

МАЗУРЕЦЬ ОЛЕКСАНДР

Хмельницький національний університет

<https://orcid.org/0000-0002-8900-0650>e-mail: exe.chong@gmail.com**МОЛЧАНОВА МАРИНА**

Хмельницький національний університет

<https://orcid.org/0000-0001-9810-936X>e-mail: m.o.molchanova@gmail.com**КЛІМЕНКО ВАЛЕРІЯ**

Хмельницький національний університет

<https://orcid.org/0000-0001-5869-4269>e-mail: ler.klimenko.8@gmail.com**СОБКО ОЛЕНА**

Хмельницький національний університет

<https://orcid.org/0000-0001-5371-5788>e-mail: olena.sobko.ua@gmail.com**СУПРУН ПАВЛО**

Хмельницький національний університет

e-mail: pasha1.apple@gmail.com

ДАТАЛОГІЧНА МОДЕЛЬ БАЗИ ДАНИХ ДЛЯ ВИЯВЛЕННЯ ГЕНДЕРНОЇ ПРИНАЛЕЖНОСТІ ЗА SVM-АНАЛІЗОМ ДОПИСІВ ІНТЕРНЕТ-МЕРЕЖ З ВИКОРИСТАННЯМ ОБ'ЄКТНО-ОРІЄНТОВАНОГО ПРОЄКТУВАННЯ

У роботі запропоновано практичний підхід до виявлення гендерної приналежності за аналізом дописів інтернет-мереж із використанням SVM-класифікаторів. Для дослідження ефективності запропонованого методу було створено об'єктно-орієнтовану програмну реалізацію в середовищі програмування PyCharm, а також виконано даталогічне моделювання структури даних. Розроблений підхід показав високу ефективність, в порівнянні з існуючим аналогом його точність вища на 0.11. Перевагою методу є можливість працювати з короткими текстами, такими як твіти, без втрати точності. Одержані результати можуть бути актуальними для різноманітних застосувань, включаючи маркетингові дослідження, аналіз громадської думки, персоналізовану рекламу, політичні дослідження, й сприти створенню безпечних і толерантних вебсередовищ.

Ключові слова: гендерна приналежність, дописи інтернет-мереж, твіти, об'єктно-орієнтована модель, даталогічна модель, цисгендер, толерантне середовище, гендер, SVM

MAZURETS OLEKSANDR, MOLCHANOVA MARYNA, KLIMENKO VALERIIA, SOBKO OLENA, SUPRUN PAVLO
Khmelnytskyi National University

DATALOGIC MODEL FOR IDENTIFYING GENDER BY SVM-ANALYSIS OF INTERNET POSTS USING OBJECT-ORIENTED PROJECTING

The article proposes the practical approach to gender identification based on the analysis of Internet posts using SVM classifiers. The input data of the approach is a set of SVM classifiers trained on English-language data and the corresponding vectorizers used during training and the post of social Internet networks for analysis. The zero step is to select the gender species to analyze and load the corresponding SVM classifier model with vectorizer. The first step is to pre-process the social media post, which includes removing stop characters and stop words, as well as checking the writing language. If the language of the post is not English, there is an automated translation into English. The next step is the vectorization of the pre-processed post of social Internet networks, after which the step of classification by a trained SVM classifier takes place. The fourth step is the formation of conclusions to the user in the original language of the post of social Internet networks, since the classifier itself works with English-language data. The initial data of approach is the assessment of the post's belonging to the specified gender for analysis.

To investigate the effectiveness of the proposed method, an object-oriented software implementation was created in the PyCharm programming environment, and data logic modeling of the data structure was also performed. The developed approach showed high efficiency, compared to the existing analogue, its accuracy is higher by 0.11. The advantage of the method is the ability to work with short texts, such as tweets, without losing accuracy. The obtained results can be relevant for a variety of applications, including marketing research, public opinion analysis, personalized advertising, political research, and contribute to the creation of safe and tolerant web environments.

Keywords: gender identity, internet posts, tweets, object-oriented model, data logic model, cisgender, tolerant environment, gender, SVM

Аналіз предметної області

Зростання використання соціальних мереж і стійка популярність онлайн-комунікацій робить завдання виявлення гендерної приналежності за дописами необхідним для різноманітних застосувань, включаючи маркетингові дослідження, аналіз громадської думки, персоналізовану рекламу, політичні дослідження тощо.

Завдяки розвитку методів машинного навчання та обробки природної мови, можливості виявлення гендерної приналежності зростають. Ці дані можуть бути використані для аналізу і розуміння поведінки користувачів в інтернеті, виявлення та прогнозування тенденцій в споживчому ринку, а також для створення

більш ефективних стратегій спілкування та взаємодії з аудиторією. Використання таких методів дозволяє уникнути стереотипів та прихованих асоціацій, а також забезпечити репрезентативність аналізу за гендерними критеріями.

Зростаюча увага до питань рівності та різноманітності також робить цей напрям актуальним, оскільки він може використовуватися для аналізу та виявлення можливих стереотипів, дискримінації або нерівності, що можуть існувати у онлайн-спільнотах [1]. Відповідно, виявлення гендерної приналежності за допомогою штучного інтелекту має великий потенціал у багатьох сферах, що робить напрям дослідження актуальним в сучасному інформаційному суспільстві. Великий обсяг структурованих даних при вирішенні задачі передбачає обов'язковий етап даталогічного моделювання бази даних, а прикладна програмна реалізація для ефективної роботи передбачає використання об'єктно-орієнтованого підходу.

