

УДК 004.93:004.032.26

## ЗАСТОСУВАННЯ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ У ЗАДАЧАХ КЛАСИФІКАЦІЇ ТЕКСТІВ

Каяфюк А.Г., аспірант

*Київський національний університет технологій та дизайну*

*Ключові слова:* класифікація текстів, згорткові нейронні мережі, TextCNN, DPCNN, обробка природної мови, 20 Newsgroups.

**Мета роботи:** дослідити доцільність використання згорткових нейронних мереж у задачах класифікації текстів та узагальнити архітектурні й інженерні рішення, які дозволяють досягати кращого співвідношення між якістю класифікації, обчислювальною вартістю та швидкодією порівняно з класичними моделями машинного навчання і ресурсоемними трансформерними підходами.

**Матеріали та методи:** Класифікація текстів є однією з ключових задач обробки природної мови, оскільки використовується у фільтрації повідомлень, тематичній класифікації документів, аналізі звернень користувачів, модерації контенту, а також у технічних задачах на кшталт категоризації лог файлів, інцидентів і результатів автоматизованого тестування. У сучасних дослідженнях найвищу якість часто забезпечують трансформерні моделі, однак їх практичне впровадження пов'язане зі значними витратами на навчання, пам'ять та вивід [3, 4]. Класичні підходи на основі TF-IDF, лінійних класифікаторів та ансамблів дерев є дешевшими, проте мають обмеження у здатності будувати ієрархічні представлення тексту. У цьому контексті згорткові нейронні мережі (CNN) займають проміжну нішу між розрідженими ознаковими моделями та важкими трансформерними архітектурами.

Ефективність CNN у текстовій класифікації пояснюється тим, що необхідна інформація часто локалізована у коротких мовних шаблонах: ключових фразах, запереченнях, типовій послідовності термінів, кодах помилок, стійких виразах і повторюваних n-грамних конструкціях. Одновимірні згортки дозволяють навчати фільтри, що автоматично виділяють такі патерни без ручного конструювання ознак. Типова архітектура TextCNN включає токенізацію, embeddinglayer, кілька паралельних згортків із різними розмірами ядра, глобальний max-pooling, конкатенацію ознак, dropout і фінальний повнозв'язний класифікаційний шар [1]. Ядра розміру 3, 4 і 5 зазвичай дають змогу захоплювати локальні залежності різної довжини.

Подальший розвиток CNN для текстів відбувався у напрямі ускладнення рецептивного поля та підвищення стабільності навчання. До поширених рішень належать глибокі стекові CNN, залишкові з'єднання,

багатоканальні моделі, дилатовані згортки, attentionpooling, а також гібридні архітектури, у яких CNN поєднуються з BiLSTM/GRU або попередньо навченими ембедінгами [3, 4]. Окремо варто виділити DPCNN, де пірамідалне зменшення довжини послідовності дозволяє будувати глибшу мережу без різкого зростання обчислювальної складності [2]. Саме такі архітектури є перспективними для задач, де потрібно вловити не лише короткі патерни, а й ширший контекст документа.

