DOI: https://doi.org/10.36910/6775-2524-0560-2025-60-19
УДК 004.02
Коровій Олександр Сергійович, аспірант
https://orcid.org/0009-0002-2835-9173
Терейковський Ігор Анатолійович, д.т.н., професор
https://orcid.org/0000-0003-4621-9668
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

МЕТОД ПОБУДОВИ НЕЙРОМЕРЕЖЕВИХ ЗАСОБІВ РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО ЗАБАРВЛЕННЯ ТЕКСТІВ В УНІВЕРСАЛЬНИХ КОМП'ЮТЕРНИХ ЗАСОБАХ.

Коровій О. С., Терейковський І. А. Метод побудови нейромережевих засобів розпізнавання емоційного забарвлення текстів в універсальних комп'ютерних засобах. У статті запропоновано метод побудови нейромережевих засобів для розпізнавання емоційного забарвлення українськомовних текстів в умовах «універсальних» обчислювальних засобів (CPU/мобільні пристрої/GPU). Метод уніфікує етапи кодування вхідних даних і декодування вихідних даних нейроморежі у ймовірнісні оцінки та формалізує алгоритм вибору архітектури залежно від ресурсних обмежень і вимог до точності. Передбачено три гілки: легковагові моделі для обмежень обчислювальних ресурсів, послідовні рекурентні моделі для помірних ресурсів і моделі на основі архітектури трансформер для сценаріїв з пріоритетом якості. Показано уніфікований конвеєр обробки текстових фрагментів: токенізація/векторизація, нейромережеве перетворення, нормування і вибір домінантної емоції. Експериментальна перевірка двох прототипів підтверджує компроміс «якість–ефективність»: RoBERTa перевершує FastText за точністю, однак вимагає більше параметрів і має вищу затримку інференсу (мілісекунди проти мікросекунд). Метод орієнтовано на українську мову та змішаний контент соціальних мереж, підтримує оптимізації для ресурс-обмежених платформ. Практична цінність полягає у керованому виборі архітектури і відтворюваному процесі розроблення, що забезпечує баланс точності, швидкодії та енергозатрат. Перспективи включають ансамблі/гібриди, роботу з сарказмом та розширення переліку емоцій. Результати демонструють, що систематизований підхід дає можливість створювати адаптивні рішення для моніторингу настроїв у реальному часі та інтеграції в мобільні й вбудовані системи. Запропонована методика сумісна з існуючими корпусами, метриками та стандартами відтворюваності і відкритої наукової практики.
Ключові слова: емоційна тональність, метод розпізнавання емоцій, обробка природної мови, нейронні мережі

Korovii O., Tereikovskyi I. Method for constructing neural network tools for recognizing the emotional tone of texts in universal computing devices. This paper proposes a method for constructing neural-network tools to recognize the emotional coloring of Ukrainian-language texts on "universal" computing platforms (CPU/mobile devices/GPU). The method standardizes the stages of input encoding and neural-network output decoding into probabilistic estimates, and formalizes an algorithm for selecting the architecture based on resource constraints and accuracy requirements. Three branches are provided: lightweight models for tight computational budgets, sequential recurrent models for moderate resources, and Transformer-based models for quality-first scenarios. A unified pipeline for processing text fragments is presented: tokenization/vectorization, neural transformation, normalization, and selection of the dominant emotion. Experimental evaluation of two prototypes confirms the "quality–efficiency" trade-off: RoBERTa outperforms FastText in accuracy, but requires more parameters and exhibits higher inference latency (milliseconds versus microseconds). The method targets Ukrainian and mixed social-media content, and supports optimizations for resource-constrained platforms. Its practical value lies in a guided architecture choice and a reproducible development process that balance accuracy, speed, and energy consumption. Future directions include ensembles/hybrids, handling sarcasm, and expanding the set of emotions. The results show that a systematized approach enables adaptive solutions for real-time mood monitoring and integration into mobile and embedded systems. The proposed methodology is compatible with existing corpora, metrics, and the standards of reproducibility and open science.
Keywords: emotional tone, method emotion recognition, natural language processing, neural networks

Постановка наукової проблеми.

