

# КІБЕРНЕТИКА та КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

УДК 519.67

DOI:10.34229/2707-451X.25.1.8

В.В. ТРЕТИНИК, Ю.А. НАДЬ

## ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ РОЗРОБКИ СИСТЕМИ СОЦІАЛЬНОЇ ДИНАМІКИ

**Вступ.** Сучасний світ характеризується швидкими змінами і частими критичними ситуаціями. Людство стикається із все більш складними викликами, такими як пандемії, війни, що можуть призвести до соціальної напруженості. Мати засоби для моніторингу і виміру напруженості дозволяє як суспільству, так і державі, реагувати вчасно та адекватно на ці виклики. Напруженість призводить до соціальних конфліктів, політичних криз, економічних спадів. Мати систему для виміру напруженості допомагає розуміти, які чинники сприяють напруженості, і вживати заходів для попередження чи пом'якшення наслідків. Тож актуально розробляти інструменти для виміру напруженості суспільства, оскільки це є важливим кроком у напрямку розуміння та управління соціальною динамікою.

У роботі [1] автори використовують методи аналізу тональності, такі як SVM (Support Vector Machines) та нейронні мережі (LSTM), для моніторингу настроїв у соціальних мережах. У роботі [2] застосовано модель BERT для автоматичної класифікації текстів із метою виявлення мови ворожнечі. У статті [3] застосовано Isolation Forest для виявлення аномальних патернів у соціальній активності. За допомогою моделі машинного навчання на основі понад 340 індикаторів, що охоплюють широкий спектр макрофінансових та соціально-економічних змінних у роботі [4] оцінюється ризик соціальних заворушень. У роботі [5] запропонована система для визначення рівня напруженості в суспільстві.

Мета даної роботи – це застосування методів машинного навчання та нейролінгвістичного програмування для задачі аналізу думок користувачів інтернет для передбачення соціальної напруженості у суспільстві.

Передбачення соціальної напруженості є комплексним завданням, для якого не існує однієї універсальної методології, що б забезпечувала точні результати. Виділимо два найпоширеніших підходи:

- використання соціологічних досліджень, опитувань та інтерв'ю з громадянами;
- використання аналізу соціальних медіа та інших джерел даних з Інтернету.

*У даній роботі розглянуто застосування методів нейролінгвістичного програмування, моделей векторного представлення слів у комбінації із алгоритмом кластеризації для створення системи визначення соціальної напруженості через аналіз коментарів Youtube каналів. Для формалізації структури системи, було використано удосконалений бізнес профіль Еріксона – Пенкера. Розроблено програмне забезпечення для задачі аналізу думок користувачів Інтернет для передбачення соціальної напруженості у суспільстві.*

**Ключові слова:** аналіз напруженості, соціальна динаміка, машинне навчання, обробка текстових даних, word2vec, k-means.

© В.В. Третиник, Ю.А. Надь, 2025

Другий підхід базується на зборі й аналізі великої кількості соціальних медіа-публікацій, повідомлень у блогах, коментарів тощо, щоб виявити ознаки соціальної напруженості. Алгоритми машинного навчання та природної мови застосовуються для автоматичного аналізу та класифікації цих даних з метою виявлення трендів та прогнозування можливих конфліктів чи напруженості. Цей метод має великі перспективи, адже він є альтернативою для повного вирішення недоліків безпосередньої взаємодії із опитуваними, як у першому підході.

У даній роботі для аналізу задач соціальної динаміки запропоновано скомбінувати підхід із застосуванням векторного представлення слів та модель кластеризації, щоб найточніше задовольнити потреби розроблюваної програми, яка оперує відкритими непоміченими текстовими даними українською мовою.

Розглянемо деякі існуючі комерційні рішення для аналізу соціальної напруженості. Datamint – платформа із аналізом соціальних медіа, новин та інших джерел у реальному часі для виявлення важливих подій та трендів [6]. Платформа надає користувачам можливість налаштувати інформаційні потоки, відстежувати ключові події та отримувати сповіщення про важливі зміни. Основні сфери використання Datamint – області кризового менеджменту, фінансової аналітики, безпеки, медіа-моніторингу та громадської безпеки.

Один з голландських проєктів, пов'язаних з аналізом напруженості суспільства, проєкт "Crisis Monitor" [7]. Він був розроблений у співпраці з урядовими та науковими організаціями Нідерландів з метою виявлення, моніторингу та аналізу суспільної напруженості та кризових ситуацій. Проєкт був завершений у 2019 році й використовував різні методи аналізу тексту, машинного навчання та інтелектуального аналізу даних для виявлення змін у настрої суспільства, спричинених різними факторами, такими як економічна криза, політичні події, природні катастрофи тощо.

