

УДК 004.85:[004.738.5:339]

**ДОСЛІДЖЕННЯ КОМБІНУВАННЯ НАВЧАННЯ З
ПІДКРІПЛЕННЯМ ТА МОВНИХ АГЕНТІВ ДЛЯ РЕАЛІЗАЦІЇ
ДІАЛОГОВИХ АГЕНТІВ**

Бовдуй Р.В.

Науковий керівник – к.т.н., доц. Кулішова Н.Є.

Харківський національний університет радіоелектроніки, каф. ШІ

м. Харків, Україна

email:bovdui.roman@nure.ua

The chatbot market has been growing rapidly lately. They allow you to optimize business processes and meet user needs. With the advent of large language models based on transformers, interest in the end-to-end approach to building such systems has increased. However, when building task-oriented dialog systems based on transformers, there is no certainty that they will pursue the intended goal, and here reinforcement learning comes to the rescue.

The paper discusses the combination of these approaches in order to improve the characteristics of task-oriented dialogue systems which are needed for business.

Таск-орієнтовані діалогові системи (TOD) це один з типів діалогових агентів, де агент намагається зрозуміти мету споживача та спробувати її реалізувати. Діалог може відбуватися через зручні інтерфейси: текстові, голосові, графічні, та інші. Такого рода агенти мають дуже широке застосування у різноманітних сферах, наприклад, в застосунках, які обробляють звернення громадян: системи бронювання білетів, персональні помічники, віртуальні медичні консультанти або фінансові консультанти.

Згідно різноманітним оглядам ринку, популярність чат-ботів постійно зростає. Завдяки такого рода системам компанії не тільки зберігають свої гроші, а і спрямовують робочу силу для вирішення більш складних завдань, тобто оптимізують бізнес процеси, що допомагає боротися з конкурентами. Відповідно до прогнозів дослідницької компанії Grand View Research ринок діалогових агентів значно зростає від 190.8 мільйонів доларів США у 2016 році до 1.25 мільярдів доларів США у 2025.

Архітектури таск-орієнтованих систем можуть бути класифіковані на два типи. Перший з них це end-to-end підхід, коли запити юзера відображаються на відповіді моделі. Зазвичай у такому підході використовуються Seq2Seq моделі, які навчаються за допомогою контрольованого навчання. Цей підхід більш гнучкий та має менше вимог до анотації вхідних даних, але загалом потребує більший об'єм даних. Непрозора структура не дозволяє інтерпретувати результати та не дає ніякого контролю.

Інший, пайплайн підхід, коли система декомпозується на кілька послідовних підзавдань: спочатку працює модуль розуміння природної

мови, який відповідає за розпізнавання намірів користувача та слотів із вхідної послідовності. Потім працює модуль відстеження стану розмови, який запам'ятовує внутрішній стан діалогу. Наступний модуль вивчення стратегії діалогу виконує дії, які задовольняють вимоги користувача. Наприкінці працює модуль генерування природної мови, який відповідає за трансформацію дії у відповідь користувачу. Цей підхід вважається легшим в інтерпретації та легшим в імплементації, але в цілому систему складніше оптимізувати глобально [1].

Останнім часом набуває популярність перший, end-to-end підхід. Це обумовлено появою різноманітних механізмів уваги та так званих трансформерів, моделей, архітектури яких покладаються на механізми уваги. За допомогою механізмів уваги, ці мовні моделі дозволяють концентруватися на різних частинах вхідної послідовності, що дозволяє ефективно обробляти речення та тексти різної довжини. Ці моделі гарно дозволяють розпаралелити обчислення та працюють доволі швидко, тому що в їх архітектурі не використовуються рекурентні клітини. Прикладами найбільш відомих архітектур є GPT, BERT, T5 [2].

Трансформери гарно зарекомендували себе у генерації текстів, перекладі, відповіді на питання на основі знань та іншому. Однак, незважаючи на всі ці успіхи, при побудові таск-орієнтованих діалогових систем не можна бути впевненими, що такі моделі будуть переслідувати ціль, поставлену даному агенту. Алгоритми навчання з підкріпленням, по своїй суті, спроектовані таким чином, щоб досягати певної мети, але через те, що алгоритми навчаються на абстрактних станах та діях, цим алгоритмам не вистачає здатності генерувати натуральні відповіді [3].

Виходячи з вищесказаного, постає питання в поєднанні мовних моделей та навчання з підкріпленням, для того, щоб отримати переваги одного підходу і нівелювати недоліки іншого, і навпаки. Це надасть змогу побудувати більш потужні таск-орієнтовані діалогові системи які будуть дійсно корисним продуктом на ринку і можуть задовольнити потреби бізнесу.

Список використаних джерел:

1. W.-C. Kwan, H. Wang, H. Wang, K.-F. Wong. A Survey on Recent Advances and Challenges in Reinforcement Learning Methods for Task-Oriented Dialogue Policy Learning. The Chinese University of Hong Kong. Jul 2022. arXiv preprint arXiv:2202.13675.

2. J. Chung, C. Gulcehre, K.H. Cho, Y. Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. Dec 2014. arXiv preprint arXiv:1412.3555.

3. F. AlMahamid, K. Grolinger. Reinforcement Learning Algorithms: An Overview and Classification. – Department of Electrical and Computer Engineering Western University. 29 Sep 2022. arXiv preprint arXiv:2209.14940.