Швидке зростання використання соціальних мереж в Україні створило безпрецедентну можливість для розуміння суспільних настроїв, емоційних реакцій на події та соціальної динаміки, особливо, під час кризових періодів. Проте існуючі системи виявлення емоцій часто не справляються з точним опрацюванням контенту українською мовою через кілька причин: лінгвістичні особливості української комунікації в соціальних мережах, наявність змішаного мовного контенту та унікальний культурний контекст, який впливає на вираження емоцій. Окрім лінгвістичних викликів, у більшості вбудованих систем (мобільні пристрої, бортові комп'ютери, дрони) доступні обчислювальні ресурси є обмеженими, дослідження та розроблення методу розпізнавання емоцій, здатного ефективно працювати, як на процесорах центрального типу (CPU), так і на графічних прискорювачах (GPU), є одним із стратегічних напрямів інноваційної діяльності в галузі сучасних інформаційних та комунікаційних технологій.

Аналіз досліджень. Виявлення емоційного забарвлення у текстових даних являє собою одну з пріоритетних дослідницьких задач у галузі комп'ютерної лінгвістики та обробки природної мови,

методологічний апарат якої зазнав суттєвої трансформації. Розвиток задачі визначення емоційного забарвлення текстів став можливим завдяки появі стабільних корпусів і чітких метрик. SemEval-2018 Task 1 («Affect in Tweets») заклав стандартні підзадачі (інтенсивність, розпізнавання та класифікація емоцій), тоді як NRC Emotion Lexicon забезпечує широку ресурсну базу для емоційних асоціацій, а GoEmotions — великомасштабну розмітку для багатоміткової класифікації (27 емоцій + нейтральна) [1–3]. Методологічний зсув від лексиконів і Support Vector Machine (SVM) до попередньо навчених мовних моделей на основі архітектури Transformer, закріпили BERT та RoBERTa, що істотно підвищують якість без складних архітектурних модифікацій. Для багатомовних і низькоресурсних сценаріїв (зокрема української) ключовою стала нейромережева модель (НММ) XLM-R, яка забезпечує ефективне міжмовне перенесення ознак [4–6]. Узагальнені огляди останніх років підтверджують провідну роль архітектури трансформер у текстовій емоційній аналітиці та підкреслюють потребу в якісні розмітці, великих навчальних корпусах, та оптимізаціях для різних обчислювальних засобів [2, 6, 12, 16].

У прикладних умовах «універсальних комп'ютерних засобів» (звичайні CPU/мобільні пристрої) є ефективність обробки вхідних сигналів нейронною мережою. Дистиляція знань (DistilBERT, TinyBERT) дає змогу суттєво зменшувати моделі зі збереженням більшості якості; MobileBERT цілеспрямовано оптимізовано для обмежених ресурсів [7–11].

Окремий внесок у формування методології для українськомовних даних демонструють праці: у [12] подано модель побудови нейромережевих засобів розпізнавання емоційних фрагментів тексту (модульність етапів та орієнтація на компроміс «якість–ресурси»); у [13] запропоновано концептуальну модель процесу визначення емоційної тональності; у [14] обґрунтовано адаптацію методу дистиляції знань для аналізу тональності, релевантну до створення компактних моделей для ресурс-обмежених систем. Сукупно ці результати становлять методологічне підґрунтя для подальшого конструювання цілісного методу побудови НМ-засобів у «універсальних» обчислювальних умовах [12–14].

У проаналізованих джерелах відсутня єдність щодо методу вибору архітектури НМ-моделі для задач емоційної класифікації в багатомовних/малоресурсних умовах; недостатньо повно висвітлено уніфіковану процедуру підготовки вхідних даних перед подачею в нейронну мережу та інтерпретації вихідних даних.

Метою статті є розроблення методу побудови нейромережевих засобів в універсальних комп'ютерних засобах за рахунок гнучкого підходу до обмежених ресурсів, що дозволить ефективне розпізнавання емоційного забарвлення текстів.

Виклад основного матеріалу дослідження. Запропонований метод ґрунтується на теоретичних положеннях, експериментах та висновках, викладених у [12-15], та інтегрує сучасні підходи до побудови систем обробки природної мови для вирішення задачі розпізнавання емоцій в умовах універсальних комп'ютерних засобів.

Ґрунтуючись на концептуальній моделі процесу розпізнавання [14], загальний процес розпізнавання емоційного забарвлення тексту за допомогою нейромережевих моделей можна представити у вигляді послідовності ключових етапів, як показано на рис. 1.