Brandwatch – одна з провідних платформ для соціального медіа-аналізу та моніторингу громадської думки [8]. Її основний продукт – обслуговування проєктів клієнтів у контексті реакцій споживачів. Вона надає різноманітні інструменти та функції для збору, аналізу та інтерпретації даних з соціальних медіа та інших джерел.

При порівнянні комерційних рішень було виявлено, що критичні параметри у вигляді вартості та регіону аналізу задовольняють запити лише меншості користувачів Інтернет. Так, безкоштовну інформацію щодо напруженості суспільства можна отримати лише для Нідерландів у період до 2019 року, а для інших потреб варто звертатись до персоналу компаній «Datamint» та «Brandwatch» із комерційними пропозиціями. Подібна нестача якісного та доступного програмного забезпечення пропонує можливість для розвитку адаптованих систем визначення напруженості у суспільстві із регіональним фокусом.

Для формалізації структури системи, було використано удосконалений бізнес профіль Еріксона – Пенкера, що має наступний вигляд (рис. 1):

Опишемо елементи, наведені на діаграмі:

Проблема роботи. Необхідність аналізувати реакцію українців та тим самим оцінювати моральний стан суспільства.

*Мета:* розроблення інструмента для оцінки настрою українського тексту.

*Процеси:*

- завантаження текстових даних з коментарів під різними україномовними YouTube каналами;
- первинна обробка тексту;
- представлення даних у векторному вигляді;
- навчання моделі кластеризатора;
- використання кластерів настрою для класифікації сентименту тексту;
- збереження результуючих даних;
- інтерфейс користувача.

