Рівень вищої освіти другий (магістерський)  
 Спеціальність 122 – Комп'ютерні науки  
 Науковий керівник: к.т.н., доц. каф. КН Олександр ПАСІЧНИК

На основі аналізу кваліфікаційної роботи на дотримання вимог академічної доброчесності (у т.ч. відсутності ознак академічного плагіату) з урахуванням результатів перевірки роботи спеціалізованим програмним засобом комісія зробила такий висновок:

№	Висновок	Позначка про відповідність
1	Ознаки академічного плагіату	
1.1	Запозичення, виявлені в роботі, є законними і не є академічним плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних, якщо потрібно). Робота приймається до захисту.	<i>відповідає</i>
1.2	Виявлені запозичення не є академічним плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). Робота приймається до захисту, але має бути відкоригована.	
1.3	Виявлені запозичення не є академічним плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота може бути допущена до захисту після того як буде відкоригована та доопрацьована і успішно пройде повторну перевірку на академічний плагіат.	
1.4	Робота містить навмисні текстові спотворення, передбачувані спроби укріття текстових запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	
2	Інші види порушень академічної доброчесності	<i>відсутні</i>

**Підтвердження:**

*Запозичення, виявлені в роботі Ольги Трохимчук, не є плагіатом, оскільки: запозичення розміщені в розділі огляду існуючих підходів, не описують безпосередньо авторську роботу і не стосуються її результатів; усі запозичення фрагментарні; до запозичень входять фрагменти, які не мають авторства і містять поширені конструкції та загальновідомі терміни, скорочення. Рівень подібності не перевищує допустимої межі. Таким чином, робота є законною та приймається до захисту.*

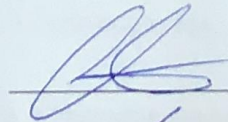
*Обсяг запозичень, визначений системами виявлення збігів/ідентичності/схожості:*

*- за системою Anti-Plagiarism: 2%;*

*- за системою StrikePlagiarism КП1: 2,1%, КП2: 0,6%.*

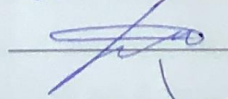
16.12.2025

Завідувач кафедри



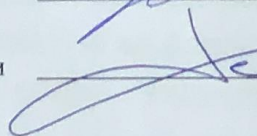
Олександр БАРМАК

Гарант освітньої програми



Руслан БАГРІЙ

Керівник кваліфікаційної роботи



Олександр ПАСІЧНИК



## ВІДГУК НАУКОВОГО КЕРІВНИКА

на кваліфікаційну роботу магістра

*зр. КНМ-24-1 Ольги ТРОХИМЧУК за темою: Метод оцінювання відповідності хештегів коротким текстам засобами обробки природної мови.*

### **1. Актуальність обраної теми**

*Оцінка відповідності хештегів коротким текстам є актуальною задачею для автоматизації аналізу соціальних мереж, контент-модерації та інформаційного пошуку. Використання нейромережесих моделей із механізмами пояснюваності дозволяє підвищити точність та прозорість оцінки релевантності, що робить тему роботи сучасною та практично значущою.*

**2. Відповідність роботи предметній області спеціальності 122 Комп'ютерні науки та загальним вимогам до наукових робіт**

*Робота цілком відповідає предметній області спеціальності 122 «Комп'ютерні науки». Вона демонструє застосування методів обробки природної мови, трансформерних архітектур та нейромережесих класифікації для оцінки відповідності хештегів, що відповідає вимогам кваліфікаційної роботи магістра.*

### **3. Професійні та особистісні якості магістранта**

*Під час виконання роботи магістрантка проявила високий рівень аналітичних здібностей та самостійності у реалізації складних алгоритмічних підходів. Робота виконана систематично та з увагою до деталей, що демонструє компетентність у сфері NLP та нейромережесих методів, здатність організувати робочий процес та дотримуватися термінів виконання завдань.*

### **4. Ступінь самостійності під час виконання кваліфікаційної роботи**

*Результати роботи є результатом самостійної наукової діяльності магістрантки. Було самостійно розроблено метод оцінювання відповідності хештегів, інтегровано механізм пояснюваності та реалізовано експериментальну перевірку ефективності на реальних даних.*

### **5. Наукова новизна та оригінальність запропонованих підходів**

*Наукова новизна роботи полягає у поєднанні трансформерної класифікації коротких текстів із показниками пояснюваності, що дозволяє враховувати узгодженість пояснень моделі з семантичним ядром хештегу. Запропоновано інтегральний показник*

релевантності, який синтезує оцінку моделі та інформативність текстових фрагментів, що забезпечує підвищення точності та прозорості рекомендацій.

#### **6. Ступінь оволодіння методами дослідження**

Магістрантка продемонструвала високий рівень володіння методами NLP, трансформерними архітектурами та інструментами пояснюваної нейромережевої класифікації.

#### **7. Повнота та якість розкриття теми роботи**

Тема роботи розкрита повністю: проведено огляд літератури, описано розроблений метод оцінювання відповідності хештегів, обґрунтовано інтегральну міру релевантності, спроектовано інтелектуальну систему та проведено експериментальне тестування. Враховано практичні аспекти застосування та обмеження системи.

