

# ІНТЕГРАЦІЯ НЕЙРОМЕРЕЖ В СИСТЕМИ РОЗПІЗНАВАННЯ МОВИ ДЛЯ ПОКРАЩЕННЯ ТОЧНОСТІ ТА ШВИДКОСТІ

Хльобас Денис Володимирович

студент групи ІАСм -1-24-1.4д,

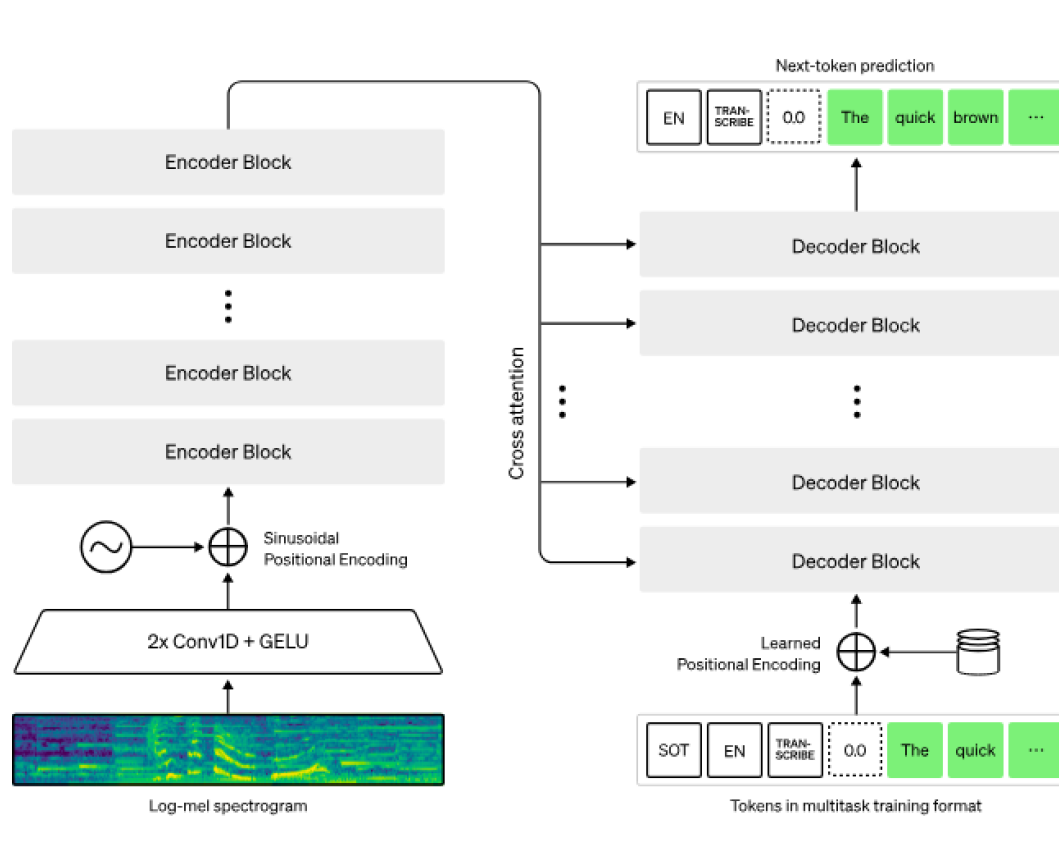
Київського столичного університету імені Бориса Грінченка,

науковий керівник – к.т.н., доц. Мельник І.Ю.

У сучасному світі технології обробки природної мови стали основою багатьох цифрових сервісів – від голосових асистентів до систем автоматичного перекладу. Ключову роль у цих процесах відіграють нейронні мережі, які суттєво підвищили якість і швидкість розпізнавання мовлення порівняно з традиційними статистичними моделями [1].

Інтеграція нейромереж у системи розпізнавання мови базується на використанні глибоких архітектур – зокрема згорткових (CNN), рекурентних (RNN, LSTM) та трансформерних моделей. Вони дозволяють враховувати як акустичні, так і контекстуальні особливості мовлення, що забезпечує вищу точність і стійкість до шуму [2].

Сучасні системи розпізнавання мови, такі як Google Speech-to-Text, Whisper від OpenAI чи DeepSpeech від Mozilla, демонструють, що саме поєднання глибокого навчання з великими мовними корпусами стало вирішальним чинником підвищення ефективності [3]. Використання трансформерів (наприклад, архітектури Wav2Vec 2.0) дало змогу моделі навчатися без попереднього маркування даних, що спростило підготовку великих датасетів і скоротило час розробки [4].



Ключовою перевагою нейромережових підходів є здатність узагальнювати дані, навчатися на різних мовах і акцентах, а також адаптуватися до конкретних користувачів. Завдяки цьому сучасні системи можуть досягати рівня точності понад 95%, що наближає їх до людського розпізнавання [1].

Разом із тим інтеграція таких систем вимагає оптимізації обчислювальних ресурсів. Для цього застосовуються методи квантизації, прунингу (скорочення кількості параметрів), а також використання апаратних прискорювачів – GPU та TPU. Це дозволяє зменшити затримку в реальному часі та забезпечити ефективне функціонування моделей на мобільних пристроях [3].

Перспективним напрямом подальших досліджень є поєднання нейромережових моделей розпізнавання мови з мовними моделями великого масштабу (LLM), що дозволить не лише точно розпізнавати, але й інтерпретувати зміст мовлення. Така інтеграція відкриває можливості для створення інтелектуальних асистентів нового покоління, здатних до контекстного розуміння та семантичного аналізу [5; 6].

Таким чином, застосування нейромереж у системах розпізнавання мови є ключовим чинником підвищення їхньої точності та швидкодії. Розвиток глибокого навчання та зростання обчислювальних потужностей створюють передумови для появи більш ефективних і адаптивних систем, що наближають взаємодію людини з комп'ютером до природного спілкування [2].

## ДЖЕРЕЛА

1. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks. – IEEE, 2013.
2. Hinton G., Deng L., Yu D. Deep neural networks for acoustic modeling in speech recognition. – IEEE Signal Processing Magazine, 2012.
3. Baevski A., Zhou H., Mohamed A., Auli M. wav2vec 2.0: A framework for self-supervised learning of speech representations. – arXiv:2006.11477, 2020.
4. OpenAI Whisper Model Card. – 2023. URL: <https://openai.com/research/whisper>
5. Mozilla DeepSpeech Project Documentation. – 2022.
6. Стрельніков, В. І., & Бондарчук, А. П. Комплексний підхід до інтелектуального управління, моделювання та виявлення мережних аномалій на основі ентропійних та нейромережових підходів. Телекомунікаційні та інформаційні технології, 2025, (2), 100-107.
7. Бушма, О. В., Машкіна, І. В., Носенко, Т. І., & Яскевич, В. О. Кваліфікаційна робота магістра: Навчально-методичний посібник для спеціальності «Комп'ютерні науки», 2024.