

ГРАФОВІ НЕЙРОННІ МЕРЕЖІ ТА АЛГОРИТМИ PAGERANK
У ЗАДАЧАХ ПРОГНОЗУВАННЯ ПОПУЛЯРНОСТІ ХЕШТЕГІВ У СОЦІАЛЬНИХ МЕРЕЖАХ

Левицька Т.О.

канд. техн. наук, доцент, ДВНЗ «Приазовський державний технічний університет»,
м. Дніпро, ORCID: <https://orcid.org/0000-0003-3359-1313>, e-mail:
levitskaya_t_a@pstu.edu;

Котихова Л.Д.

асистент, ДВНЗ «Приазовський державний технічний університет», м. Дніпро,
ORCID: <https://orcid.org/0009-0006-5008-622X>, e-mail: kotikhova_l_d@pstu.edu

У статті було досліджено сучасні графові нейронні мережі (GNN, Graph Neural Networks) та алгоритми PageRank як інструменти для вирішення задачі передбачення популярності хештегів у соціальних мережах. Розглянуто можливості й особливості їх застосування. В рамках дослідження було проаналізовано наявні методи та підходи для вирішення цієї задачі й виявлено можливість їх вдосконалення. Актуальність дослідження зумовлена стрімким зростанням обсягів інформації в соціальних мережах та необхідністю ефективного аналізу її поширення. Також було розглянуто поняття хештегів та їх структурні особливості й поведінку в соціальних медіа, представлено модель прогнозування їх популярності в соціальних мережах на основі графових нейронних мереж та алгоритмів PageRank. Проведено аналіз сучасних архітектур GNN, розглянуто етапи та особливості їх роботи. А також було розглянуто модифікації PageRank, зокрема Time-constrained Personalized PageRank. При порівнянні з іншими моделями для вирішення аналогічних завдань було виділено переваги та недоліки пропонованої моделі з використанням графових нейронних мереж в комбінації з алгоритмом PageRank. Перевагою представленої моделі прогнозування є зокрема те, що вона допоможе також врахувати вплив користувачів та їх роль в популяризації контенту, що дозволяє краще зрозуміти динаміку розповсюдження інформації у соціальних мережах. Крім цього наведено результати дослідження, які підтверджують ефективність комбінованого підходу в порівнянні з класичними алгоритмами машинного навчання. Також в дослідженні було виявлено сильні та слабкі сторони пропонованої моделі й наведено практичні рекомендації для використання. Отримані результати можуть бути корисними, наприклад, для вирішення задач контентного аналізу, таргетованої реклами та управління інформаційними потоками в соціальних медіа.

Ключові слова: соціальні мережі, прогнозування, хештег, графові нейронні мережі, PageRank.

Постановка проблеми

Аналіз соціальних мереж – це комплексна галузь, що поєднує в собі знання з соціології, мережевого аналізу, теорії графів, штучного інтелекту, інформаційного пошуку та маркетингу. В його основі лежить вирішення задач, пов'язаних зі структурами, які можна описати за допомогою графів – моделей, що допомагають досліджувати взаємозв'язки між елементами.

Поєднання цих знань разом з сучасними методами аналізу та візуалізації становить основу соціального аналізу, дозволяючи вивчати суспільство через мережеві зв'язки між різними людьми [1].

Разом зі зростаючим інтересом до аналізу соціальних мереж можемо спостерігати і розширення доступу до мережевих даних. Це в свою чергу сприяє тому, що з'являються нові методи аналізу і глибшого їх розуміння. До того ж, результати аналізу соціальних мереж застосовують у різних галузях. Зокрема, особливо важливими вони є в сучасній соціології, економіці, антропології, біології, демографії, комунікації, географії, історії, інформатиці, організаційних дослідженнях, політології, медицині, соціальній психології, лінгвістиці та інших галузях.

База такого аналізу соціальних медіа – зв'язки різного виду (родинні, бізнесові, політичні тощо). Це дозволяє оцінити структури мережі та її внутрішню динаміку. Такі структури можуть як сприяти

економічному розвитку, так і мати загрози. Наприклад, поширення епідемій або фінансових криз через щільні зв'язки між ключовими суб'єктами.