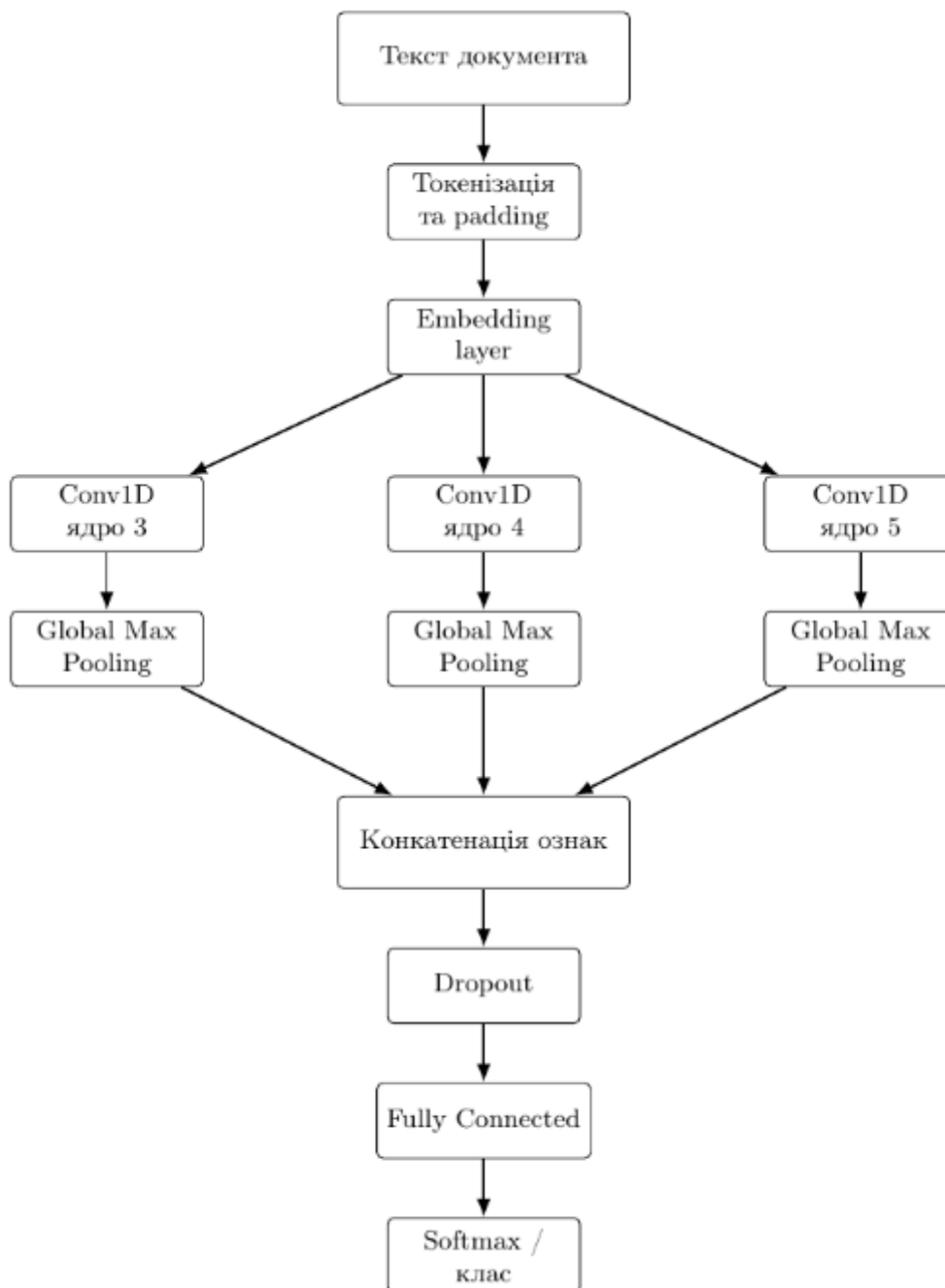


Рисунок 1 – Архітектура TextCNN для задачі класифікації текстів

Для практичного використання у задачах класифікації текстів доцільно порівнювати CNN із сильним базовим варіантом TF-IDF +

LinearSVC. Якщо набір даних невеликий або тексти короткі й добре описуються ключовими словами, класична модель може залишатися дуже конкурентною. Проте за наявності шуму, варіативних формулювань, технічної лексики або залежності від локальних комбінацій токенів CNN часто демонструють кращу узагальнювальну здатність. Для зменшення вартості навчання доцільно обмежувати максимальну довжину послідовності, очищати очевидний шум, використовувати помірний розмір словника, dropout, earllystopping, weightdecay, а також mixedprecision і мінібатчі помірною розміру; у режимах обмежених ресурсів доцільними є також low-resource та fasttextclassification підходи [5]. У випадку довгих документів корисною є стратегія нарізання на вікна з подальшим агрегуванням прогнозів.