Автоматизація виявлення гендерної приналежності за SVM-аналізом дописів інтернет-мереж сприятиме створенню безпечних і толерантних вебсередовищ.

Сучасний стан досліджень

Проблемою виявлення гендерної приналежності за аналізом дописів інтернет-мереж наразі займається велика кількість науковців по всьому світу, зокрема багато досліджень зосереджено в напрямку гендерної лінгвістики [2]. Зважаючи на те, що мова є важливим засобом вираження особистості, існує кілька характеристик тексту, які можуть вказувати на можливі гендерні відмінності у способі спілкування.

Лінгвістична гендерологія розглядає мову як ключовий засіб для розуміння гендеру як окремої дисципліни та вивчає взаємозв'язок між мовою та гендером як складовими соціокультурного контексту [3]. Це дозволяє вивчати, як сама мова відтворює та формує стереотипи, ролі та ідеї про гендер, а також розглядати мовленнєву поведінку як важливий аспект гендерної ідентичності.

Наука гендерна лінгвістика вивчає мову й комунікацію з урахуванням гендерних аспектів з метою продукції нових знань у сфері лінгвістики. Це включає дослідження гендерних відмінностей у вживанні мови, аналіз гендерного впливу на мовленнєві структури та практики, а також розробку й застосування гендерно-чутливих методів дослідження [2].

На сьогоднішній день термін «гендер» використовується у різних сферах суспільства. У лінгвістиці, вперше, слово «гендер» описувало граматичну відмінність за ознаками чоловічої, жіночої й середньої статей в англійській мові. У науковому контексті термін «гендер» використовується для позначення соціальної статі [1].

Гендерна ідентичність є внутрішнім сприйняттям статевої сутності та вираження гендерних характеристик через одяг, поведінку й зовнішній вигляд. Це психосоціальний конструкт, який виникає у ранньому віці і може відрізнятися від статевої приналежності, присвоєної при народженні. Хоча гендер традиційно розглядали як бінарний, сучасна розуміння визнає багато різновидів гендерної ідентичності, включаючи трансгендерність, небінарність та інші форми [4].

Різноманітність гендерів охоплює широкий спектр термінів і понять, які використовуються для характеристики гендерної ідентичності людей. Гендери розподіляються на види:

- цисгендерність – коли гендерна ідентичність співпадає з призначеною статтю або гендером тіла;
- трансгендерність – гендерна ідентичність, яка відрізняється від біологічної статі при народженні;
- гендерквірність (або небінарність) – гендерна ідентичність, що не вписується у традиційні бінарні уявлення про стать, або може поєднувати елементи різних гендерних ідентичностей;
- інтерсексуальність – гендерна ідентичність людини, народженої з репродуктивними органами, рівнем гормонів або статевими хромосомами, які не є виключно чоловічими або жіночими [4].

Наведені вище різноманітні гендери мають певні риси на письмі, які можуть сприйматися як характеристики, що сприяють відокремленню ознак гендерів й враховуються при автоматизованому аналізі текстів з метою виявлення гендерних відмінностей за допомогою методів NLP.

Прикладом роботи в наведеному напрямку є «Hacker Factor: Gender Guesser» [5]. Цей інструмент намагається визначити гендер автора на основі використаних слів. Він аналізує поданий текст за двома типами письма: формальним та неформальним. Формальне письмо включає художню та наукову літературу, статті та новини. Неформальне письмо включає тексти блогів і чат-кімнат. Цей метод базується на роботі вчених з Іллінойського технологічного інституту та Бар-Іланського університету, які розробили метод оцінки гендеру на основі використання слів, де враховували вагові частоти слів та частин мови, щоб визначити гендер автора. Ці дослідження показали, що навіть в різних жанрах письма (наприклад, художня та наукова література, блоги) є специфічні частоти слів для кожної статі.

З урахуванням надбань науковців в галузі виявлення гендерної приналежності за дописами соціальних інтернет-мереж, здебільшого науковцями використовується обраний вище підхід на основі SVM. Зважаючи на значну увагу науковців, напрям актуальний і потребує продовження досліджень.

Мета дослідження полягає в проектуванні та розробці програмного забезпечення для виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP, що включає аспекти даталогічного та об'єктно-орієнтованого моделювання.

Основна частина

Виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP виконується для аналізу і класифікації текстових даних з метою визначення гендерної приналежності, яка здійснюється на основі особливостей мови та стилю письма, що є характерними для певних груп

користувачів. Схема та кроки методу наведені на рисунку 1. Відповідний метод працює шляхом перетворення вхідних даних у вигляді набору натренованих класифікаторів SVM та вектори заторів [6], що використовувались під час навчання й допису соціальних інтернет-мереж для аналізу в вихідні дані у вигляді оцінки приналежності допису для аналізу до визначеного гендеру [7].