Одне з найважливіших завдань під час аналізу соціальних мереж – прогнозування популярності хештегів та трендів. Хештеги – це багатофункціональні елементи, які передають різноманітні міжособистісні значення в соціальних медіа. Наприклад, допомагають групувати пости у соціальних мережах за певною тематикою. Додавання хештегів до постів стало результатом лінгвістичних інновацій.

Структура хештега впливає на його сприйняття. Особливістю хештегів є те, що вони утворені злиттям слів. А це потребує від читача уваги та концентрації. Тому ефективніше використовувати односкладові хештеги. Вони лаконічні, динамічні й краще впливають на аудиторію.

На думку дослідників, функції хештегів відповідають трьом основним функціям мови, визначеним М. Галлідеєм [2]:

– Концептуальна – передача досвіду мовця про внутрішній і зовнішній світ.

– Міжособистісна – встановлення та підтримка соціальних стосунків, вплив на співрозмовника.

– Текстова – створення зв'язного висловлювання.

Таким чином, при постійному розвитку соціальних мереж та зростанні обсягів інформації

надзвичайно актуальним лишається питання передбачення популярності трендів чи хештегів.

Аналіз останніх досліджень та публікацій

Прогнозування та передбачення загалом відіграє важливу роль в багатьох сферах людської діяльності. Цьому питанню присвячено багато досліджень. Особливо в роботі [3] автори акцентують увагу на важливості застосування наукового підходу для керування соціально-економічними процесами. Автори розглядають побудову моделі для прогнозування популярності публікацій в соціальних мережах. Для цього вони використовують динамічну рекурентну нейронну мережу.

Також в роботі [4] акцентується увага на впливовості та можливостях використання соціальних мереж. Наприклад, розглянуто їх як засіб для громадської самоорганізації.

В роботі [5] автор відзначає, що хоч дослідження Інтернету та соціальних медіа в Україні розпочалося не так давно, але актуальність їх вивчення зростає з кожним днем. В зв'язку з цим автор розглядає різноманітні методи для аналізу даних соціальних медіа в залежності від їх типу.

Також науковці приділяли увагу й дослідженню саме ролі хештегів. Наприклад, в роботі [6] було проведено статистичний аналіз хештегів та розглянуто стилістичні, синтаксичні та статистичні особливості хештегів.

Проаналізовані дослідження приділяли значну увагу ролі користувача в соціальній мережі та його/її впливу на інших користувачів. У сучасній літературі бракує підходів, які б були спрямовані на ідентифікацію впливових користувачів у соціальних мережах. Також недостатньо уваги приділено аналізу рейтингу користувачів щодо їх взаємодії з поширеним контентом.

Мета статті

Метою статті є дослідження можливостей застосування моделей графових нейронних мереж та алгоритмів PageRank для передбачення можливої популярності хештегів в соціальних мережах. Об'єкт дослідження – соціальні мережі, представлені у вигляді графових структур. Предметом дослідження є методи та моделі прогнозування популярності хештегів в соціальних мережах.

Виклад основного матеріалу

Протягом останніх десятиліть дослідники створюють та вдосконалюють різноманітні нейронні мережі для роботи з графовими даними – так звані графові нейронні мережі (Graph Neural Networks, GNN). Наявні досягнення в цій галузі значно підвищили їхню ефективність та здатність до моделювання складних взаємозв'язків.

Сьогодні GNN знаходять практичне застосування у різних сферах, зокрема:

- відкриття нових антибактеріальних засобів;
- моделювання фізичних процесів;
- виявлення фейкових новин;
- прогнозування дорожнього трафіку;
- створення рекомендаційних систем.

Наведені можливі варіанти використання графових нейронних мереж демонструють потужність графових структур у моделюванні складних систем. Також вони відкривають нові можливості для інтеграції GNN в реальні проєкти.