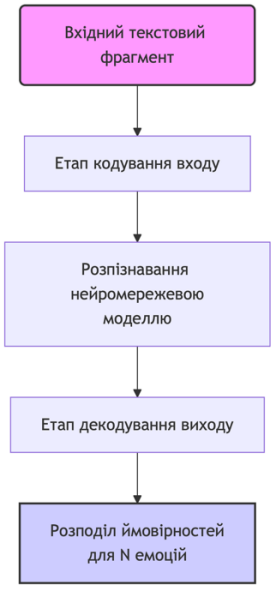

Рис. 1. Загальна блок-схема процесу розпізнавання емоцій

Процес ініціюється подачею на вхід текстового фрагмента. На етапі кодування відбувається трансформація сирого тексту у числовий формат (вектор або послідовність векторів), придатний для обробки нейронною мережею. Найбільш узагальнено процес кодування можна описати так:

$$V = E(T), \quad (1)$$

де:

- V – це вихідний числовий формат: вектор або послідовність векторів.[
- E – це функція кодування, яка перетворює текст у вектори.
- T – це вхідний сирий текст.

Ця формула (1) лаконічно відображає, що векторне представлення V є результатом застосування функції кодування E до тексту T.
Нейромережева модель, обробляє цей числовий вхід (1) для виявлення прихованих патернів.
На етапі декодування вихідні сигнали моделі перетворюються у фінальний, інтерпретований результат — розподіл ймовірностей, що показує ступінь присутності кожної емоції у тексті:

$$P = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^N, \quad (2)$$

де N — це кількість цільових емоцій, які нейронна мережа навчена розпізнавати, а кожна компонента p_i відповідає ймовірності присутності і-тої емоції в тексті.

Відповідно до робіт [13, 14] першим кроком потрібно обрати архітектуру нейромережевої моделі (НММ), так як від цього етапу залежить процедура процесингу вхідних/вихідних даних. Для цього було розроблено алгоритм зображений на рис. 2.