Вхідними даними методу є набір навчених на англійських даних класифікаторів SVM та відповідних векторизаторів, що використовувались під час навчання та допис соціальних інтернет-мереж для аналізу. Нульовим кроком є вибір виду гендера для аналізу та завантаження відповідної моделі класифікатора SVM з векторизатором. Першим кроком є попередня обробка допису соціальних інтернет-мереж, що включає в себе видалення стоп-символів і стоп-слів, а також перевірку мови написання. Якщо мова допису не англійська, відбувається автоматизований переклад на англійську. Наступним кроком здійснюється векторизація попередньо обробленого допису соціальних інтернет-мереж, після чого відбувається крок класифікації натренованим SVM-класифікатором [7]. Четвертим кроком є формування висновків користувачу мовою оригіналу допису соціальних інтернет-мереж, оскільки сам класифікатор працює із англійськими даними. Вихідними даними методу є оцінка приналежності допису для аналізу до визначеного гендеру.

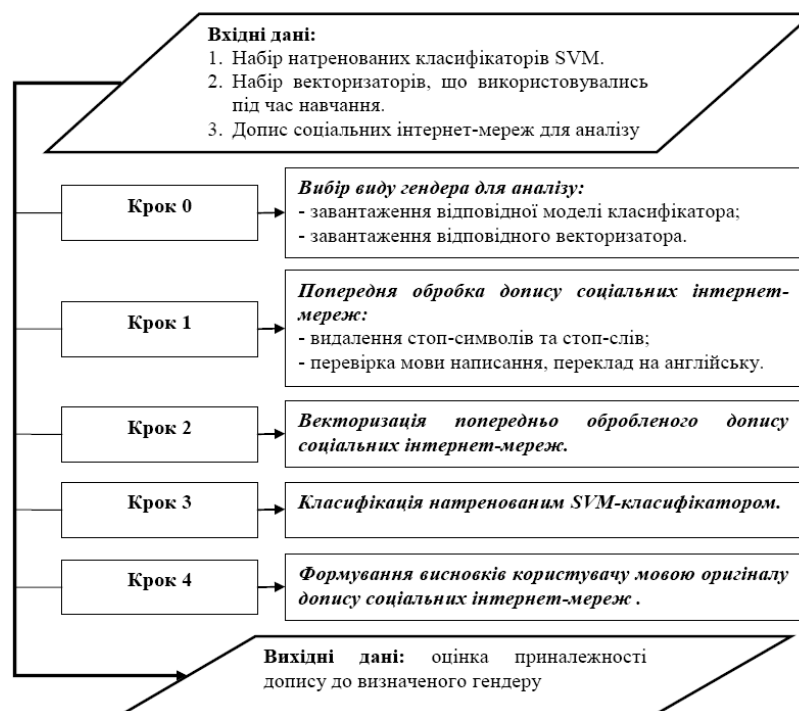


Рис. 1. Кроки виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP

На рисунку 2 наведено проектну архітектуру інформаційної системи з виявлення гендерної приналежності за дописами соціальних інтернет-мереж і взаємозв'язок компонентів. Наведена система складається із 3-х підсистем та бази даних.

Підсистема виявлення гендерної приналежності за дописом є головною підсистемою інформаційної системи, та призначена для виявлення гендеру за користувацьким текстовим дописом. Також включає в себе можливості вибору виду гендеру для аналізу, вибір датасету для аналізу, виведення статистики щодо обраного датасету, вибір гендеру за обраним видом, вибір тестового допису з датасету або уведення «вручну», визначення гендерної приналежності автора за дописом. Аналіз допису для виявлення гендеру здійснюється на підставі класифікатора SVM, попередньо натренованого засобами підсистеми навчання моделей SVM.

Підсистема навчання моделей SVM призначена для тренування SVM-моделей, які призначені для виявлення гендерів. Для кожного виду гендеру повинен бути натренований свій класифікатор та відповідний векторизатор. Також дана підсистема включає в себе такі можливості, як вибір користувацьких параметрів для SVM-моделі, вибір датасету для навчання класифікаторів, навчання моделей SVM за користувацькими параметрами, навчання моделей SVM автоматизованим підбором параметрів, виведення статистики навчання, збереження навченої моделі та векторизатора.

Підсистема редактора датасета за гендерами призначена для редагування датасетів, що використовуються для навчання і тестування класифікаторів. Вмісти датасетів та решти параметрів класифікаторів знаходяться у базі даних, тому дана підсистема є також редактором вмісту бази даних. Також виконує такі функції: вибір виду гендеру для аналізу, вибір гендеру за обраним видом, виведення дописів датасету для обраного гендеру, деталізація обраного допису датасету, визначення мови допису, додавання нового допису, редагування обраного допису, видалення обраного допису.