Наприклад, застосування графових нейронних мереж для прогнозування популярності тем і хештегів у соціальних мережах є новим підходом, що надає можливості для моделювання складних взаємодій між користувачами та контентом на графах. Використання GNN дозволяє ефективно виявляти наявні приховані структури та залежності в даних.

Завдяки графовим нейронним мережам можливе більш глибоке розуміння складних взаємодій у соціальних мережах. Вони дозволяють враховувати структуру графа та його атрибути. І це дає більшу точність у порівнянні з традиційними методами. GNN моделюють нелінійні залежності між вузлами і здатні враховувати динаміку змін у мережах. Завдяки цьому вони більш ефективні при прогнозуванні популярності тем і хештегів.

Дослідження свідчать про ефективність GNN: ці моделі забезпечують вищу точність передбачень популярності контенту порівняно з іншими підходами, такими як SVM або Random Forest. Наприклад, GNN демонструють покращену продуктивність при врахуванні динаміки взаємодій між користувачами та контентом для моделювання соціальних графів. В той же час, PageRank добре працює на статичних графах, але його точність нижча в динамічних середовищах. Варто зазначити, що GNN забезпечує високу точність і ефективність у динамічних та складних графових структурах, але водночас вимагає більше обчислювальних ресурсів.

Для підвищення ефективності пропонується застосувати комбінований метод. Комбінація PageRank з GNN дозволяє скористатися сильними сторонами обох підходів: PageRank надає надійний спосіб оцінки важливості вузлів, а GNN враховує складні взаємодії і динамічні зміни у мережі. Це забезпечує більш точні та ефективні результати в порівнянні з використанням кожного методу окремо.

Комбінація PageRank і GNN покращує загальну продуктивність і точність прогнозів, враховуючи як структуру мережі, так і динаміку змін.

Графові нейронні мережі є сучасним підходом до аналізу даних, представлених у вигляді графів. Вони поєднують можливості традиційних нейронних мереж з гнучкістю роботи з графовими структурами, що

дозволяє враховувати як властивості окремих вузлів і ребер, так і структуру графа в цілому.

Графи також можна зустріти й в повсякденному житті. Об'єкти реального світу часто визначаються через їхні взаємозв'язки з іншими. Тому набір таких об'єктів та зв'язків між ними природно моделюється у формі графа.

При використанні графових структур для аналізу соціальних зв'язків графи взаємодій моделюють соціальні зв'язки як набір вузлів (користувачів) і ребер (взаємодій між ними). Основні принципи побудови таких графів включають:

- Вузли (Nodes), які представляють елементи соціальної мережі, такі як окремі користувачі мережі та сторінки. Вузли можуть мати атрибути, такі як ім'я, кількість підписників, активність тощо.

- Ребра (Edges), що відображають взаємодії між різними вузлами, такими як підписки, лайки, коментарі або повідомлення. Ребра можуть бути спрямованими (наприклад, підписка одного користувача на іншого) або не спрямованими (взаємна дружба).

- Ваги ребер (Edge Weights): ребра можуть бути зваженими, де вага відображає інтенсивність або частоту взаємодії (наприклад, кількість повідомлень між певними користувачами).

Виділяють такі типи графів:

- Однорівневі графи: всі вузли мають однаковий тип (наприклад, тільки користувачі).

- Багаторівневі графи: включають вузли різних типів (наприклад, користувачі та публікації).

Основна концепція GNN полягає в тому, що граф G розглядається як вхідна структура:

$$G = (V, E),$$

де V – множина вузлів (nodes),

E – множина ребер (edges), що з'єднують вузли.

Кожен вузол $v \in V$ може мати власні атрибути h_v (наприклад, вектор ознак). Ребра можуть бути спрямованими або не спрямованими, з вагами або без них.

Завдання GNN може полягати в наступному:

- Прогнозування на рівні вузлів: класифікація або регресія для кожного вузла (наприклад, виявлення впливових користувачів у соціальній мережі).