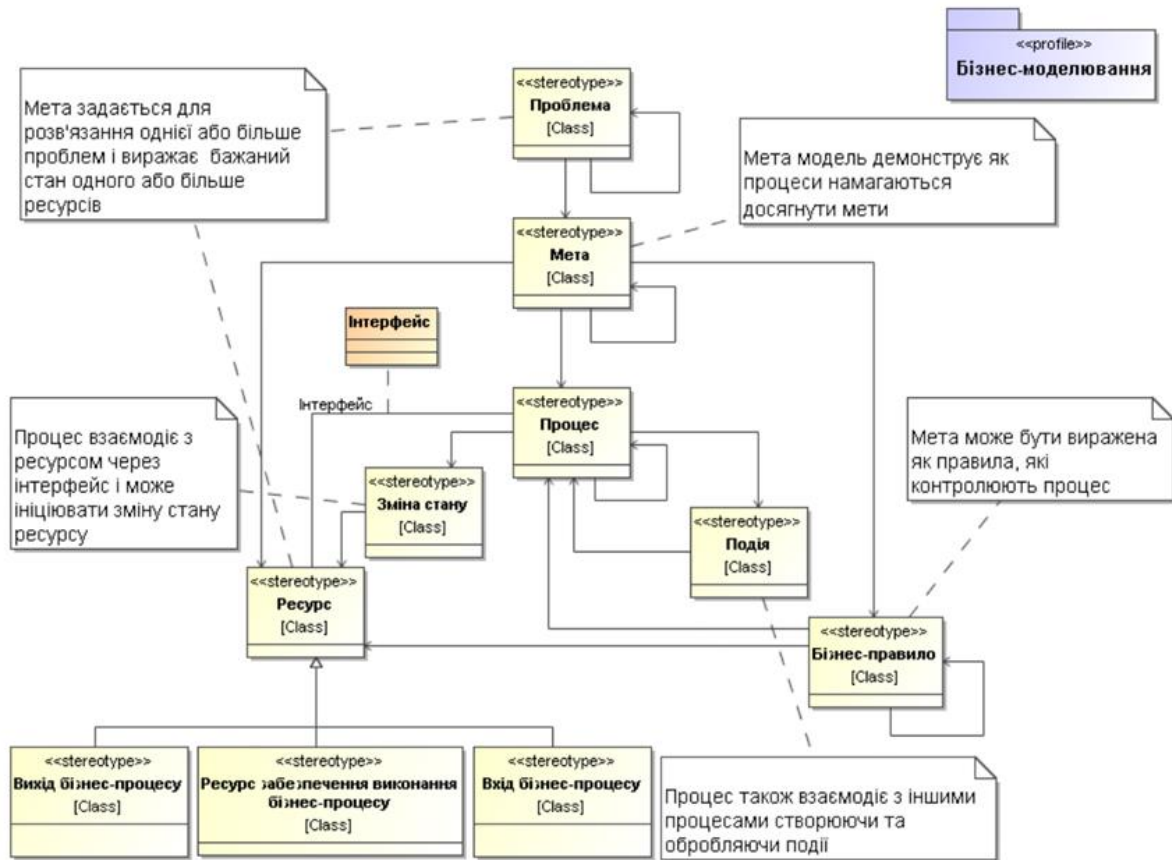


РИС. 1. Удосконалений бізнес-профіль Еріксона – Пенкера

*Зміни стану:*

- необроблений текст -> оброблений текст;
- ініціалізована модель векторного представлення -> навчена модель векторного представлення;
- оброблений текст -> векторизований текст;
- ініціалізована модель кластеризації -> навчена модель кластеризації;
- векторизований текст -> розбиття векторів на кластери;
- кластеризовані вектори -> визначений настрій тексту.

*Ресурси:*

- вихід бізнес-процесу: навчена модель, середній коефіцієнт настрою коментарів до відео;
- ресурс забезпечення виконання бізнес-процесу: оброблені методами НЛП коментарі з-під відео, модель векторного огортання, модель кластеризації;
- вхід бізнес-процесу: необроблений текст з коментарів, ініціалізована модель.

*Події:*

- поява нової ефективнішої версії моделі під час процесу навчання, що вплине на результати отримані у процесі аналізу настрою з коментарів;
- зміна навчального набору даних, що вплине на результати роботи програми.

*Бізнес правила:*

BR1 – швидкість обробки коментаря не перевищує 0.1с;

BR2 – коментарі у вигляді емоjіs обробляються;  
 BR3 – коментарі іноземними мовами аналізуються;  
 BR4 – набір навчальних даних має містити збалансовану кількість коментарів на кожну дату;  
 BR5 – кількість іноземних коментарів для визначення настрою у конкретну дату не має перевищувати 20 % усіх коментарів.

Для детальнішого опису взаємодії інтерфейсів наведемо діаграму діяльності з водними доріжками (рис. 2).