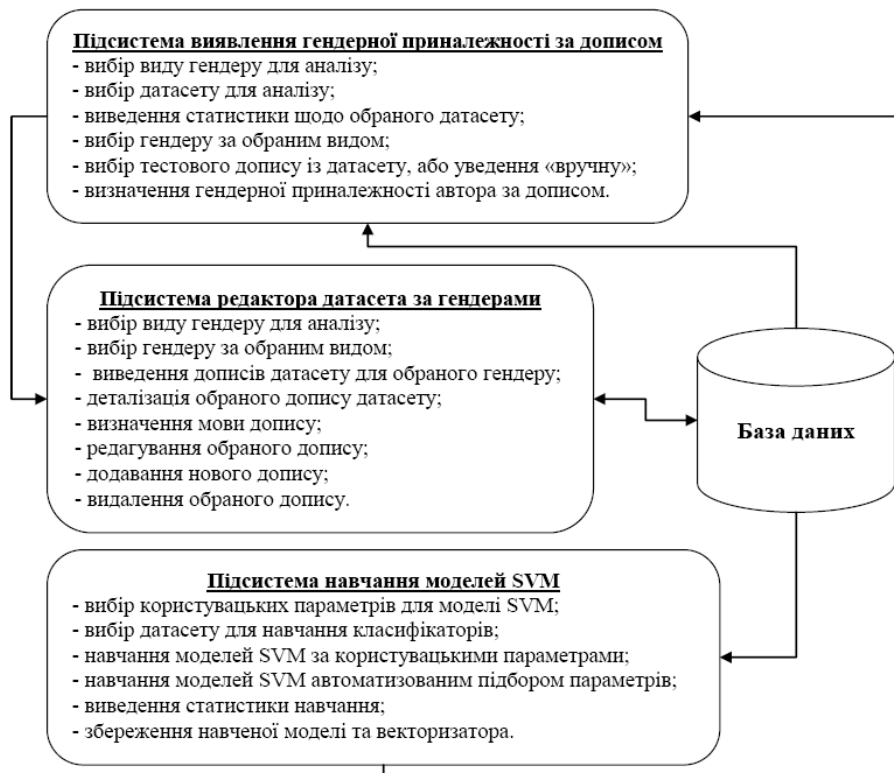


Рис. 2. Архітектура інформаційної системи з виявлення гендерної приналежності за дописами

На рисунку 3 зображено спроектовану даталогічну модель бази даних для виявлення гендерної приналежності за дописами, яка візуально демонструє структуру бази даних, включаючи її таблиці, поля і зв'язки між ними. База даних містить таблиці: genders, subject_types, webservices, subjects, posts_for_analysis, post_analysis, posts_for_training, subject_gender_analysis, gender_analysis_for_1_subject, types_of_genders.

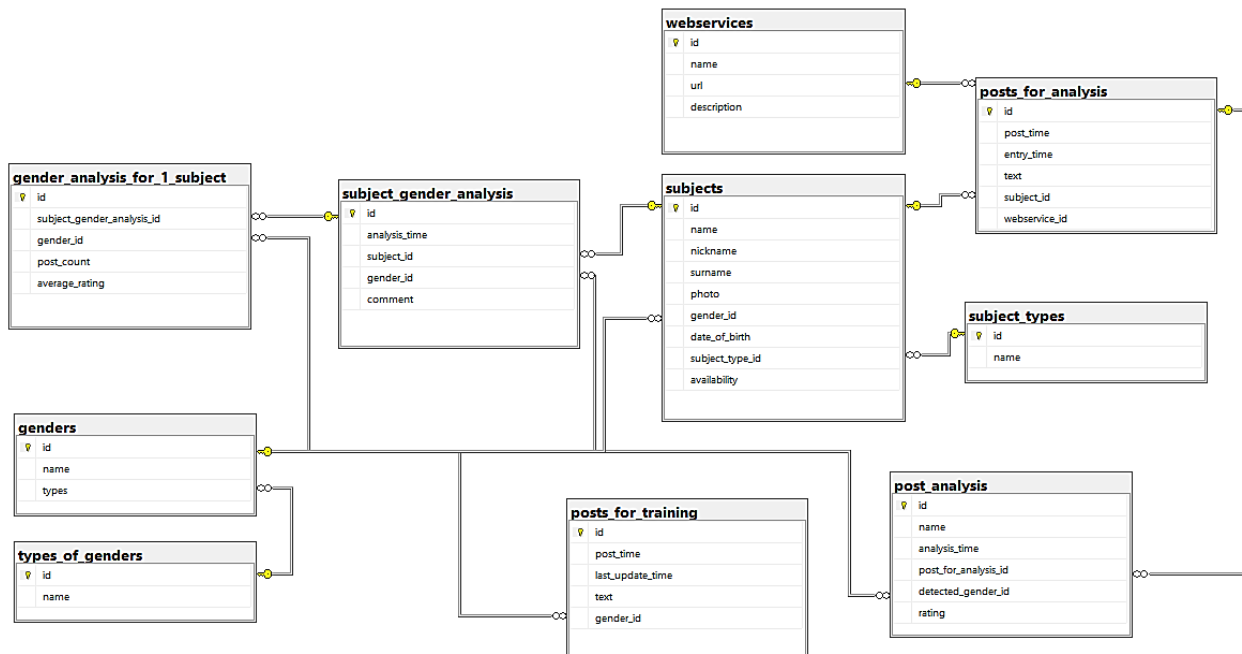


Рис. 3. Даталогічна модель бази даних для виявлення гендерної приналежності за дописами

Таблиця «Subjects» описує атрибути, які використовуються для зберігання інформації про суб'єктів, що підлягатимуть аналітичній обробці. Включає його ім'я, нікнейм, прізвище, фотографію, дату народження, ідентифікатор гендеру, тип суб'єкта та його доступність. Атрибути таблиці: унікальний ідентифікатор суб'єкта «id» (INT), ім'я суб'єкта «name» (VARCHAR), нікнейм суб'єкта «nickname» (VARCHAR), прізвище суб'єкта «surname» (VARCHAR), фотосвітлина суб'єкта «photo» (VARCHAR), гендер «gender_id» (INT), дата народження суб'єкта «date_of_birth» (DATE), тип суб'єкта «subject_type_id» (INT), доступність суб'єкта «availability» (BIT).