- Прогнозування на рівні ребер: аналіз зв'язків між вузлами (наприклад, рекомендація потенційних друзів або контенту).

- Прогнозування на рівні графа: класифікація графа (наприклад, задача визначення типу соціальної мережі).

Ідея роботи GNN полягає в тому, що GNN використовують принцип агрегації сусідньої інформації: кожен вузол отримує інформацію від своїх сусідів і оновлює власне представлення.

Робота GNN складається з кількох ключових етапів:

1. Ініціалізація вузлів: початкове представлення вузлів задається їх атрибутами (або випадково).

2. Агрегація сусідніх вузлів: кожен вузол збирає інформацію зі своїх сусідів. Типові функції агрегації: Середнє значення (mean), Сума (sum), Максимум (max).

Наприклад, для вузла v :

$$m_v = \text{AGGREGATE}(\{h_u | u \in \text{Neighbors}(v)\}),$$

де h_u – вектор ознак сусіднього вузла.

3. Оновлення вузлів: на основі агрегованої інформації вузол оновлює свій стан:

$$h_v^{(t+1)} = \text{UPDATE}(h_v^{(t)}, m_v),$$

де t – поточний крок ітерації.

4. Численні шари (layers): мережа складається з декількох шарів, що також дозволяє вузлам отримувати інформацію й від найбільш віддалених сусідів.

5. Фінальна проєкція: на цьому етапі представлення вузлів, ребер або графа передається у функцію активації чи класифікації.

Загалом можна виділити основні важливі переваги GNN:

- Гнучкість: підходить для роботи з різними типами графів (статичні, динамічні, багаторівневі).

- Ефективність: інтегрує локальну та глобальну інформацію графа.

- Широкий спектр застосувань: соціальні мережі, біоінформатика, рекомендаційні системи, аналіз знань.

Графові нейронні мережі є інноваційним інструментом для аналізу графових структур. Вони надають широкі можливості для прогнозування соціальних зв'язків, визначення ключових елементів мережі та побудови моделей, що здатні враховувати складну взаємодію між елементами систем [7].

Для оцінки значущості вузлів у графі пропонується використання алгоритму PageRank. Його гнучкість і простота роблять його популярним у багатьох областях, від ранжування вебсторінок до аналізу соціальних мереж. Завдяки можливості адаптації до різних типів даних PageRank залишається одним із найважливіших алгоритмів у сфері графового аналізу [8]. Таким чином впровадження PageRank дозволить провести аналіз впливовості користувачів та хештегів.

Основна ідея алгоритму полягає у визначенні важливості вебсторінок на основі кількості та якості посилань на них [9].

PageRank базується на припущенні, що важливість вузла залежить від кількості вхідних зв'язків і ваги цих зв'язків. Якщо вузол має багато вхідних ребер від важливих вузлів, він сам стає важливим.

Нехай $G = (V, E)$ – граф із вузлами V і ребрами E .

Вагу вузла v визначають за формулою:

$$PR(v) = (1 - d) + d \sum_{u \in \text{In}(v)} \frac{PR(u)}{\text{OutDegree}(u)},$$

де PR – ранг вузла v ;

In (v) – множина вузлів, які мають зв'язки до v ;

OutDegree (u) – кількість вихідних зв'язків вузла u ;

d – коефіцієнт демпінгу (зазвичай $d = 0.85$), який враховує ймовірність переходу на випадковий вузол.

Для обчислення PageRank використовується матрична форма:

$$PR = d * A^T * PR + (1 - d) * 1,$$

де A – нормалізована матриця суміжності графа;

1 – вектор, де кожен елемент дорівнює 1.

Значення PR(v) обчислюються ітераційно, доки не досягнуть збіжності (зазвичай за критерієм зміни менше порогу ϵ).

Вузли, що мають високе значення PageRank, мають більше вхідних зв'язків або пов'язані з іншими значущими вузлами. Це може відображати популярність хештегів й трендів, вплив користувачів у соціальних мережах або вагу інформаційних джерел. Хештеги з найвищими значеннями PageRank є найбільш значущими, і це може свідчити про потенційну можливість зростання їх популярності.