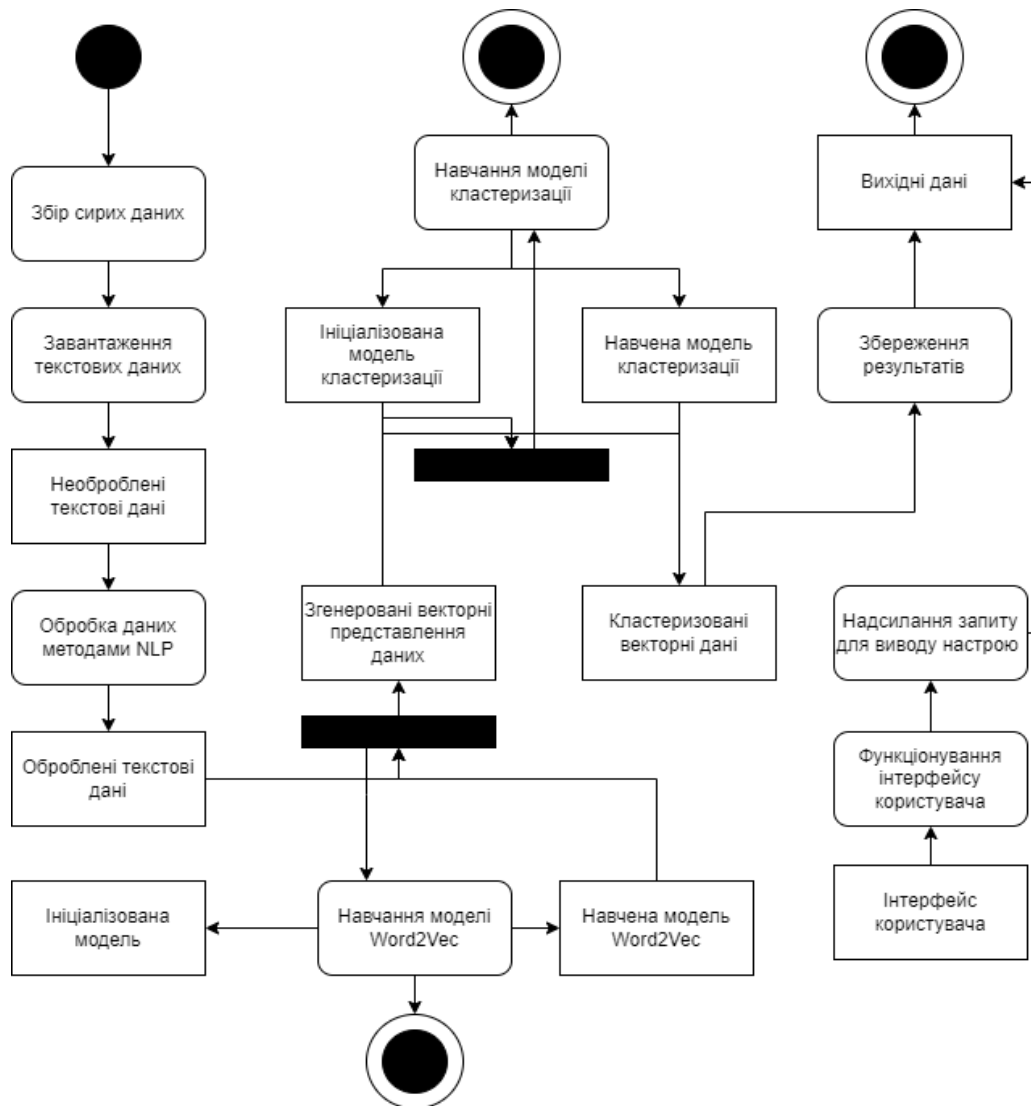


РИС. 2. Діаграма діяльності першого рівня у нотатції UML

Оскільки дані до системи поступатимуть у сирому вигляді, тому для вдалого огортання слів у вектори та подальшу кластеризацію за настроєм використаємо методи нейролінгвістичного програмування, а саме:

- токенизація даних;
- лематизація даних;

- видалення стоп-слів;
- визначення мови (використана вбудована бібліотека Python langdetect).

Дані трьох категорій збиралися за допомогою YouTube API:

- дані щодо каналу: id каналу, id відео, що там опубліковані, назви відео;
  - дані щодо відео: id відео, кількість коментарів, кількість лайків, кількість переглядів дата публікації;
  - дані щодо коментаря: id коментаря, текст коментаря, дата публікації коментаря, ім'я автора.
- На рис. 3 показано приклад оброблених даних.