#### **8. Логічність, послідовність, аргументованість, літературна грамотність викладу матеріалу**

Робота має логічну структуру та послідовне викладення матеріалу. Теоретичні положення підкріплено аргументованим аналізом сучасних підходів, результати експериментів представлені у зрозумілому вигляді з подальшим поясненням, виклад грамотний і відповідає академічним стандартам.

#### **9. Можливість практичного застосування кваліфікаційної роботи, окремих її частин**

Розроблений метод та інтелектуальна система можуть бути використані для автоматичного оцінювання релевантності хештегів у коротких текстах соціальних мереж та платформ мікроблогінгу. Інтеграція механізму пояснюваності дозволяє підвищити довіру користувачів і забезпечує можливість адаптації під різні домени та специфічні набори хештегів.

#### **10. Висновок про можливість допуску кваліфікаційної роботи до захисту, на яку оцінку заслуговує робота**

Робота виконана на високому науковому та практичному рівні, усі завдання реалізовані, наукова новизна підтверджена. Роботу можна допустити до захисту, рекомендована оцінка – добре.

Науковий керівник \_\_\_\_\_ к.т.н., доц. каф. КН Олександр ПАСІЧНИК



## ВІДГУК ОПОНЕНТА

### на кваліфікаційну роботу магістра

зр. КНМ-24-1 Ольги ТРОХИМЧУК за темою: Метод оцінювання відповідності хештегів коротким текстам засобами обробки природної мови.

#### 1. Актуальність обраної теми

Аналіз відповідності хештегів коротким текстам є важливою складовою автоматизації цифрових комунікацій, підвищення точності інформаційного пошуку та рекомендацій у соціальних мережах. Використання нейромережесвих методів із пояснюваними механізмами забезпечує сучасний підхід до підвищення якості рекомендацій і робить тему роботи актуальною та практично значущою.

#### 2. Відповідність роботи предметній області спеціальності 122 Комп'ютерні науки та загальним вимогам до наукових робіт

Тема роботи кваліфікаційної магістра відповідає предметній області спеціальності 122 «Комп'ютерні науки». Робота демонструє застосування трансформерних моделей, методів обробки природної мови, нейромережесвої класифікації та інтегрованих показників релевантності для оцінювання хештегів у коротких текстах, що відповідає вимогам кваліфікаційної роботи магістра.

#### 3. Повнота розкриття мети та завдань дослідження

Мета роботи та поставлені завдання реалізовані повністю. Розроблено метод оцінювання релевантності хештегів, інтегровано механізм пояснюваності на основі лексичного й семантичного аналізу тексту, а також проведено експериментальну перевірку результатів на різних наборах даних.

#### 4. Наявність наукової новизни

Наукова новизна роботи полягає у поєднанні трансформерної класифікації коротких текстів із оцінкою пояснюваності, що дозволяє враховувати узгодженість внутрішніх пояснень моделі із семантичним ядром хештегу. Введено інтегральний показник релевантності, який синтезує оцінку моделі та вагу інформативних фрагментів тексту, що суттєво розширює можливості існуючих підходів.

## 5. Зміст кожного розділу роботи

Робота складається з чотирьох розділів. Перший розділ присвячено аналізу сучасного стану досліджень у сфері оцінювання релевантності хештегів та коротких текстів. Другий розділ описує розроблений метод та формування інтегральної міри релевантності. Третій розділ містить проєктування інтелектуальної системи, включаючи її структуру, елементи та метрики оцінювання. Четвертий розділ присвячено експериментальній перевірці методу та оцінці ефективності системи.

## 6. Ступінь розкриття теми роботи

Тема роботи розкрита повністю: проведено огляд літератури, розроблено метод оцінювання релевантності хештегів із врахуванням пояснюваності, створено інтелектуальну систему та проведено експериментальне тестування.

## 7. Якість оформлення кваліфікаційної роботи

Робота оформлена належним чином, розділи структуровані логічно, таблиці та рисунки коректно ілюструють результати. Посилання на джерела оформлені відповідно до академічних стандартів, що забезпечує високу якість викладу матеріалу.

## 8. Недоліки кваліфікаційної роботи

Метод залежить від прямого лексичного збігу тексту та ядра хештегу, обмежений англомовними короткими текстами та навчальною вибіркою. Це може знижувати узагальнювальну здатність у інших мовах, доменах або при нестандартних форматах повідомлень. Рекомендується додатково протестувати систему на мультидомених корпусах, а також провести аналіз впливу розміру та структури хештегів на ефективність інтегральної міри релевантності.

## 9. Загальний висновок (допускається чи не допускається до захисту), якої оцінки заслуговує кваліфікаційна робота

Робота виконана на високому науковому та практичному рівні, усі завдання реалізовані, наукова новизна підтверджена. Роботу можна допустити до захисту, рекомендована оцінка – добре.

Опонент (прізвище, ім'я, по батькові, посада, місце роботи)

Ткачук Єлизавета Тимонівна, д.т.н., проф., професор  
«15» 12 2025 р. каф. КІС

Підпис