

МЕТОДИ НАВЧАННЯ НЕЙРОНОЇ МЕРЕЖІ НА НЕПОВНИХ ДАНИХ

Ковальова В.Ю.

Науковий керівник – канд. техн. наук, доц. Єсілевський В.С.
Харківський національний університет радіоелектроніки, каф. ПМ,
м. Харків, Україна
e-mail: viktoriia.kovalova@nure.ua

Attention-based deep learning networks enhance handling incomplete and low-quality data, advancing machine learning. The paper explores challenges and methodologies, emphasizing attention mechanisms for improved model accuracy. Incomplete and low-quality datasets in healthcare, finance, and environmental studies pose challenges to traditional models. Attention-based models address this by focusing on relevant data features, minimizing imperfections. Specialized neural network architectures are discussed for enhanced data processing efficiency. Continued research is crucial for optimizing these models, including advanced regularization and data augmentation.

У сучасному світі обсяги даних стрімко зростають, але значна їх частина є неповною або низькоякісною, що ускладнює процес їх аналізу та використання у прийнятті рішень [1]. Механізм уваги в глибокому навчанні дозволяє моделям зосереджуватись на релевантних частинах вхідних даних, покращуючи якість обробки інформації та ефективність використання даних, оскільки модель може "ігнорувати" нерелевантну або відсутню інформацію, зосереджуючись на тій, що має значення для виконання завдання.

Застосування таких мереж може істотно підвищити якість роботи систем у таких сферах, як медицина, фінанси, екологія, де дані часто є неповними або пошкодженими. Це дозволить використовувати штучний інтелект більш ефективно, забезпечуючи високу точність аналізу та прогнозування.

Проблематикою неповних та низькоякісних даних є відсутність частини інформації у датасетах, втрати даних під час передачі або відсутності записів. Дані можуть бути зашумлені, містити помилки або недостовірну інформацію, що ускладнює їх аналіз та інтерпретацію.

Цілями використання механізму уваги для неповних та низькоякісних даних є підвищення точності: здатність моделі точно ідентифікувати та використовувати корисну інформацію з неповних або низькоякісних датасетів [2]. Адаптація до різноманітності даних: моделі повинні бути гнучкими, щоб адаптуватися до різноманітних форматів і якостей даних, з якими вони працюють. Зменшення впливу шуму: ефективне відкидання або мінімізація впливу нерелевантної інформації на процес навчання та висновку.

Для досягнення цих цілей розробляються спеціалізовані архітектури нейронних мереж, які інтегрують механізми уваги на різних рівнях. Такі моделі тренуються на датасетах, які спеціально підготовлені для імітації неповних або низькоякісних умов, з використанням технік, таких як аугментація даних, для покращення їх робастності.

Одним з основних викликів є розробка універсальних методів, які можуть ефективно працювати з широким спектром неповних та низькоякісних даних [3]. Основні завдання включають вдосконалення механізмів уваги для кращої обробки таких даних, розробку нових архітектур нейронних мереж, які зможуть ефективно використовувати доступну інформацію та знижувати вплив шуму та прогалин у даних [4]. Перспективи полягають у подальшому удосконаленні технік глибокого навчання, зокрема механізмів уваги, для підвищення ефективності обробки даних в різних доменах застосування.

Підходи, засновані на механізмі уваги, мають широкий спектр потенційних застосувань, пропонуючи революційні рішення для роботи з неповними та низькоякісними даними у різних галузях, включаючи автоматичний переклад, обробку природної мови, розпізнавання зображень, медичну діагностику та багато інших.

Мережі глибокого навчання на основі механізму уваги відкривають нові перспективи для ефективною обробки неповних та низькоякісних даних, пропонуючи значні переваги у точності, адаптивності та надійності моделей. Однак, для реалізації їх повного потенціалу необхідні подальші дослідження та розробки в цій області.

Список використаних джерел:

1. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. Attention Is All You. 2017. URL:<https://doi.org/10.48550/arXiv.1706.03762>.
2. Tom B. Brown, Benjamin Mann, et al. Language Models are Few-Shot Learners. 2020. URL: <https://doi.org/10.48550/arXiv.2005.14165>.
3. Bengio Y., Courville A., Goodfellow I. Deep Learning. MIT Press, 2016. 800 p.
4. Valentyn Yesilevskyi et al. "Transfer learning in aircraft classification". "INFORMATION SYSTEMS AND TECHNOLOGIES" IST-2019. the 8-th International Scientific and Technical Conference September 9–14, 2019 Kobleve-Kharkiv, Ukraine. p. 132–135.