|    | comment_id                  | video_id   | text_processed                                 | author_name             | comment_published_date | original_language |
|----|-----------------------------|------------|--|-------------------------|------------------------|-------------------|
| 9  | UgzVdWJaG02wXnryNEp4AaABAg  | Oi8LqRgBJA | депутат одеський область барвіненко            | Олег                    | 2023-05-14T06:43:20Z   | uk                |
| 10 | UgxSH6t--dodFxc8bdB4AaABAg  | Oi8LqRgBJA | косити німецький фашист російський те...       | Sufo                    | 2023-05-14T04:51:18Z   | uk                |
| 11 | Ugz8FomYETBRV-NeKJV4AaABAg  | Oi8LqRgBJA | півтори година вздумайте зменшувати ...        | Yuliya Subotenko        | 2023-05-14T04:32:05Z   | uk                |
| 12 | Ugw4bFXNIVgHV6ErV7x4AaABAg  | Oi8LqRgBJA | yellow_heart face_blowing_a_kiss               | Oksana Rotsheld         | 2023-05-13T22:05:32Z   | en                |
| 13 | Ugy7axL2aCIAAD82_14AaABAg   | Oi8LqRgBJA | прет друг топ                                  | taras karakos           | 2023-05-13T17:25:20Z   | mk                |
| 14 | UgzkSA0ZI4hnG9iSlt4AaABAg   | Oi8LqRgBJA |  | Dmytro Dmytrovych       | 2023-05-13T17:03:05Z   | no                |
| 15 | UgyjFpjOLKGYcP0AG_N4AaABAg  | Oi8LqRgBJA | лакучая кива                                   | Eugene Sychuk           | 2023-05-13T14:19:03Z   | ru                |
| 16 | UgwDFMvhSvCA-966Pup4AaABAg  | Oi8LqRgBJA | ахаха хлопець гумор дотепний                   | Валентина Кутепова      | 2023-05-13T13:36:11Z   | uk                |
| 17 | UgzI7Gg9AOU9upyDowB4AaABAg  | Oi8LqRgBJA | луценко шизофрения верит что придумал          | Татуличкин Александр    | 2023-05-13T12:45:23Z   | bg                |
| 18 | Ugz6HKREFtgnFvkiVYF4AaABAg  | Oi8LqRgBJA | face_with_tears_of_joy face_with_tears_of_j... | Алена Яцела             | 2023-05-13T11:07:28Z   | en                |
| 19 | Ugxw-fifUNixeorT9ZB4AaABAg  | Oi8LqRgBJA | момент типовий маячний арика береза            | Spirid 10               | 2023-05-13T10:57:27Z   | uk                |
| 20 | UgxgFW12VZpBi5MRGXN4AaABAg  | Oi8LqRgBJA | ксп контрольно спостережний пост ком...        | Dhdhdhdhd Hdhhdhdbdbd   | 2023-05-13T10:36:38Z   | ru                |
| 21 | Ugx1EePBVhpPyyYr-94AaABAg   | Oi8LqRgBJA | луценко думати вистачити запитати кол...       | Dhdhdhdhd Hdhhdhdbdbd   | 2023-05-13T10:33:52Z   | uk                |
| 22 | UgxbDQRyZ8wyLhnFCgB4AaABAg  | Oi8LqRgBJA | ласка зробити футболка білий колір ріж ...     | Олексій Ав'юцький       | 2023-05-13T09:59:09Z   | uk                |
| 23 | UgwlrM8EblaK22bLmdI4AaABAg  | Oi8LqRgBJA | дякувати                                       | Олексій Ав'юцький       | 2023-05-13T09:56:56Z   | uk                |
| 24 | UgyOGeLMKi8RCTnDaFJ4AaABAg  | Oi8LqRgBJA | півтори година сладости жизни                  | Roman Dutchak           | 2023-05-13T09:16:24Z   | uk                |
| 25 | UgxAfyU0s48mjg27mCB4AaABAg  | Oi8LqRgBJA | півтори година сладости жизни                  | Roman Dutchak           | 2023-05-13T09:15:57Z   | uk                |
| 26 | UgytHjAcQlncldZ3qB4AaABAg   | Oi8LqRgBJA | розрив бомба кива істеричка поїхавшою...       | Гайкин. Лимонов         | 2023-05-13T09:13:20Z   | uk                |
| 27 | Ugx5i7I2bjSQMfL0pCx4AaABAg  | Oi8LqRgBJA | ефір чаплигою бомба grinning_squinting...      | Наталія                 | 2023-05-13T08:41:39Z   | en                |
| 28 | UgyFLngxMbsfeivfDB4AaABAg   | Oi8LqRgBJA | заставці луценко костюмі піти воювати ...      | Volodymyr Zakolodyazhny | 2023-05-13T08:08:35Z   | uk                |
| 29 | Uqww-VV-ne7HF3TaRo3x4AaABAg | Oi8LqRgBJA | додати кон передача цікавіший grinning ...     | Illia Diliqodin         | 2023-05-13T07:42:32Z   | uk                |

РИС. 3. Приклад оброблених даних

В роботі налаштовані ключові гіперпараметри моделі *word2vec* (розмір вектора, розмір контекстного вікна, мінімальна кількість слів) та методу кластеризації *k-means* (кількість кластерів, метод вибору початкових центроїдів та метрика відстані).

Проведено кластеризацію та присвоєні значення сентименту словам. Категорії слів, залежно від кластерів, будуть наступними (рис. 4):

- негативне емоційне забарвлення: кластер № 2;
- нейтральне емоційне забарвлення: кластер № 3, кластер № 1;
- позитивне емоційне забарвлення: кластер № 0.

Напруженість суспільства напряму відповідатиме кількості негативних коментарів. Оцінка емоційного забарвлення речення заокруглюється до 1(позитивна), 0(негативна) та 0.5 (нейтральна) в залежності від кількості позитивно/негативно/нейтрально-визначених слів.

Розглянемо дані у масштабі днів. Користувач отримує на графіку середнє значення настрою по дням (де -1 немає соціальної напруженості, +1 висока соціальна напруженість) (рис. 5). На графіку позначено середнє значення напруженості за проміжок часу, поденна напруженість та згладжений середній показник, що показує середнє значення у проміжку 7-ми днів. За графіком видно, що соціальна напруженість має високі показники у грудні 2021 року та починаючи з другої половини лютого 2022 року. Останні піки характеризуються початком повномасштабного вторгнення, проте пік у грудні 2021 – менш-очевидний.