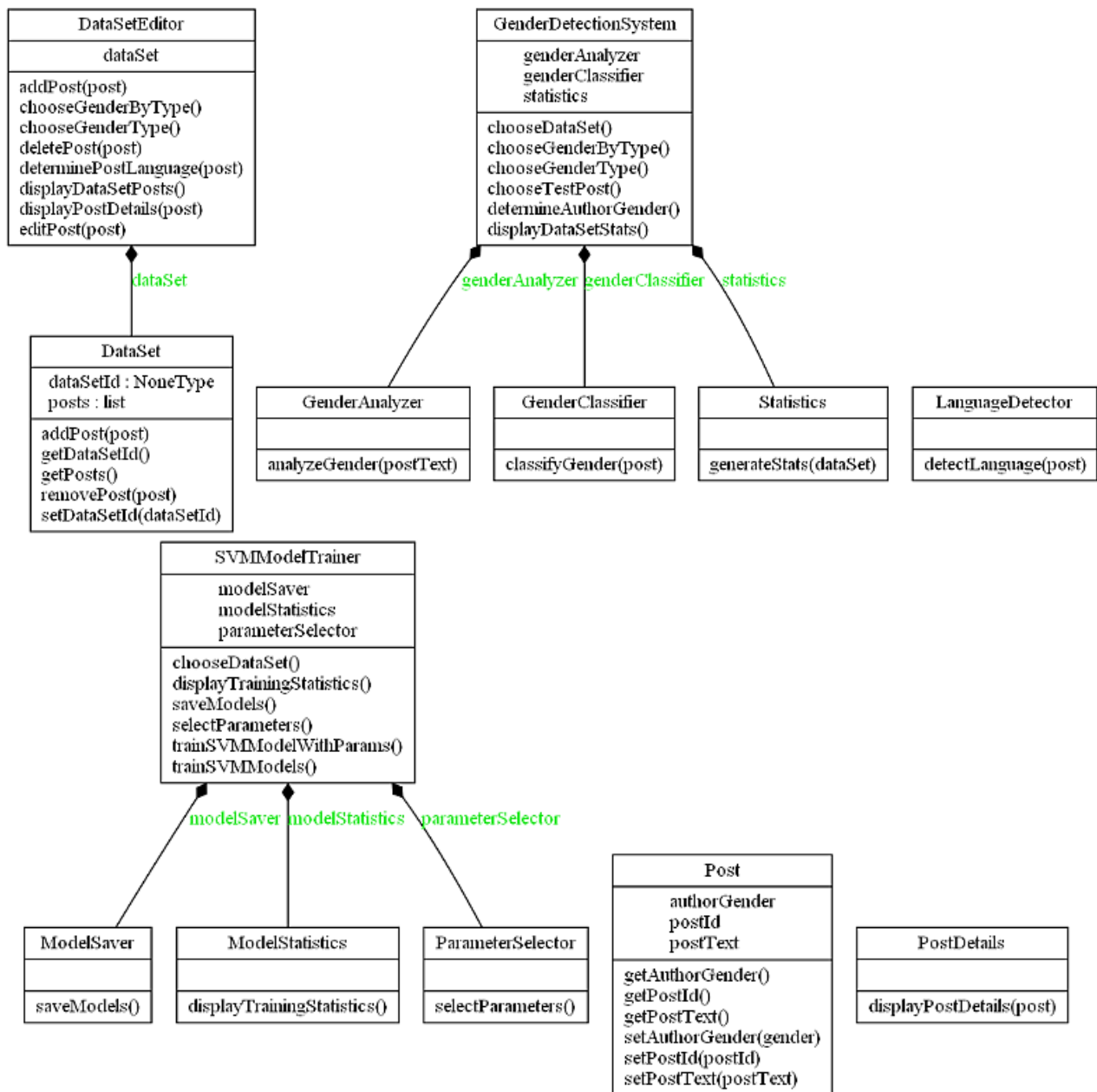


Рис. 4. Об'єктно-орієнтована діаграма класів для прикладного виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP

Таблиця «Posts_for_analysis» містить структуровані дані про публікації, які підлягають аналітичній обробці. До неї включено такі атрибути: дату і час публікації, дату та час внесення публікації в базу даних, текст публікації, суб'єкт, який є власником публікації, та вебсервіс розміщення публікації. Атрибути таблиці «Posts_for_analysis»: унікальний ідентифікатор публікації «id» (INT), дата та час публікації «post_time» (DATETIME), дата та час внесення публікації в базу даних «entry_time» (DATETIME), текст публікації «text» (TEXT), власник публікації «subject_id» (INT), вебсервіс розміщення публікації «webservice_id» (INT).