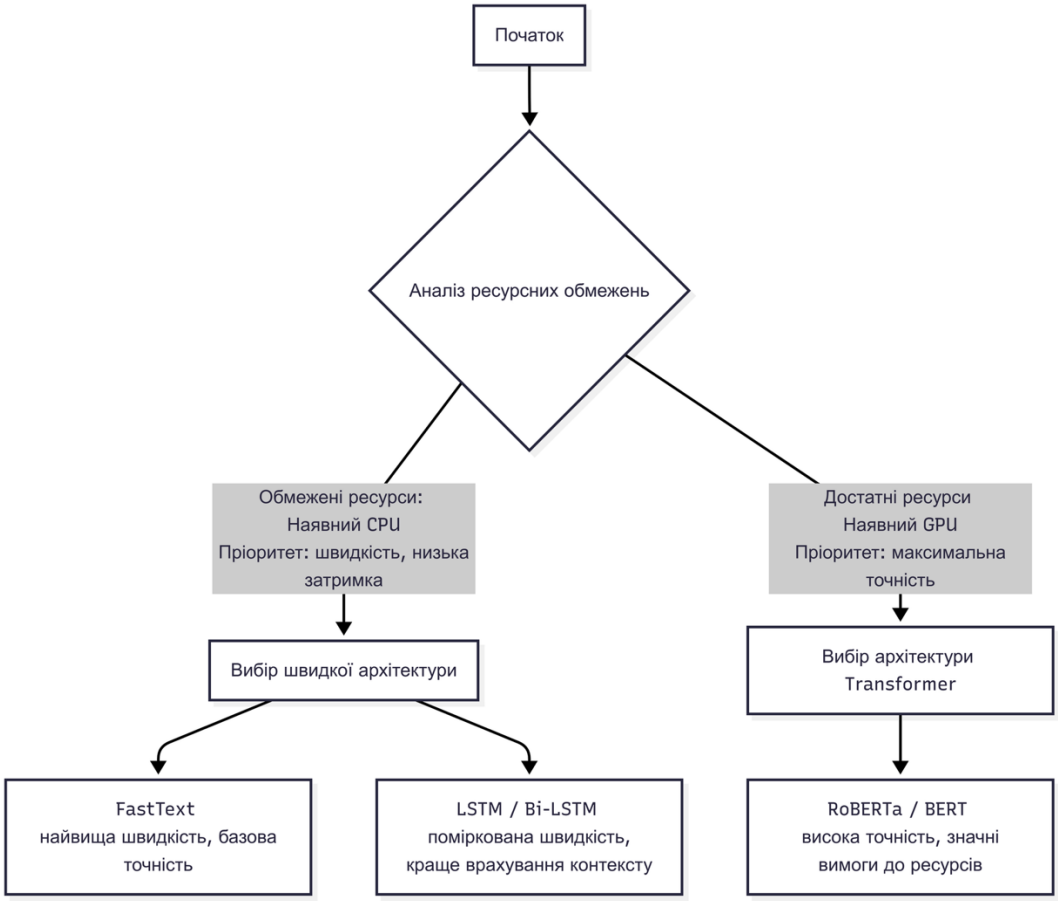

Рис 2. Алгоритм вибору архітектури НММ

В основі алгоритму зображеного на рисунку 1, лежить аналіз ресурсних обмежень. Перш за все треба врахувати вимоги до швидкості, наявності обчислювальних ресурсів, та рівня точності розпізнавання. При наявності лише CPU, мобільного пристрою, обмеженнь енергоспоживання, в цьому випадку краще використовувати архітектуру – FastText. Якщо потужність CPU дозволяє більш об'ємні обчислення, тоді краще використовувати – LSTM/Bi-LSTM. В іншому випадку, коли важлива висока точність, а також є можливість використовувати GPU, тоді є доцільним застосування НММ на базі архітектури – Transformer, представниками цієї архітектури є попередньо навчені моделі – RoBERTa/BERT/XLM.

Після визначення оптимального типу архітектури нейромережевої моделі (НММ) відповідно до ресурсних обмежень та вимог до продуктивності, наступним критичним етапом є її навчання. Цей процес не є довільним, а реалізується згідно зі структурованою методологією, що базується на положеннях, викладених у [13, 14]. Зокрема, загальна послідовність дій та взаємозв'язок етапів регламентуються концептуальною моделлю процесу розпізнавання емоцій [14], тоді як практичні аспекти реалізації, включаючи формування даних та вибір гіперпараметрів, ґрунтуються на підходах, апробованих у дослідженні моделі побудови нейромережевих систем[13]. Таке поєднання забезпечує системність та відтворюваність процесу навчання.

Специфіка реалізації етапів кодування та декодування залежить від обраної архітектури нейронної мережі. Розглянемо їх для трьох основних класів моделей.

1. FastText

Процес кодування для архітектури на основі усереднених векторних представлень (FastText) передбачає сегментацію вхідного тексту на лексичні одиниці (токени), в даному контексті токен – це повноцінне слово. Кожен токен відображається у відповідне йому векторне представлення (embedding) з попередньо навченого словника. Для отримання єдиного вектора, що характеризує весь текстовий фрагмент, застосовується агрегація (зазвичай усереднення) векторів усіх його

токенів. Цей підхід, відомий як "Bag-of-Embeddings", є обчислювально ефективним, проте ігнорує порядок слів у тексті [13].

Отриманий на етапі кодування єдиний вектор подається на вхід простого НММ на основі архітектури FastText, що, як правило, складається з одного або декількох повнозв'язних шарів. Останній шар генерує вектор необроблених значень (логітів), який далі перетворюється на розподіл ймовірностей (2), блок-схему зображено на рис 6.