|     | words       | vectors   | cluster |
|-----|-------------|---|---------|
| 228 | сподіватися | [-1.73448548e-02 1.14627900e-02 5.58654182...   | 0       |
| 229 | московський | [-5.95676852e-03 9.36023742e-02 7.29537308...   | 2       |
| 230 | кацап       | [-0.03218418 0.10035323 0.07897378 0.01465...   | 2       |
| 231 | прекрасний  | [0.02278283 0.052255 -0.05809831 0.045017...    | 0       |
| 232 | мати        | [-0.09669054 0.08858611 0.13569699 -0.16536...  | 1       |
| 233 | співати     | [-4.73527387e-02 6.06367588e-02 6.15637451...   | 0       |
| 234 | винний      | [-0.0052504 -0.06900981 0.10649006 -0.00417...  | 2       |
| 235 | випадок     | [-0.09018847 0.02358929 0.12478289 -0.14740...  | 3       |
| 236 | приклад     | [-0.06608076 -0.12355306 0.09856348 -0.08377... | 3       |
| 237 | crying_face | [-0.0683188 -0.0158913 0.04481011 0.106441...   | 2       |
| 238 | повага      | [0.07517032 -0.07795046 0.05134587 0.01117...   | 0       |
| 239 | вдячний     | [0.00821893 -0.03159884 0.01045158 0.04558...   | 0       |

РИС. 4. Присвоєні словам значення сентименту після кластеризації

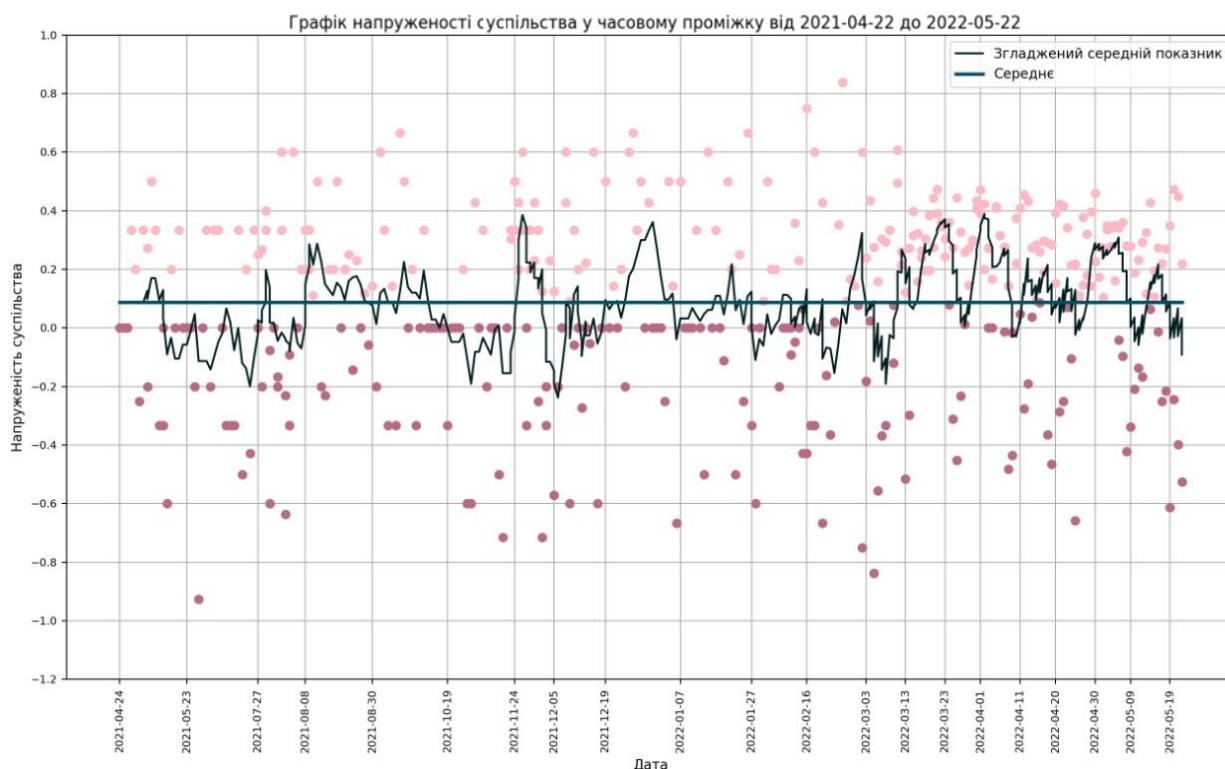


РИС. 5. Графік настрою у проміжку заданому користувачем

**Висновок.** У роботі розроблено архітектуру та програмне забезпечення системи соціальної динаміки на основі методів машинного навчання. Проведено розбиття на чотири модулі: обробка текстових даних, тренування моделі Word2Vec, тренування моделі K-Means та користувацький інтерфейс; проведено навчання моделей із різними гіперпараметрами, налаштованими вручну. Графік напруженості суспільства показує тенденції соціальної динаміки українців. На основі нього можна виділити час, за який нація адаптується до негативних сценаріїв та шаблон поведінки, притаманний

великому прошарку суспільства в контексті зацікавленості у негативних новинах та тривалості їх впливу. Такі характеристики особливо важливі для визначення подальшої поведінки соціуму та передбачення тенденції напруженості у майбутньому.