Таблиця «Webservices» зберігає дані про вебсервіси. Включає назву вебсервісу, гіперпосилання на нього та його опис. Атрибути таблиці: унікальний ідентифікатор вебсервісу «id» (INT), назва вебсервісу «name» (VARCHAR), гіперпосилання на вебсервіс «url» (VARCHAR), опис вебсервісу «description» (TEXT).

Таблиця «Post analysis» описує атрибути, які використовуються для зберігання інформації про результати аналізу публікацій: назву аналізу, його дату і час, публікацію для аналізу, гендер власника публікації, та її оцінку. Атрибути таблиці «Post analysis»: унікальний ідентифікатор аналізу публікації «id» (INT), назва аналізу публікації «name» (VARCHAR), дата і час аналізу «analysis_time» (DATETIME), публікація для аналізу «post_for_analysis_id» (INT), гендер власника публікації «detected_gender_id» (INT), оцінка публікації «rating» (FLOAT).

Таблиця «Posts_for_training» надає пояснення атрибутів, що використовуються для опису публікацій, які застосовуються для тренування моделі: дату та час публікації, дату та час останнього оновлення публікації, текст публікації та гендер власника публікації. Атрибути таблиці «Posts_for_training»: унікальний ідентифікатор публікації для навчання «id» (INT), дата і час публікації «post_time» (DATETIME),

дата і час останнього оновлення публікації «last_update_time» (DATETIME), текст публікації «text» (TEXT), гендер власника публікації «gender_id» (INT).

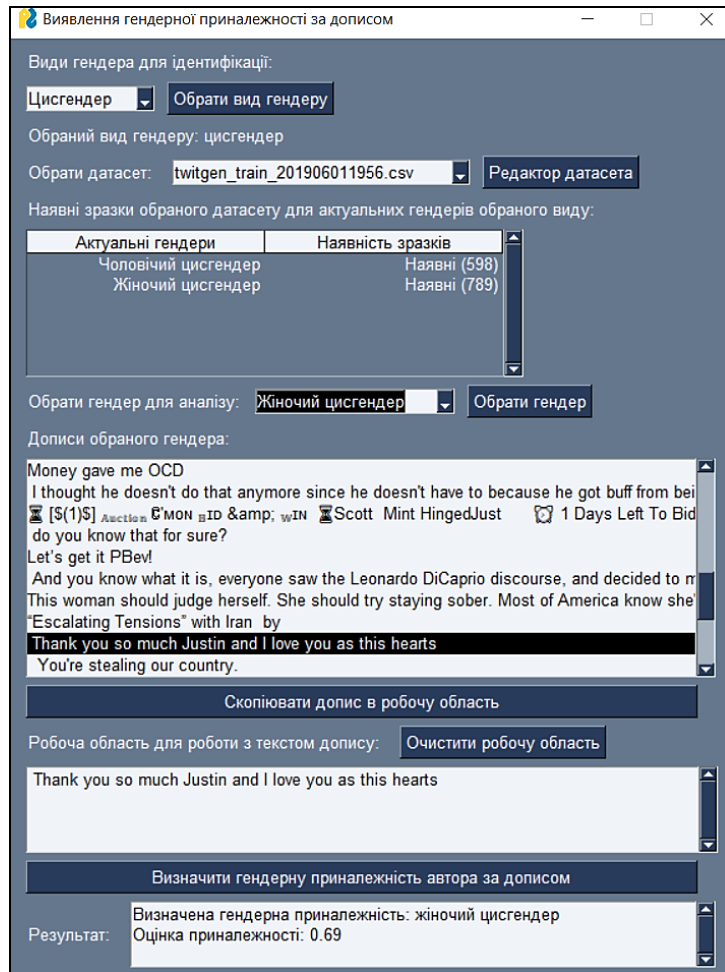


Рис. 5. Визначення гендерної приналежності за дописами засобами NLP

Таблиця «Subject_gender_analysis» містить інформацію про аналітичні процедури, що стосуються гендерної ідентифікації суб'єктів. Включаючи дату та час підрахунку, суб'єкт та його гендер, а також коментар до аналізу. Атрибути таблиці: унікальний ідентифікатор аналізу приналежності суб'єкта до гендеру «id» (INT), дата і час підрахунку «analysis_time» (DATETIME), суб'єкт «subject_id» (INT), гендер «gender_id» (INT), коментар до аналізу «comment» (VARCHAR).

Таблиця «Gender_analysis_for_1_subject» містить узагальнену інформацію про результати аналізу гендерної ідентифікації суб'єктів. Вона включає в собі аналіз приналежності суб'єкта до гендеру, гендер суб'єкта, кількість публікацій, що вказують на цей гендер, та середню оцінку всіх аналізів, що вказують на цей гендер. Атрибути таблиці «Gender_analysis_for_1_subject»: унікальний ідентифікатор аналізу приналежності суб'єкта до одного гендеру «id» (INT), аналіз приналежності суб'єкта до гендеру «subject_gender_analysis_id» (INT), гендер суб'єкта «gender_id» (INT), кількість публікацій, що вказують на цей гендер «post_count» (INT), середня оцінка всіх аналізів, що вказують на цей гендер «average_rating» (FLOAT).

Для розробки програмного забезпечення для виявлення гендерної приналежності за використано інтегроване середовище програмування PyCharm, мову програмування Python і мову запитів SQL. Програмні складові системи відповідають розробленій вище проектній архітектурі системи. Схема діаграми класів наведена на рисунку 4.

