DOI 10.62731/mcnd-20.12.2024.002

ПОРІВНЯЛЬНИЙ АНАЛІЗ АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ: RNN, LSTM, GRU

Коломоєць Сергій Олексійович
ORCID ID: 0000-0003-3741-4517
аспірант, асистент кафедри Інформаційних систем та технологій
Факультет інформатики та обчислювальної техніки
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Україна

Науковий керівник: Чумаченко Олена Іллівна
ORCID ID: 0000-0003-3006-7460
професор, д-р. техн. наук, професор кафедри Штучного інтелекту
Інститут прикладного системного аналізу
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Україна

У традиційних штучних нейронних мережах (ШНМ) існує проблема пов'язана зі збереженням інформації, яка була до них. Це наче починати читати книгу зсередини і не знати що було в попередніх главах. Не має жодної збереженої інформації про попередні дії, що в свою чергу впливає на класифікацію теперішніх подій.

Проблему традиційних ШНМ розв'язують рекурентні нейронні мережі (Recurrent Neural Networks RNN), за рахунок циклів, які зберігають інформацію [1]. Однак зі збільшенням кількості інформації RNN втрачають зв'язок між інформацією та стають неточними. Однак було розроблено архітектуру довгої короткочасної пам'яті (Long Short-Term Memory LSTM) – підтип RNN, який може навчатися довгостроковим залежностям [2]. Створення LSTM вирішило проблемо зі збереженням інформації, однак ця архітектура є складною та енергозатратною для обчислень, тому було створено спрощену мережу для досягнення балансу між складність та ефективністю – закриті рекурентні блоки (Gated Recurrent Unit GRU).

Однією з проблем, яку GRU вирішує порівняно з LSTM, є його простіша архітектура, що призводить до скорочення часу навчання і вимагає меншої кількості параметрів для вивчення [3]. GRU досягає

цього, об'єднуючи вентилі забування і введення в єдиний вентиль оновлення, що спрощує механізм вентилювання порівняно з LSTM.

Спрощена конструкція робить GRU більш ефективними в обчислювальному плані і простішими в реалізації, водночас пропонуючи конкурентну ефективність у виявленні довготривалих залежностей у послідовних даних. Крім того, GRU менш схильні до перенастроювання через меншу складність, що робить їх особливо придатними для задач з обмеженими навчальними даними або обчислювальними ресурсами. Схематичне порівняння архітектур RNN, LSTM та GRU зображено на рисунку 1. Порівняння архітектур LSTM та GRU відображено в таблиці 1.

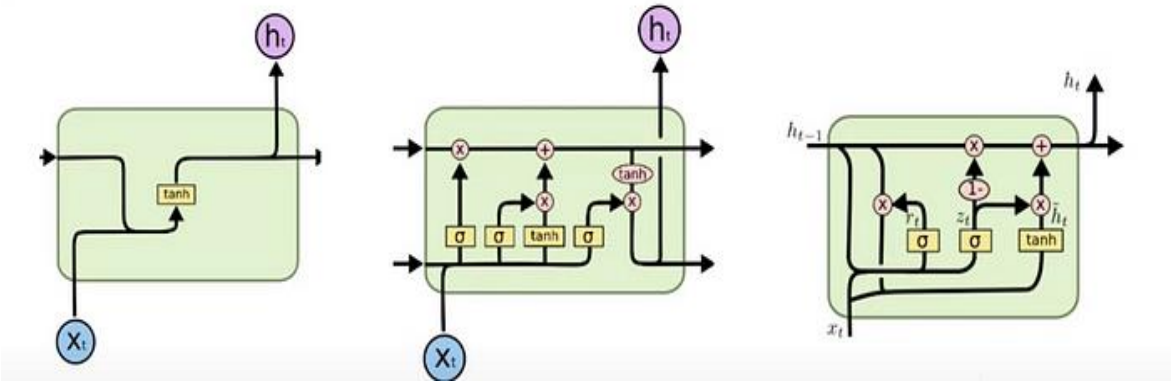

Рис. 1. **Порівняння архітектур RNN (ліворуч), LSTM (по центру), GRU (праворуч), де x – вхідний сигнал, h – вихідний сигнал**

Таблиця 1

Порівняння архітектур LSTM та GRU

Характеристики	LSTM	GRU
складність	3 фільтри, стан комірки	ворота скидання та оновлення
кількість параметрів	більша кількість	менша кількість
представлення пам'яті	наявність стану комірки	оновлення прихованого стану
швидкість навчання	повільніші, краще підходять для довгострокових залежностей	швидші, кращі при роботі з невеликими даними

Висновки. GRU пропонують простішу та ефективнішу в обчислювальному плані альтернативу LSTM для моделювання послідовних даних. Незважаючи на свою компактну архітектуру, GRU демонструють чудову продуктивність у відстеженні часових

залежностей і добре підходять для задач, що вимагають обробки в реальному часі або працюють з обмеженими обчислювальними ресурсами.

Список використаних джерел:

1. What is RNN? - Recurrent Neural Networks Explained - AWS. Amazon Web Services, Inc. URL: https://aws.amazon.com/what-is/recurrent-neural-network/?nc1=h_ls
2. GeeksforGeeks. What is LSTM - Long Short Term Memory? - GeeksforGeeks. GeeksforGeeks. URL: https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/
3. Anishnama. Understanding Gated Recurrent Unit (GRU) in Deep Learning. Medium. URL: https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2