**Авторські внески:** Третиник В.В. – концепція, структура, аналіз даних, написання вступної частини та висновків; Надь Ю.А. – збір даних, візуалізація даних, розробка архітектури та програмного забезпечення системи соціальної динаміки на основі методів машинного навчання.

**Примітка.** Для отримання даних були використані такі відеоматеріали:

Жданов – <https://www.youtube.com/channel/UC111NXIcDs0VGfyre6EiPmA>;  
 STERNENKO – [https://www.youtube.com/channel/UC5HBd4I\\_kpba5b0O1pK-Bfg](https://www.youtube.com/channel/UC5HBd4I_kpba5b0O1pK-Bfg);  
 Роман Ратушний – [https://www.youtube.com/channel/UCB-1Cgjgn\\_xw1-g9m2GF6jQ](https://www.youtube.com/channel/UCB-1Cgjgn_xw1-g9m2GF6jQ);  
 Цимбалюк – <https://www.youtube.com/channel/UCB9WOAf7gaCLiFnB2R46Tlw>;  
 Gordonua – <https://www.youtube.com/channel/UCCcprrrcbdaj14kYPjcbj9w>;  
 Палає – <https://www.youtube.com/channel/UCCnxINydeDs-74iVMRU-Qw>;  
 Utoronto – [https://www.youtube.com/channel/UCF\\_ZiWz2Vcq1o5u5i1TT3Kw](https://www.youtube.com/channel/UCF_ZiWz2Vcq1o5u5i1TT3Kw);  
 Ragulivna – <https://www.youtube.com/channel/UCf4A8MGpasfQTa28WncfreQ>;  
 ISLNDTV – <https://www.youtube.com/channel/UCFsMTNtpoT5hvyNK1IUxqAQ>;  
 Arestovych – <https://www.youtube.com/channel/UCjWY2g76QZf7QLEwx4cB46g>;  
 ЄПИТАННЯ – <https://www.youtube.com/channel/UCLzoT0z2WwyDUAx8LHYzDJw>;  
 DeepStateUA – <https://www.youtube.com/channel/UCMGRXVYg1wMA6e4zTGX-DwQ>;  
 MarkusUA – [https://www.youtube.com/channel/UCw9EjcSur6DF\\_hDU5uD8tRQ](https://www.youtube.com/channel/UCw9EjcSur6DF_hDU5uD8tRQ).

#### Список літератури

1. Harshali P.P., Atique M. Sentiment Analysis for Social Media: A Survey. *2nd International Conference on Information Science and Security (ICISS)*, 14-16 December 2015, Seoul, Korea (South). <https://doi.org/10.1109/ICIS-SEC.2015.7371033>
2. Kapil P., Ekbal A., Das D. Investigating Deep Learning Approaches for Hate Speech Detection in Social Media. *International Conference on Computational Linguistics and Intelligent Text Processing*, 2019. <https://doi.org/10.48550/arXiv.2005.14690>
3. Macis L., Tagliapietra M., Meo R., Pisano P. Breaking the trend: Anomaly detection models for early warning of socio-political unrest. *Technological Forecasting and Social Change*. Vol. 206. 2024.
4. Redl C., Hlatshwayo S. Forecasting Social Unrest: A Machine Learning Approach. *International Monetary Fund*. 2021. Vol. 2021, Iss. 263. P. 29. <https://doi.org/10.5089/9781557758873.001>
5. Shchokoliev M., Tretynik V. The System of Determining the Overall Estimate of the Level of Tension in Society. *Digital Transformation, Cyber Security and Resilience of Modern Societies, Studies in Big Data*. Vol. 84. P. 475–485. <https://doi.org/10.1007/978-3-030-65722-2>
6. Dataminr: About us. <https://www.dataminr.com/about> (звернення: 30.01.2025)
7. Crisis Monitor. <https://www.cbs.nl/en-gb/about-us/innovation/project/social-tensions-indicator-gauging-society> (звернення: 30.01.2025)
8. Brandwatch Overview. <https://www.brandwatch.com/> (звернення: 30.01.2025)

Одержано 30.01.2025

**Третиник Віолета Вікентіївна,**

кандидат фізико-математичних наук,  
 доцент кафедри прикладної математики НТУУ «КПІ ім. Ігоря Сікорського»

<https://orcid.org/0000-0002-3538-8207>

[viola.tret@gmail.com](mailto:viola.tret@gmail.com)

**Надь Юлія Артурівна,**

розробник  
 GXperts GmbH, Відень.

