

АНАЛІЗ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ ПРИРОДНОЇ МОВИ

ANALYSIS OF NEURAL LANGUAGE MODELS

Мовна модель – статистична модель, яка дозволяє для будь-якої послідовності токенів мови визначити ймовірність того, що ця послідовність зустрінеться в природньому застосуванні цієї мови. Корисною властивістю такої моделі є можливість визначити яке продовження послідовності є найімовірнішим, що дозволяє використовувати таку модель в багатьох галузях, пов'язаних з обробкою природної мови - зокрема генерування текстів [1].

Нейронні мережі - це сімейство алгоритмів машинного навчання, які використовують комбінацію з лінійних та нелінійних перетворень даних та стохастичну оптимізацію для апроксимації довільних функцій. Через розвиток апаратних прискорювачів векторних обчислень та значне збільшення доступних обсягів даних, такий підхід дозволив значно покращити якість результатів в багатьох галузях, пов'язаних з когнітивно-складними задачами (машинний зір, класифікація і регресія багатовимірних даних, обробка природної мови тощо).

Найбільшого застосування в обробці природної мови (і мовному моделюванні зокрема) набули архітектури рекурентних нейронних мереж (Recurrent Neural Networks, RNN) та мереж з довго-короткотривалою пам'яттю (Long-Short Term Memory), проте з 2017 року їх витісняють підходи на базі трансформерів (Transformers) [2].

Якість мовних моделей можна оцінити за допомогою перплексії (perplexity), а оскільки ця міра легко обчислюється і не залежить від імплементації мовної моделі, то вона є хорошим інструментом для порівняння різних підходів. Експерименти з великими мовними моделями показують, що використання великих репрезентативних обсягів тексту покращують результати [3], тому тренувальні дані є важливим критерієм аналізу.

Станом на 2021 рік, поширена відсутність вільних даних та сирцевого коду в публікаціях призвела до галузевої кризи відтворюваності [4], тож наявність та доступність даних та коду тренування моделі є важливим критерієм при оцінці підходів до моделювання мови.

Література.

1. A. K. Yadav. Sentence generation from a bag of words using N-gram model [Електронний ресурс] / A. K. Yadav, S. K. Borgohain. - 2014. - Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/7019414>.
2. A. Vaswani. Attention Is All You Need [Електронний ресурс] / A. Vaswani et al. - 2017. - Режим доступу до ресурсу: <https://arxiv.org/pdf/1706.03762.pdf>.
3. A. Radford. Language Models are Unsupervised Multitask Learners [Електронний ресурс] / A. Radford et al. - 2017. - Режим доступу до ресурсу: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.
4. M. Hutson. Artificial intelligence faces reproducibility crisis. / M. Hutson. - 2018. - Режим доступу до ресурсу: <https://www.science.org/doi/10.1126/science.359.6377.725>.