Підсистема виявлення гендерної приналежності за дописом (рисунок 5). Клас «GenderDetectionSystem» представляє основну підсистему, яка координує процес виявлення гендерної приналежності. Він містить інші класи, необхідні для виконання цього завдання: «GenderAnalyzer», «GenderClassifier», «Statistics».

Підсистема редактора датасета за гендерами. Клас «DataSetEditor» відповідає за редагування та управління датасетом. Він містить методи для додавання, редагування і видалення дописів, а також для визначення мови допису, такі як: editPost(), addPost(), deletePost(), determinePostLanguage(). Клас «PostDetails» відображає деталізовану інформацію про окремий допис, таку як текст та інші атрибути: displayPostDetails().

Підсистема навчання моделей SVM. Клас «SVMModelTrainer» відповідає за навчання моделей SVM

та управління параметрами навчання. Він містить методи для вибору параметрів, навчання моделей і збереження навчених моделей: `selectParameters()`, `trainSVMModels()`, `saveModels()`. Клас «ParameterSelector» допомагає вибрати параметри для навчання моделей `selectParameters()`. Клас «ModelStatistics» відображає статистику навчання, використовує метод `displayTrainingStatistics()`. Клас «ModelSaver» зберігає навчені моделі, використовує метод `saveModels()`.

Експерименти та дискусія

Для дослідження ефективності прикладного виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP запропонованим методом, було виконане порівняння роботи створеного програмного застосунку й існуючого рішення «Gender Guesser» [17]. «Gender Guesser» погано працює з текстами довжиною менше 300 слів та дозволяє на думку авторів виявити цисгендер з точністю близько 60-70%. Однак зважаючи на специфіку набору навчальних даних, тестування буде проведено на твітах. Для експерименту взято 50 твітів з валідаційної вибірки випадковим чином. Після вибору даних, 27 твітів належать категорії «чоловічий цисгендер» та 23 твіта категорії «жіночий цисгендер». «Gender Guesser» надісланий текст оцінює за двома типами написання: формальним і неформальним. Офіційне написання включає в себе художні та науково-популярні історії, статті та новини. Неформальне письмо включає тексти в блогах і чатах. Отже, для експерименту буде використано колонку неформального результату. Результати виявлення цисгендеру за допомогою програми «Gender Guesser» та створеного методу наведено в таблицях 1 і 2 відповідно. У таблиці 3 наведено спільні результати у відсотковому значенні, що також проілюстровані на діаграмі на рисунку 6.

Таблиця 1

Результати виявлення цисгендеру за допомогою «Gender Guesser»

	Чоловічий цисгендер	Жіночий цисгендер
Кількість коректно ідентифікованих дописів	15	12
Кількість некоректно ідентифікованих дописів	12	11
Разом дописів:	27	23

Таблиця 2

Результати виявлення цисгендеру за допомогою розробленого методу

	Чоловічий цисгендер	Жіночий цисгендер
Кількість коректно ідентифікованих дописів	18	16
Кількість некоректно ідентифікованих дописів	9	7
Разом дописів:	27	23

Таблиця 3

Порівняння точності виявлення цисгендеру розробленим методом та «Gender Guesser»

Коректно ідентифіковані дописи	Чоловічий цисгендер	Жіночий цисгендер
«Gender Guesser»	0.555555556	0.52173913
Розроблений метод	0.666666667	0.695652174

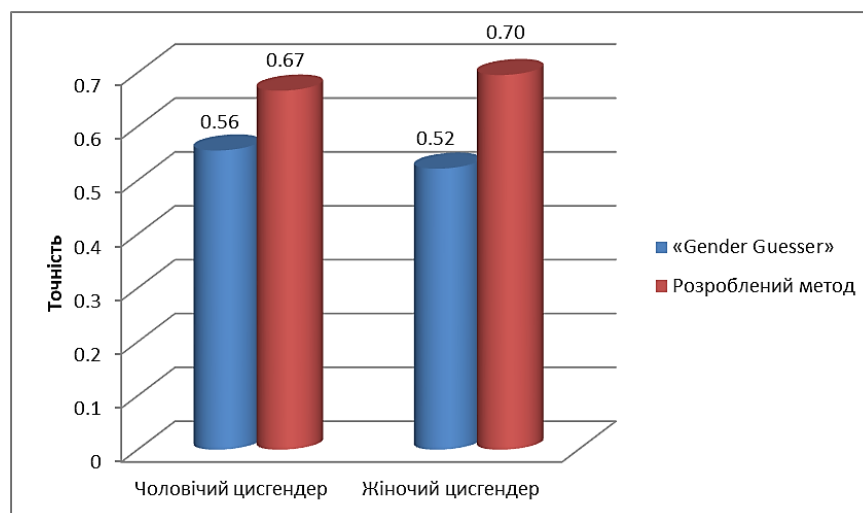


Рис. 6. Діаграма результатів щодо виявлення цисгендеру альтернативними підходами

Як видно з таблиці 3 та рисунку 6, розроблений метод показав вищу ефективність. Це може бути пов'язано із прикладами, на яких навчались класифікатори, адже навчання SVM було здійснено на базі твітів, а в застосунку з яким відбувалось порівняння вказано, що для збільшення точності необхідно використовувати тексти понад 300 слів, а це є доволі великий обсяг.