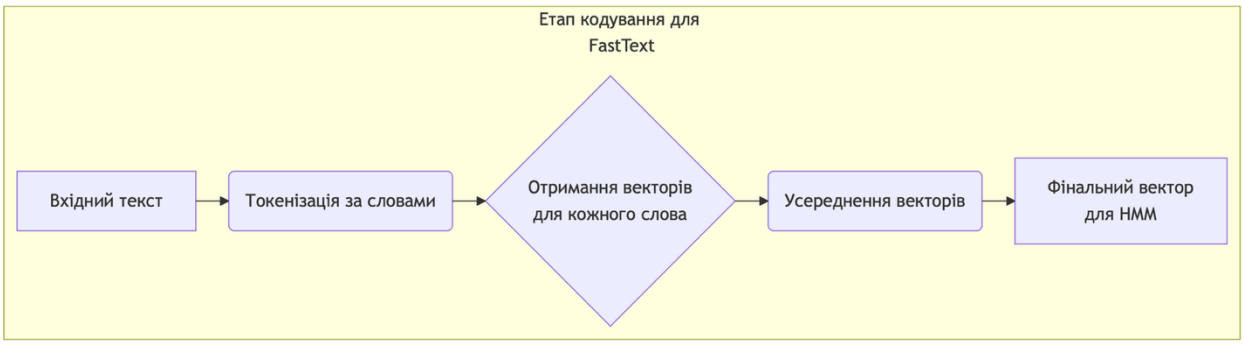

Рис. 3. Блок-схема кодування вхідних даних для архітектури FastText

2. LSTM/Bi-LSTM

На відміну від підходу що використовується у FastText, рекурентні мережі обробляють текст як впорядковану послідовність. Текст так само сегментується на токени, і кожен токен перетворюється на вектор. Однак агрегація не відбувається; на вхід LSTM-шарів подається повна послідовність векторів, що дозволяє моделі аналізувати контекстуальні зв'язки та залежності між словами, рис 4. Рекурентні шари послідовно обробляють вхідні вектори, акумулюючи інформацію у своєму прихованому стані. Фінальний прихований стан (або агрегація всіх станів) використовується, як репрезентація всього тексту і подається на класифікаційну "голову" для генерації вектора логітів, рис 6.

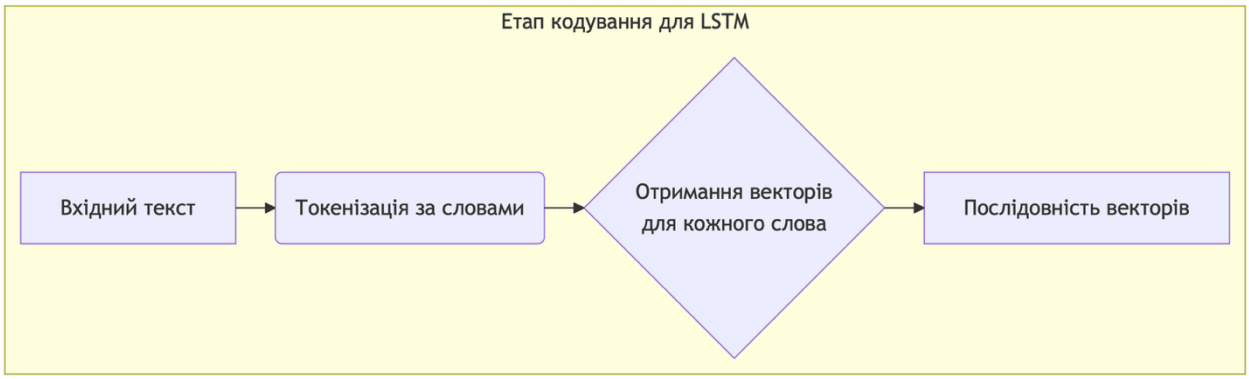

Рис. 4. Блок-схема кодування вхідних даних для архітектури LSTM

3. Transformer

НММ на основі архітектури Transformer використовують більш гранулярний підхід до токенізації, відомий як subword-токенізація (напр., Byte Pair Encoding або SentencePiece), який ефективно працює з рідкісними та невідомими словами. До послідовності субодиниць додаються спеціальні службові токени (напр., [CLS], [SEP]), що позначають початок послідовності та розділяють речення. На вхід моделі подаються не самі вектори, а їхні числові ідентифікатори (ID) з відповідними масками уваги, див. рис 5. Після обробки вхідних тензорів трансформерними шарами, для задачі класифікації всього тексту зазвичай використовується векторний стан, що відповідає службовому токену [CLS]. Вважається, що цей вектор агрегує інформацію про всю послідовність. Він подається на останній повнозв'язний шар НММ для отримання вектора логітів.