Алгоритм PageRank широко застосовується для виявлення впливових користувачів у соціальних мережах та для ранжування контенту. Він базується на структурі мережі й використовує кількість та якість посилань (взаємодій) між вузлами (користувачами або постами). Проте його недоліки полягають у статичному підході: PageRank не враховує тимчасові аспекти і може бути менш точним в умовах високої динаміки соціальних мереж. Однак існують покращені версії, такі як Time-constrained Personalized PageRank (TPPR), які враховують часову змінність, що дозволяє підвищити точність прогнозування популярності в реальних умовах соціальних мереж [10].

PageRank можна використовувати як одну з ознак при прогнозуванні майбутньої популярності хештегів чи трендів. Комбінація з графовими нейронними мережами (GNN) дозволить поєднати PageRank та ознаки вузлів, що допоможе покращити прогнози.

При використанні графових нейронних мереж для прогнозування для оцінки точності передбачення (регресії) використовуються наступні параметри:

– Середньоквадратична помилка (MSE): міра відхилення прогнозів від реальних значень.

– Середня абсолютна помилка (MAE): середня абсолютна різниця між прогнозами та реальними значеннями.

– Коефіцієнт детермінації (R^2): відображає частку дисперсії, поясненої моделлю.

Етап оцінки містить в собі такі кроки:

1. Тестування моделі: перевірка на тестовому наборі та порівняння прогнозів із фактичними значеннями.

2. Візуалізація: побудова ROC-кривої для класифікації, аналіз залишків (різниця між прогнозами та реальними значеннями) для регресії. При візуалізації результатів пропонується дотримуватись наступних принципів:

– Розмір вузлів пропорційний їх значенню PageRank.

– Колір вузлів може вказувати на категорію чи інші властивості.

3. Аналіз помилок: виявлення вузлів та категорій, для яких модель дає неправильні результати, та виявлення аномальних даних.

GNN надають можливість використовувати структуру зв'язків та індивідуальні характеристики вузлів, що забезпечує значно глибший аналіз й більш точні прогнози.

Для оцінки ефективності запропонованої моделі GNN важливо порівняти її результати з іншими методами прогнозування популярності трендів. У цьому контексті можна порівняти графову нейронну мережу з класичними підходами, такими як:

– Лінійна регресія: може бути використана для прогнозування популярності хештегів (трендів) на основі простих ознак, таких як кількість твітів чи кількість згадувань. Цей підхід є швидким та простим в реалізації. Але він не враховує контексту зв'язків між трендами, що може призвести до зниження точності прогнозів на реальних, складних даних.

– Метод на основі контентного аналізу: використання таких методів, як TF-IDF, дозволяє визначати важливість слів у хештегах. Це надає можливість з'ясувати, які слова чи фрази найбільше пов'язані з популярними трендами. Однак цей підхід має обмеження, бо не враховує динаміку змін в популярності трендів або їх взаємодії.

– Графові методи (без нейронних мереж): алгоритми, які використовують тільки графову структуру (наприклад, алгоритм PageRank чи центральність вузлів), здатні допомогти оцінити важливість окремих хештегів на основі їх зв'язків в графі. Такі методи можуть давати хороші результати при визначенні ключових трендів, якщо взаємодія між ними має великий вплив. Проте такі підходи не мають можливості глибокого навчання, тому не здатні захопити складні залежності.

– Графові нейронні мережі (GNN): забезпечують більш глибоке розуміння складних взаємодій між хештегами та трендами. Вони не тільки оцінюють індивідуальну популярність хештегів (трендів), а й враховують, як вони пов'язані між собою. Це дозволяє моделі швидко адаптуватися до змін у мережі та передбачати взаємний вплив трендів. Недолік підходу полягає в тому, що графові нейронні мережі потребують значних обчислювальних ресурсів. Також вони можуть

бути складними для налаштування, якщо дані мають низьку якість або містять багато шуму.

