

Джерела та література

1. Гурчунова Є. Що таке SMM, як скласти стратегію та просувати компанію в соціальних мережах. URL: <https://aboutmarketing.info/internet-marketynh/shcho-take-smm-yak-sklasty-stratehiyu-ta-prosuvaty-kompaniyu-v-sotsialnykh-merezhakh/>

2. Соцмережі та війна – особливості ведення бізнес-акаунтів під час воєнного стану. URL: <https://odesseo.com.ua/smm-and-war/>

Антон Березний

Наук. керівник – к. пед. н., доцент Агейчева А.О.,

м. Полтава

NMT-СИСТЕМИ У МАШИННОМУ ПЕРЕКЛАДІ: ТЕОРІЯ І ПРАКТИКА

Сьогодні у світі широко застосовуються технології штучного інтелекту. Не обійшли вони увагою і таку галузь як перекладацька діяльність. Нейронні мережі машинного перекладу – це і є штучний інтелект. Навчені нейронні мережі вже сьогодні якісно перекладають тексти замість людей за основними мовними парами. Кількість мов, які можуть перекладатися за допомогою нейронно-мережових систем стрімко розширюється.

Нейронний машинний переклад (NMT) – це алгоритм, який використовується для перекладу слів з однієї мови на іншу. На сьогодні вважається, що високоякісний NMT може визначити контекст перекладу та використовувати моделі, щоб запропонувати більш точний переклад. Google Translate, DeepL, Yandex, Baidu Translate є добре відомими прикладами NMT, доступними для громадськості через автоматичний переклад браузера.

NMT-моделі застосовують глибоке навчання та навчання ознакам. Для їхньої роботи необхідна лише частка пам'яті для використання традиційних

моделей статистичного машинного перекладу (SMT). Крім того, на відміну від звичайних систем перекладу, усі частини моделі нейронного перекладу навчаються разом від початку до кінця (end-to-end), що дає змогу максимізувати продуктивність перекладу [1;2;3].

Уперше програми глибокого навчання почали застосовувати в 1990-х роках у задачах із розпізнавання мовлення.

Перша наукова стаття про використання нейронних мереж у машинному перекладі з'явилася 2014 року, а впродовж кількох наступних із її допомогою вдалося вирішити чимало інших задач. Станом на 2017 рік це такі системи:

- NMT-системи з великим словником (*Large-Vocabulary NMT*)
- NMT-системи з урахуванням частин слів (*Subword-NMT*)
- Багатомовні NMT-системи (*Multilingual NMT*)
- NMT-системи з багатьма джерелами (*Multi-Source NMT*)
- NMT-системи без джерел (*Zero-Resource NMT*)
- NMT-системи на рівні символів (*Character-dec NMT*)
- Повносимвольні NMT-системи (*Fully Character-NMT*)
- NMT-системи без підготовки (*Zero-Shot NMT*)
- Google
- Додаток для захоплення зображень.

У 2015 році NMT-система вперше з'явилася на відкритому конкурсі машинного перекладу (OpenMT '15). На WMT '15 теж уперше взяв участь NMT-претендент; наступного року серед переможців було вже 90 % NMT-систем. Від 2017 року Європейське патентне відомство використовує нейронний машинний переклад для миттєвого надання доступу до інформації, вміщеної у глобальній патентній системі.

NMT-система є фактично подальшим розвитком статистичного підходу на рівні фраз, який функціонує на базі незалежних підкомпонентів [4]. Нейронний машинний переклад (NMT) не є кардинальною відмовою від статистичного машинного

перекладу (SMT). Його основним вихідним пунктом є використання векторного подання слів і внутрішніх станів — «вбудовування» (embeddings), «подання в неперервному просторі» (continuous space representations).

Структура NMT-моделей простіша порівняно зі структурою моделей на основі фраз: у ній немає окремої мовної моделі, моделі перекладу та моделі перевпорядкування, а є лише одна модель послідовностей, яка передбачає одне слово за раз. Однак це передбачення послідовності слів спирається одразу на все вихідне речення та на всю вже створену цільову послідовність.

Перші спроби моделювання послідовності слів зазвичай проводилися за допомогою рекурентної нейронної мережі (RNN). Двонаправлена рекурентна нейронна мережа, так званий *кодувальник* (encoder), використовується для кодування вихідного речення для другої RNN, відомого як *розкодувальник* (decoder), а та, відповідно, використовується для передбачення слів цільовою мовою [5].

Перед рекурентними нейронними мережами постають труднощі під час кодування довгих вхідних даних в один вектор. Їх можна нівелювати за допомогою механізму уваги (attention mechanism) [6], який дає декодувальнику можливість зосереджуватися на різних частинах вхідних даних під час генерації кожного вихідного слова. Існують моделі покриття (coverage models) для вирішення проблем у таких механізмах уваги, які призводять до генерування надто довгого або надто короткого перекладу [7] (наприклад, ігнорування наявної інформації про вирівнювання). Порівняно краще обробляють довгі неперервні послідовності згорткові нейронні мережі (convnets), але певний час їх не використовували через наявність кількох недоліків. У 2017 році ці недоліки вдалося успішно подолати за допомогою «механізмів уваги».

Домінувальною архітектурою для кількох мовних пар залишається модель Transformer, яка ґрунтується на механізмі уваги. На рівнях «самоуваги» цієї моделі досліджуються залежності між словами послідовності шляхом аналізу зв'язків

між усіма словами в парних послідовностях і безпосереднього моделювання цих зв'язків. Цей підхід простіший, ніж механізм селекції, на якому працюють рекурентні нейронні мережі. А його простота дала дослідникам змогу розробляти високоякісні моделі перекладу за допомогою моделі Transformer, навіть в умовах, коли кількість вхідних даних невелика.

Джерела та література

1. *Kalchbrenner, Nal; Blunsom, Philip (2013). Recurrent Continuous Translation Models. Proceedings of the Association for Computational Linguistics: 1700 – 1709. URL: <https://aclanthology.org/D13-1176/>*
2. *Sutskever, Ilya; Vinyals, Oriol; Le, Quoc Viet (2014). «Sequence to sequence learning with neural networks». URL: <https://arxiv.org/abs/1409.3215>*
3. *Kyunghyun Cho; Bart van Merriënboer; Dzmitry Bahdanau; Yoshua Bengio (3 September 2014). «On the Properties of Neural Machine Translation: Encoder–Decoder Approaches». URL: <https://arxiv.org/abs/1409.1259>*
4. *Wolk, Krzysztof; Marasek, Krzysztof (2015). Neural-based Machine Translation for Medical Text Domain. Based on European Medicines Agency Leaflet Texts. Procedia Computer Science. 64 (64): 2—9. URL: <https://arxiv.org/abs/1509.08644>*
5. *Bahdanau Dzmitry; Kyunghyun Cho; Bengio Yoshua (2014). «Neural Machine Translation by Jointly Learning to Align and Translate». URL: <https://arxiv.org/abs/1409.0473>*
6. *Bahdanau, Dzmitry; Cho, Kyunghyun; Bengio, Yoshua (2014-09-01). «Neural Machine Translation by Jointly Learning to Align and Translate». URL: <https://arxiv.org/abs/1601.04811>*
7. *Coldewey, Devin (29 серпня 2017). DeepL schools other online translators with clever machine learning. URL: <https://techcrunch.com/2017/08/29/deepl-schools-other-online-translators-with-clever-machine-learning/>*