Однак, точність розробленого методу також можна покращити. Вибір стилю письма часто залежить від різних факторів, таких як освіта, досвід, професія, вік та національність автора. Наприклад, представниця

жіночого цисгендеру, яка працює у галузі, де переважають особи чоловічого цисгендеру може прийняти їх стиль письма. Автори з великим досвідом часто використовують професійні стилі письма, незалежно від гендерних особливостей. Для більшої точності класифікації необхідно більше даних для навчання і вони повинні бути більше збалансовані.

Висновки

У статті розглянуто поточний стан наукового напряму виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP, де на основі опрацьованого матеріалу запропоновано створення використання методу SVM для аналізу дописів соціальних інтернет-мереж з метою виявлення гендерної приналежності авторів, що є актуальним для різноманітних застосувань, включаючи маркетингові дослідження, аналіз громадської думки, персоналізовану рекламу, політичні дослідження і багато іншого, й сприяє створенню безпечних і толерантних вебсередовищ. Відповідний метод працює шляхом перетворення вхідних даних у вигляді набору натренованих класифікаторів SVM та векторизаторів, що використовувались під час навчання й допису соціальних інтернет-мереж для аналізу у вихідні дані у вигляді оцінки приналежності допису для аналізу до визначеного гендеру.

У межах дослідження було спроектовано та розроблено програмне забезпечення для виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP, що включало аспекти датового та об'єктно-орієнтованого моделювання. Було виконано дослідження ефективності створеного методу виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP. Для дослідження ефективності прикладного виявлення гендерної приналежності за дописами соціальних інтернет-мереж засобами NLP запропонованим методом, було виконане порівняння роботи створеного програмного застосунку й існуючого рішення. Розроблений метод показав високу ефективність, у порівнянні із існуючим аналогом його точність вища на 0.11. Перевагою методу визначено те, що він може працювати з короткими текстами без втрати точності, зокрема такими як твіти. Подальші дослідження будуть спрямовані на виявлення додаткових ознак класифікації, а також на роботу з іншими видами гендерів.

Література

1. Понятійний апарат гендерних досліджень [Електронний ресурс]. – 2024. – Режим доступу : <https://learn.ztu.edu.ua/mod/resource/view.php?id=187636>.
2. Гендерна лінгвістика як наука [Електронний ресурс]. – 2024. – Режим доступу : <https://studfile.net/preview/7157868/page:5/>
3. Гендерна лінгвістика [Електронний ресурс]. – 2024. – Режим доступу : https://uk.wikipedia.org/wiki/Гендерна_лінгвістика
4. Гендер. Поняття, суть, види та ознаки [Електронний ресурс]. – 2024. – Режим доступу : https://termin.in.ua/hender/#Vidi_ci_tipi_genderiv
5. Gender Guesser [Електронний ресурс]. – 2024. – Режим доступу : <https://www.hackerfactor.com/GenderGuesser.php>
6. Молчанова М. О. Алгоритм виявлення аб'юзивного вмісту в україномовному аудіоконтенті для імплементації в об'єктно-орієнтовану інформаційну систему / М. О. Молчанова, О. В. Мазурець, О. В. Собко та ін. // Вісник Хмельницького національного університету» серія: Технічні науки. – 2024. – № 1 (331). – С. 101–106.
7. Slobodzian V., Molchanova M., Kovalchuk O. An Approach Based on the Visualization Model for the Ukrainian Web Content Classification. ACIT. 2022. P. 400–405.

References

1. Poniatiyniyi aparat hendernykh doslidzhen [Elektronnyi resurs]. – 2024. – Rezhym dostupu do resursu: <https://learn.ztu.edu.ua/mod/resource/view.php?id=187636>.
2. Henderna linhvistyka yak nauka [Elektronnyi resurs]. – 2024. – Rezhym dostupu do resursu: <https://studfile.net/preview/7157868/page:5/>
3. Henderna linhvistyka [Elektronnyi resurs]. – 2024. – Rezhym dostupu do resursu: https://uk.wikipedia.org/wiki/Henderna_linhvistyka
4. 4. Hender. Poniattia, sut, vydy ta oznaky [Elektronnyi resurs]. – 2024. – Rezhym dostupu do resursu: https://termin.in.ua/hender/#Vidi_ci_tipi_genderiv
5. Gender Guesser [Elektronnyi resurs]. – 2024. – Rezhym dostupu do resursu: <https://www.hackerfactor.com/GenderGuesser.php>
6. Alhorytm vyivlennia ab'iuzyvnoho vmistu v ukrainomovnomu audiokontenti dlia implementatsii v obiektno-oriietovanu informatsiinu systemu / [M. O. Molchanova, O. V. Mazurets, O. V. Sobko ta in.]. // Visnyk Khmelnytskoho natsionalnoho universytetu» serii: Tekhnichni nauky. – 2024. – №1 (331). – S. 101–106.
7. An Approach Based on the Visualization Model for the Ukrainian Web Content Classification / [V. Slobodzian, M. Molchanova, O. Kovalchuk ta in.]. // ACIT. – 2022. – S. 400–405.