В межах дослідження було також виявлено сильні та слабкі сторони пропонованої моделі.

Сильні сторони:

- Робота з графовими структурами: GNN ефективно працюють з даними у вигляді графів, адже враховують зв'язки між різними трендами. Це дозволяє моделі вивчити, як хештеги та тренди взаємодіють між собою, що є важливим для точного прогнозування в динамічних соціальних мережах.

- Гнучкість і адаптивність: GNN здатні адаптуватися до змін у мережах та вивчати складні нелінійні залежності між трендами. Це дозволяє точніше прогнозувати їх майбутню популярність.

- Можливість обробки великих даних: графові нейронні мережі здатні обробляти великі об'єми даних з багатьма зв'язками. Це робить їх потужним інструментом для прогнозування популярності хештегів та трендів й на великих платформах, таких як Twitter чи YouTube.

- Передбачення складних залежностей: модель здатна прогнозувати не лише індивідуальні зміни популярності трендів, а також й їх взаємний вплив у мережі.

Слабкі сторони:

- Необхідність в обчислювальних ресурсах: GNN є складними для обчислювання, а для їх навчання необхідні значні ресурси. Особливих ресурсів потребують великі графи з численними зв'язками між вузлами.

- Залежність від якості даних: GNN є чутливими до якості вхідних даних. Якщо дані про хештеги не є достатньо точними або містять багато шуму (наприклад, відсутність точних даних про популярність), то це може призвести до зниження точності прогнозів.

- Потрібен великий обсяг даних для навчання: для того, щоб навчити модель правильно розпізнавати залежності, потрібно мати великий обсяг даних. Для малих наборів даних GNN можуть не дати значущих результатів.

- Можливі проблеми при інтерпретації результатів: графові нейронні мережі можуть бути достатньо складними для коректної інтерпретації. Проблеми можуть виникнути в ситуаціях, де важливо розуміти, чому модель зробила той чи інший прогноз. Це може заважати належному поясненню результатів прогнозу користувачам.

Але незважаючи на наявність слабких сторін, запропонована модель GNN демонструє хороші результати при прогнозуванні популярності хештегів або трендів. Особливо результативним є використання моделі в тих ситуаціях, коли наявні складні залежності між трендами. Однак ефективність пропонованої моделі значним чином залежить від якості вхідних даних та наявності достатніх обчислювальних ресурсів для навчання на великих графах.

Висновки

В роботі було розглянуто останні дослідження та методи для вирішення задачі прогнозування популярності хештегів в соціальних мережах. Після цього було запропоновано модель з використанням графових нейронних мереж та алгоритму PageRank. Важливим внеском цієї роботи є досягнення ранжування користувачів на основі поширення впливу в соціальних мережах.

Також у процесі дослідження було зроблено порівняння з іншими моделями нейронних мереж, де розроблена модель надає більш високу точність передбачення ніж інші.

Загалом при використанні графових нейронних мереж в динамічних соціальних мережах вони мають переваги над іншими розглянутими методами. А використання алгоритмів PageRank допомагає підвищити інтерпретованість результатів.

Отримані в результаті дослідження висновки мають практичне значення для використання в інформаційній безпеці та маркетинговій аналітиці. Так практичні рекомендації для соціальних мереж та маркетологів, наприклад, можуть складатись з декількох ключових напрямків. По-перше, для аналізу поточних популярних хештегів та трендів можна використовувати модель для виявлення актуальних тем, які активно обговорюються. Наприклад, у мережі Twitter доцільно буде виділяти ключові хештеги та аналізувати їхні зв'язки. Це дозволить розділити аудиторію за інтересами й пропонувати користувачам персоналізований контент. Ще одним важливим завданням є моніторинг впливових вузлів та впливових користувачів. PageRank або інші графові метрики можуть допомогти визначити тих користувачів, які найбільше впливають на розповсюдження інформації в мережі. Спільна робота маркетологів з такими користувачами допоможе підтримувати та просувати тренди, що може сприяти просуванню брендів.