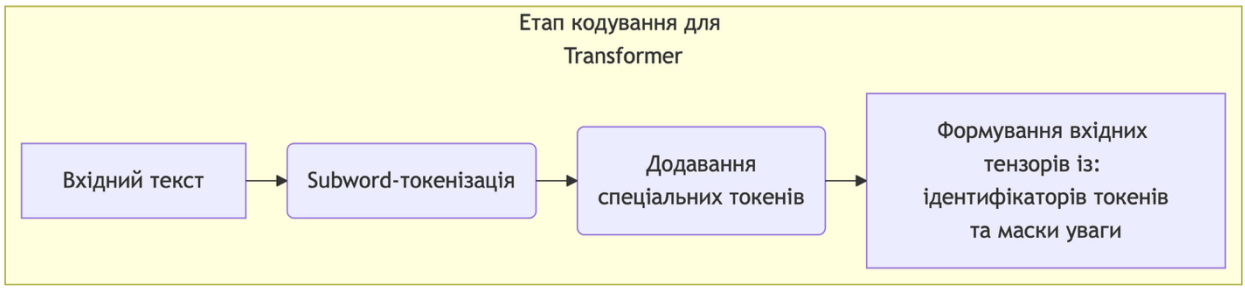

Рис. 5. Блок-схема кодування вхідних даних для архітектури Transformer

Незалежно від обраної архітектури (FastText, LSTM чи Transformer), фінальний етап перетворення вихідних даних моделі у ймовірнісну оцінку є уніфікованим для задач мультикласової класифікації. Останній повнозв'язний шар нейронної мережі генерує вектор необроблених числових значень, відомих як - логіти (*англ.* logits).Для перетворення цього вектора у розподіл ймовірностей (2) застосовується функція активації Softmax. Якщо $z = (z_1, z_2, ..., z_n)$ — це вектор логітів, де N — кількість емоційних класів, то ймовірність p_i для i-го класу обчислюється за формулою (3):

$$p_i = \text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}, \quad (3)$$

Ця функція нормалізує вектор логітів таким чином, що кожен елемент вихідного вектора знаходиться в діапазоні, а сума всіх елементів дорівнює 1. Таким чином, вихідний елемент можна інтерпретувати як ймовірність належності тексту до відповідного емоційного класу.

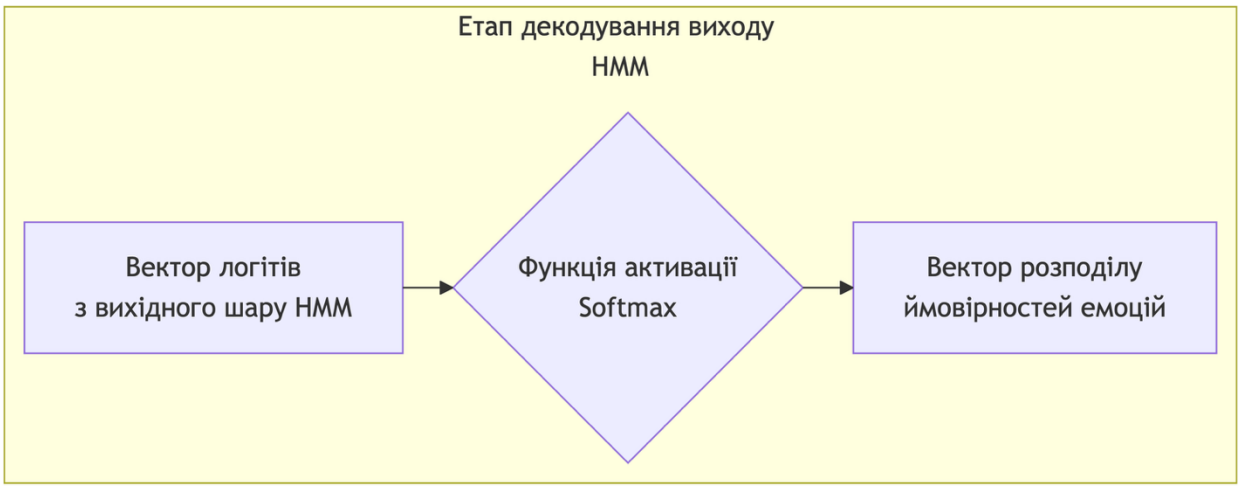

Рис. 6. Блок-схема декодування логітів у розподіл ймовірностей

Клас емоції, що відповідає максимальному значенню у фінальному векторі ймовірностей, вважається результатом розпізнавання моделі. Це дозволяє однозначно визначити домінантну емоцію у проаналізованому текстовому фрагменті.

Експериментальна апробація методу

Практична перевірка та валідація запропонованого методу була проведена в рамках попереднього дослідження [13]. У зазначеній роботі було реалізовано два прототипи, що відповідають двом гілкам запропонованого алгоритму вибору: легковаговий класифікатор на основі FastText для систем з обмеженими ресурсами та високоточний класифікатор на основі архітектури RoBERTa-base для систем, де пріоритетом є якість розпізнавання.

Результати експерименту, представлені в [13], наочно демонструють ключовий компроміс: модель RoBERTa досягла значно вищої якості (F1-score = 0.91) порівняно з FastText (F1-score = 0.73), однак ціною суттєво більших обчислювальних витрат (125 млн параметрів проти 3 млн) та значно більшого часу інференсу (9 мс проти 5 мкс). Ці результати підтверджують доцільність

запропонованого методу, який формалізує вибір оптимального інструменту залежно від практичних умов його застосування.