УДК 519.67

В.В. Третиник<sup>1\*</sup>, Ю.А. Надь<sup>2</sup>

## Використання методів машинного навчання для розробки системи соціальної динаміки

<sup>1</sup> НТУУ «КПІ ім. Ігоря Сікорського»

<sup>2</sup> GXperts GmbH, Відень

\* Листування: [viola.tret@gmail.com](mailto:viola.tret@gmail.com)

**Вступ.** Сучасний світ характеризується швидкими змінами і частими критичними ситуаціями. Людство стикається із все більш складними викликами, такими як пандемії, війни, що можуть призвести до соціальної напруженості. Мати засоби для моніторингу і виміру напруженості дозволяє як суспільству, так і державі, реагувати вчасно та адекватно на ці виклики. Напруженість призводить до соціальних конфліктів, політичних криз, економічних спадів. Мати систему для виміру напруженості допомагає розуміти, які чинники сприяють напруженості, і вживати заходів для попередження чи пом'якшення наслідків. Тож актуально розробляти інструменти для виміру напруженості суспільства, оскільки це є важливим кроком у напрямку розуміння та управління соціальною динамікою.

**Мета роботи** – застосування методів машинного навчання та нейролінгвістичного програмування для задачі аналізу думок користувачів Інтернет для передбачення соціальної напруженості в суспільстві.

**Результати.** У даній роботі для аналізу задач соціальної динаміки запропоновано скомбінувати підхід із застосуванням векторного представлення слів та модель кластеризації, щоб найточніше задовольнити потреби розроблюваної програми, яка оперує відкритими непоміченими текстовими даними українською мовою. У роботі розроблено архітектуру та програмне забезпечення системи соціальної динаміки на основі методів машинного навчання. Проведено розбиття на чотири модулі: обробка текстових даних, тренування моделі Word2Vec, тренування моделі K-Means та користувацький інтерфейс; проведено навчання моделей із різними гіперпараметрами, налаштованими вручну. Наведено графік напруженості суспільства, що показує тенденції соціальної динаміки українців.

**Ключові слова:** аналіз напруженості, соціальна динаміка, машинне навчання, обробка текстових даних, word2vec, k-means.

UDC 519.67

Violeta Tretynyk<sup>1\*</sup>, Yulia Nad<sup>2</sup>

## Using Machine Learning Methods to Develop a System of Social Dynamics

<sup>1</sup> The National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

<sup>2</sup> GXperts GmbH, Vienna

\* Correspondence: [viola.tret@gmail.com](mailto:viola.tret@gmail.com)

**Introduction.** The modern world is characterized by rapid changes and frequent critical situations. Humanity faces increasingly complex challenges, such as pandemics, wars, which can lead to social tensions. Having means for monitoring and measuring tension allows both society and the state to respond in a timely and adequate manner to these challenges. Tension leads to social conflicts, political crises, and economic downturns. Having a system for measuring tension helps us understand what factors contribute to tension and take steps to prevent or mitigate the effects. Therefore, it is urgent to develop tools to measure societal tensions, as this is an important step in the direction of understanding and managing social dynamics.

**The purpose of the article.** The purpose of this paper is to apply the methods of machine learning and neurolinguistic programming to the task of analyzing the opinions of Internet users to predict social tension in society.

**Results.** In this work, for the analysis of the problems of social dynamics, it is proposed to combine the approach using the vector representation of words and the clustering model in order to most accurately meet the needs of the developed program, which operates on open, unobserved text data in the Ukrainian language. The architecture and software of the social dynamics system based on machine learning methods were developed. It is divided into four modules: text data processing, Word2Vec model training, K-Means model training and user interface; models were trained with different manually adjusted hyperparameters. A graph of social tensions is presented, showing trends in the social dynamics of Ukrainians.

**Keywords:** tension analysis, social dynamics, machine learning, text data processing, word2vec, k-means.