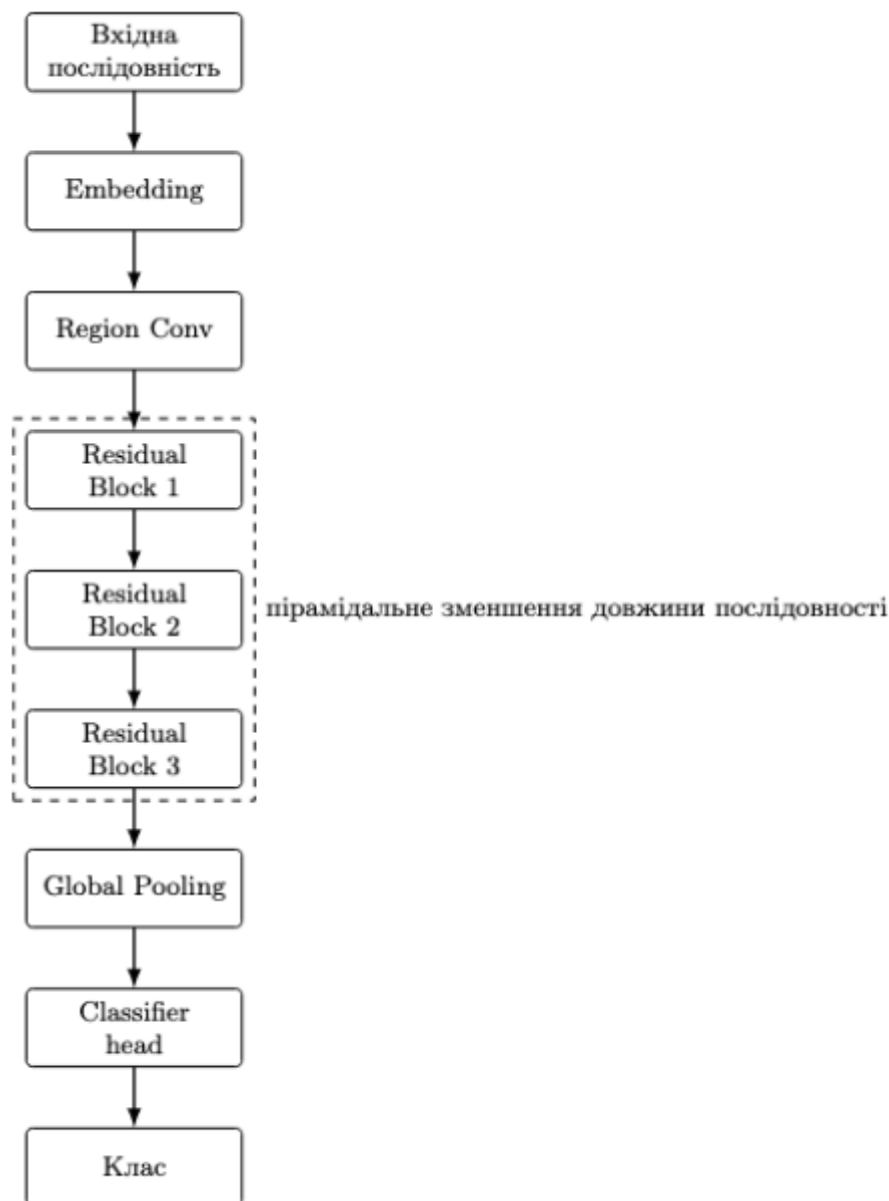


Рисунок 2 – Схема DPCNN із пірамідальним зменшенням довжини послідовності

У рамках ілюстративного експериментального конвеєра для задачі тематичної класифікації може бути використаний корпус 20 Newsgroups. Для нього доцільно реалізувати щонайменше три архітектури: базовий TextCNN, глибшу DeepCNN та DPCNN. Такий набір дає можливість оцінити, як змінюється якість класифікації при переході від поверхневого виділення локальних ознак до глибших ієрархічних представлень. Окремою перевагою CNN є добра паралелізація згорток на GPU і прийнятна швидкодія на CPU, що робить їх придатними для систем із жорсткими вимогами до latency, пакетного скорингу та внутрішніх корпоративних сервісів.

Висновки: Згорткові нейронні мережі залишаються актуальним класом моделей для класифікації текстів, особливо в тих сценаріях, де трансформери є надто дорогими або надлишковими, а класичні моделі вже досягають межі своїх можливостей [3, 4].

Найбільшу практичну цінність CNN мають у задачах, де класи сигналізуються локальними текстовими патернами, важливими є швидкість інференсу, контроль обчислювальних витрат і можливість масштабованого впровадження.

Найперспективнішими виглядають гібридні та поглиблені архітектури CNN, що поєднують багатомасштабні згортки, залишкові з'єднання, попередньо навчені подання та раціонально налаштований навчальний конвеєр [2–5].

#### Список використаних джерел

1. Kim Y. Convolutional Neural Networks for Sentence Classification // Proceedings of EMNLP. 2014. DOI: 10.3115/v1/D14-1181.
2. Johnson R., Zhang T. Deep Pyramid Convolutional Neural Networks for Text Categorization // Proceedings of ACL. 2017. URL: <https://aclanthology.org/P17-1052/>.
3. Taha K., Yoo P. D., Yeun C. et al. A comprehensive survey of text classification techniques and their research applications: Observational and experimental insights // Computer Science Review. 2024. Vol. 54. 100664. DOI: 10.1016/j.cosrev.2024.100664.
4. Allam H., Haggag O., Hassan H., Elmasry H. Text Classification: How Machine Learning Is Shaping the Future of NLP // Information. 2025. Vol. 16(2). 130. DOI: 10.3390/info16020130.
5. Mao Y., Liu P., Ma Y. et al. Low-Resource Fast Text Classification Based on Intra-Class and Inter-Class Distance Calculation // Proceedings of COLING. 2025. URL: <https://aclanthology.org/2025.coling-main.70/>.