Висновки. У ході дослідження розроблено комплексний метод для систематизованого конструювання та вибору нейромережевих засобів розпізнавання емоційного забарвлення текстів. Ключовою особливістю методу є запропонований алгоритм вибору архітектури (легковагові vs. трансформерні), що ґрунтується на аналізі доступних обчислювальних ресурсів (CPU/GPU) та вимог до продуктивності, дозволяючи створювати рішення з оптимальним балансом між точністю та ефективністю. Метод формалізує весь процес розробки, включаючи уніфіковані процедури кодування вхідних даних та декодування вихідних сигналів моделі, що забезпечує його гнучкість та адаптивність. Ефективність запропонованого підходу підтверджується результатами експериментальних досліджень [13], які продемонстрували практичну доцільність вибору різних архітектур для різних умов експлуатації.

Розроблений метод має значну практичну цінність, оскільки надає розробникам і дослідникам інструментарій для обґрунтованого створення спеціалізованих систем аналізу текстів, зокрема для мов з обмеженими ресурсами, як-от українська. Перспективи подальших досліджень спрямовані на розширення функціональних можливостей методу та підвищення якості розпізнавання. Це включає дослідження гібридних та ансамблевих архітектур для досягнення кращого компромісу між точністю та швидкістю; покращення стійкості моделей до складних лінгвістичних явищ, таких як сарказм та іронія; розширення набору розпізнаваних емоцій для більш гранулярного аналізу; а також формування та стандартизація репрезентативних навчальних корпусів для української мови, що є критично важливим для подальшого прогресу в цій галузі.

Список бібліографічного опису

1.Mohammad, S. M., Bravo-Marquez, F., Salameh, M., & Kiritchenko, S. (2018). SemEval-2018 Task 1: Affect in Tweets. In Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018) (pp. 1-17). New Orleans, Louisiana: Association for Computational Linguistics. DOI: 10.18653/v1/S18-1001

2.Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a Word-Emotion Association Lexicon. Computational Intelligence, 29(3), 436-465. DOI: 10.1111/j.1467-8640.2012.00460.x

3.Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 4040-4054). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.372

4.Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171-4186). Minneapolis, Minnesota: Association for Computational Linguistics. DOI: 10.18653/v1/N19-1423

5.Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692. DOI: 10.48550/arXiv.1907.11692

6.Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 8440-8451). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.747

7.Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108. DOI: 10.48550/arXiv.1910.01108

8.Jiao, X., Yin, Y., Shang, L., Jiang, X., Chen, X., Li, L., Wang, F., & Liu, Q. (2020). TinyBERT: Distilling BERT for Natural Language Understanding. In Findings of the Association for Computational Linguistics: EMNLP 2020 (pp. 4163-4174). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.372

9.Sun, Z., Yu, H., Song, X., Liu, R., Yang, Y., & Zhou, D. (2020). MobileBERT: a Compact Task-Agnostic BERT for Resource-Limited Devices. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 2158-2170). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.195

10. Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2018). Quantization and Training of Neural Networks for Efficient Integer-Only Inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2704-2713). Salt Lake City, UT, USA. DOI: 10.1109/CVPR.2018.00286

11. Kim, S., Gholami, A., Yao, Z., Mahoney, M. W., & Keutzer, K. (2021). I-BERT: Integer-only BERT Quantization. arXiv preprint arXiv:2101.01321. DOI: 10.48550/arXiv.2101.01321

12. Дичка, І. А., Терейковський, І. А., Коровій, О. С., Терейковська, Л. О., & Романкевич, В. О. (2023). Оцінка Ефективності Засобів Розпізнавання Емоційної Тональності Фрагментів Тексту. ВЧЕНІ ЗАПИСКИ ТНУ ІМЕНІ В.І. ВЕРНАДСЬКОГО. СЕРІЯ: ТЕХНІЧНІ НАУКИ, 34(3), 20-27. DOI: https://doi.org/10.32782/2663-5941/2023.3.1/20

13. Tereikovskyi, I. A., & Korovii, O. S. (2025). A Model for Constructing Neural Network Systems for Recognizing Emotions of Text Fragments. Herald of Advanced Information Technology, 8(2), 197-208. DOI: https://doi.org/10.15276/hait.08.2025.12