Перелік використаних джерел

- [1] Mercer S. Social Network Analysis and Complex Dynamic Systems. *Motivational Dynamics in Language Learning*. Multilingual Matters, 2014. Pp. 73-82. DOI: <https://doi.org/10.21832/9781783092574-010>.
- [2] Halliday M. A. K., Hasan R. Language, context, and text: aspects of language in a social-semiotic perspective. Oxford : Oxford University Press, 1985. 126 p.
- [3] Прогнозування поведінки об'єктів соціальних мереж / І. Івченко та ін. *Вісник Хмельницького національного університету. Серія: Технічні науки*. 2024. № 341(5). С. 317-321. DOI: <https://doi.org/10.31891/2307-5732-2024-341-5-46>.
- [4] Руденко А. Соціальні мережі як ефективний засіб громадської самоорганізації в сучасній Україні. Київ, 2018. 17 с. URL: <https://niss.gov.ua/sites/default/files/2018->

- [05/Rudenko-52ea7.pdf](#) (дата звернення: 21.08.2025).
- [5] Туманов О. О. Статистичні методи аналізу даних соціальних медіа. *Бізнесінформ*. 2020. № 2. С. 266-272. DOI: <https://doi.org/10.32983/2222-4459-2020-2-266-272>.
- [6] Рябова К. Статистичний аналіз хештегів (на матеріалі соціальних мереж Twitter, Instagram і Google). URL: https://kamts1.kpi.ua/sites/default/files/files/riabova_statustuchnuv.pdf (дата звернення: 17.07.2025).
- [7] Hashtag recommendation for enhancing the popularity of social media posts / Chakrabarti P., Malvi E., Bansal S., Kumar N. *Social network analysis and mining*. 2023. № 13(1). Article 21. DOI: <https://doi.org/10.1007/s13278-023-01024-9>.
- [8] A Gentle Introduction to Graph Neural Networks / Sanchez-Lengeling B., Reif E., Pearce A., Wiltchko A. B. *Distill*. 2021. DOI: <https://doi.org/10.23915/distill.00033>.
- [9] Heavey A., Karagöz E. Analysing Credibility of Twitter Users Using the PageRank Algorithm. Bachelor's Thesis in Computer Science. Stockholm, 2017. 27 p.
- [10] Роль хештегов у соціальних мережах. URL: <https://webnauts.pro/uk/blog/rol-heshtegov-u-soczialnyh-merezhah/> (дата звернення: 28.08.2025).

GRAPH NEURAL NETWORKS AND PAGERANK ALGORITHMS IN THE PROBLEMS OF PREDICTING THE POPULARITY OF HASHTAGS IN SOCIAL NETWORKS

Levytska T.O.

PhD (Engineering), associate professor, SHEI «Priazovskyi state technical university», Dnipro, ORCID: <https://orcid.org/0000-0003-3359-1313>, e-mail: levitskaya_t_a@pstu.edu;

Kotykova L.D.

assistant, SHEI «Priazovskyi state technical university», Dnipro, ORCID: <https://orcid.org/0009-0006-5008-622X>, e-mail: kotykova_l_d@pstu.edu

The article explores modern Graph Neural Networks (GNN) and PageRank algorithms as tools for solving the task of predicting the popularity of hashtags in social media. It examines the capabilities and specific features of their application. As part of the study, existing methods and approaches to this problem were analyzed, revealing opportunities for improvement. The relevance of the research is driven by the rapid growth of information volumes in social networks and the need for effective analysis of its dissemination. The concept of hashtags, their structural characteristics, and behaviour in social media are also discussed. A predictive model based on Graph Neural Networks and PageRank algorithms is presented. An analysis of current GNN architectures has been carried out, and the stages and features of their work have been reviewed. Modifications of PageRank, particularly Time-constrained Personalized PageRank, are also considered. In comparison with other models designed for similar tasks, the advantages and disadvantages of the proposed approach – combining GNNs with PageRank – are highlighted. A key advantage of the presented predicting model is that it will also help to account for user influence and their role in the popularization of content, which allows us also to understand better the dynamics of the distribution of information in social networks. Additionally, the research presents results that confirm the effectiveness of the combined approach compared to classical machine learning algorithms. The study also identifies the strengths and weaknesses of the proposed model and provides practical recommendations for its application. The results may be useful, for example, for tasks such as content analysis, targeted advertising, and managing information flows in social media.