14. Коровій, О., & Терейковський, І. (2024). Концептуальна модель процесу визначення емоційної тональності тексту. КОМП'ЮТЕРНО-ІНТЕГРОВАНІ ТЕХНОЛОГІЇ: ОСВІТА, НАУКА, ВИРОБНИЦТВО, (55), 115-123. DOI: https://doi.org/10.36910/6775-2524-0560-2024-55-14
15. Korovii, O., & Petrashenko, A. (2022). Adaptation of Distilling Knowledge Method in NLP for Sentiment Analysis. In Shkarlet, S., Moroz, B., & Palagin, A. (Eds.), Proceedings of the 11th International Scientific and Practical Conference on Information Security and Information and Communication Systems (ISEM 2021) (Vol. 463, pp. 248-258). Springer, Cham. DOI: 10.1007/978-3-031-08114-7_22
16. Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: a review. Artificial Intelligence Review, 53(6), 4335–4385. DOI: https://doi.org/10.1007/s10462-019-09794-5

References:

1.Mohammad, S. M., Bravo-Marquez, F., Salameh, M., & Kiritchenko, S. (2018). SemEval-2018 Task 1: Affect in Tweets. In Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018) (pp. 1-17). New Orleans, Louisiana: Association for Computational Linguistics. DOI: 10.18653/v1/S18-1001
2.Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a Word-Emotion Association Lexicon. Computational Intelligence, 29(3), 436-465. DOI: 10.1111/j.1467-8640.2012.00460.x
3.Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 4040-4054). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.372
4.Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171-4186). Minneapolis, Minnesota: Association for Computational Linguistics. DOI: 10.18653/v1/N19-1423
5.Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692. DOI: 10.48550/arXiv.1907.11692
6.Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 8440-8451). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.747
7.Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108. DOI: 10.48550/arXiv.1910.01108
8.Jiao, X., Yin, Y., Shang, L., Jiang, X., Chen, X., Li, L., Wang, F., & Liu, Q. (2020). TinyBERT: Distilling BERT for Natural Language Understanding. In Findings of the Association for Computational Linguistics: EMNLP 2020 (pp. 4163-4174). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.372
9.Sun, Z., Yu, H., Song, X., Liu, R., Yang, Y., & Zhou, D. (2020). MobileBERT: a Compact Task-Agnostic BERT for Resource-Limited Devices. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 2158-2170). Online: Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.195
10. Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2018). Quantization and Training of Neural Networks for Efficient Integer-Only Inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2704-2713). Salt Lake City, UT, USA. DOI: 10.1109/CVPR.2018.00286
11. Kim, S., Gholami, A., Yao, Z., Mahoney, M. W., & Keutzer, K. (2021). I-BERT: Integer-only BERT Quantization. arXiv preprint arXiv:2101.01321. DOI: 10.48550/arXiv.2101.01321
12. Dychka, I. A. Tereikovskyi, O. S. Korovii, L. O. Tereikovska, and V. O. Romankevych, "Evaluation of the effectiveness of means for recognizing the emotional tonality of text fragments," Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences, vol. 34 (73), no. 3, part 1, pp. 130-135, 2023.DOI: https://doi.org/10.32782/2663-5941/2023.3.1/20
13. Tereikovskyi, I. A., & Korovii, O. S. (2025). A Model for Constructing Neural Network Systems for Recognizing Emotions of Text Fragments. Herald of Advanced Information Technology, 8(2), 197-208. DOI: https://doi.org/10.15276/hait.08.2025.12
14. Korovii, O., & Tereikovskyi, I. (2024). Conceptual Model of the Process for Determining the Emotional Tone of Text. Computer-Integrated Technologies: Education, Science, Production, (55), 115–123. DOI: https://doi.org/10.36910/6775-2524-0560-2024-55-14
15. Korovii, O., & Petrashenko, A. (2022). Adaptation of Distilling Knowledge Method in NLP for Sentiment Analysis. In Shkarlet, S., Moroz, B., & Palagin, A. (Eds.), Proceedings of the 11th International Scientific and Practical Conference on Information Security and Information and Communication Systems (ISEM 2021) (Vol. 463, pp. 248-258). Springer, Cham. DOI: 10.1007/978-3-031-08114-7_22
16. Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: a review. Artificial Intelligence Review, 53(6), 4335–4385. DOI: https://doi.org/10.1007/s10462-019-09794-5