Keywords: social networks, prediction, hashtag, graph neural networks, PageRank.

References

- [1] S. Mercer, "Social Network Analysis and Complex Dynamic Systems," in *Motivational Dynamics in Language Learning*, Multilingual Matters Publ., 2014, pp. 73-82. doi: 10.21832/9781783092574-010.
- [2] M.A.K. Halliday, and R. Hasan, *Language, context, and text: aspects of language in a social-semiotic perspective*. Oxford : Oxford University Press, 1985.
- [3] I. Ivchenko, O. Haidaienko, N. Knyrik, H. Morozova, and D. Grybeniuk, "Prohnozuvannia povedinky ob'ektiv sotsialnykh merezh" ["Forecasting the behavior of social media objects"], *Visnyk Khmelnytskoho natsionalnoho universy-tetu. Seriya: Tekhnichni nauky – Herald of Khmelnytskyi National University. Technical sciences*, № 341(5), pp. 317-321, 2024. doi: 10.31891/2307-5732-2024-341-5-46. (Ukr.)
- [4] A. Rudenko, *Sotsialni merezhi yak efektyvnyi zasib hromadskoi samoorhanizatsii v suchasnii Ukraini* [Social networks as an effective means of community self-organization in modern Ukraine]. Kyiv, Ukraine, 2018. [Online]. Available: <https://niss.gov.ua/sites/default/files/2018-05/Rudenko-52ea7.pdf>. Accessed on: August 21, 2025.
- [5] O.O. Tumanov, "Statystychni metody analizu danykh sotsialnykh media" ["The Statistical Methods for Analyzing Social Media Data"], *Biznesinform– Business Inform*, № 2, pp. 266-272, 2020. doi: 10.32983/2222-4459-2020-2-266-272. (Ukr.)

- [6] K. Riabova, Statystychnyi analiz kheshtehiv (na materiali sotsialnykh merezh Twitter, Instagram i Google) (Statistical analysis of hashtags (based on social networks Twitter, Instagram and Google)). [Online]. Available: https://kamts1.kpi.ua/sites/default/files/files/riabova_statustuchnuy.pdf. Accessed on: July 17, 2025. (Ukr.)
- [7] P. Chakrabarti, E. Malvi, S. Bansal, and N. Kumar, "Hashtag recommendation for enhancing the popularity of social media posts," *Social network analysis and mining*, № 13(1), article 21, 2023. doi: 10.1007/s13278-023-01024-9.
- [8] B. Sanchez-Lengeling, E. Reif, A. Pearce, and A.B. Wilschko, "A Gentle Introduction to Graph Neural Networks," *Distill*, 2021. doi: 10.23915/distill.00033.
- [9] A. Heavey, and E. Karagöz, "Analysing Credibility of Twitter Users Using the PageRank Algorithm," Bachelor's Thesis in Computer Science, Stockholm, 2017.
- [10] Rol kheshtehov u sotsialnykh merezhakh (The role of hashtags in social networks). [Online]. Available: <https://webnauts.pro/uk/blog/rol-heshtegov-u-soczi-alnyh-merezhah/>. Accessed on: August 28, 2025. (Ukr.)

Стаття надійшла 09.09.2025

Стаття прийнята 28.09.2025

Стаття опублікована 30.10.2025

Цитуйте цю статтю як: Левицька Т. О., Котихова Л. Д. Графові нейронні мережі та алгоритми PageRank у задачах прогнозування популярності хештегів у соціальних мережах. *Вісник Приазовського державного технічного університету. Серія: Технічні науки*. 2025. Вип. 51. С. 50-56. DOI: <https://doi.org/10.31498/2225-6733.51.2025.344600>.