

**А.О. Колесник**, канд. техн. наук, доц. (ДБТУ, Харків)

**М.І. Крупей**, ст. викл. (ДБТУ, Харків)

**О.М. Муравйова**, ст. викл. (ДБТУ, Харків)

## **ВИКОРИСТАННЯ ІНСТРУМЕНТІВ КОМП'ЮТЕРНОЇ ЛІНГВІСТИКИ ДЛЯ ВИВЧЕННЯ ОСНОВ ПЕРЕКЛАДУ МАЙБУТНІМИ ФАХІВЦЯМИ СФЕРИ УПРАВЛІННЯ**

Готовність до ділового спілкування та розуміння документів, укладених іноземною мовою, стає важливим показником кваліфікації спеціаліста будь-якої сфери трудової діяльності. Особливо це стосується навчання професіоналів, чия галузь передбачає теорію та практику управління людьми та організаціями.

Потреба у мовній підготовці конкурентоспроможних менеджерів набуває істотного значення і є складовою загальної системи їх професійного навчання в закладі вищої освіти. Ця необхідність значно ускладнює вимоги до іноземної мови як навчального предмета, оскільки викликає необхідність активізації іншомовної освіти шляхом застосування активних методів навчання. Особливого значення у професійній підготовці фахівців з управління та адміністрування набуває навчання основ перекладу.

Зважаючи на тенденцію молодого покоління до використання машинного перекладу, одним із завдань викладача є необхідність сформулювати в студента уявлення про ті можливості, які дають електронні інструменти для полегшення пошуку необхідних інформаційних ресурсів у наукових та професійних цілях та навчання необхідним для цих цілей видам перекладацької діяльності.

Розглянемо загальні інструменти комп'ютерної лінгвістики, які можна використовувати для вирішення поставлених завдань.

1. Семантичні мережі: WordNet, CYC, Free-base/Wikidata та ін. Багато з цих ресурсів спочатку були розроблені для англійської мови. Деякі були розширені для роботи з іншими мовами, наприклад, BabelNet 4.021 підтримує 284 мови. Багато з них можна завантажити безкоштовно.

Робота даних ресурсів базується на моделюванні слів і понять, які є елементами однієї схеми (одного графа) та відносини яких представлені як синоніми, антоніми та гіпоніми. Існує безліч назв таких графів: графи знань, семантичні мережі, онтології тощо. Ці

графи можуть використовуватися для представлення лексичних знань та уявлень про світ, що виходять за межі лінгвістики [1].

2. Глибокі (нейронні) мережі – це «клас алгоритмів машинного навчання, які навчаються глибше (абстрактніше) розуміти дані» [2]. Наразі такі мережі показують велику ефективність. На вхідний вузол нейронної мережі подаються будь-які мовні дані у закодованому вигляді: літери, короткі високочастотні морфеми, біграми та слова. Застосування цього підходу залежить від наявності у розпорядженні дослідника великого масиву анотованих текстів.

Моделі глибоких мереж, таких як BERT, GPT, ELMo, ERNIE досягли рекордних результатів під час вирішення широкого кола завдань. Найбільш розробленим алгоритмом є системи NER (Name Entity Recognition, NER), тобто фізичних осіб, організацій, географічних об'єктів тощо. Системи NER та BERT, засновані на словниках і правилах, можна порівняти з рівнем людської точності.

Системи нейронного машинного перекладу використовуються основними онлайн-перекладачами (Google Translate, Bing Microsoft, DeepL, Amazon Translate, Yandex Translate). Однак ці системи досі мають багато проблем.

По-перше, залишається проблема перекладу рідкісних слів та абревіатур. По-друге, іноді хороша мова перекладу може «розслабити» читача, і серйозна помилка в перекладі залишиться непоміченою, що може призвести до грубих помилок та спотворення сенсу. По-третє, більшість сучасних систем перекладають речення, не враховуючи загального контексту.

Найбільш успішною розробкою в цій галузі є онлайн перекладач DeepL, особливо його платні, розширені версії, проте переклад, виконаний у контексті, також потребує ретельної перевірки та аналізу з боку фахівця-мовника.

Для перекладу вузькоспеціалізованої автентичної лексики іноземною мовою, доцільно виділити наступні джерела: електронні словники, серед яких найбільшою популярністю користуються Мультитран, Оксфордський словник Мюррея, засоби машинного перекладу, які виконують дії з перетворення тексту для перекладу на еквівалентний за змістом текст рідною мовою: онлайн-словник Promt, Яндекс перекладач, DEEPL.

Системи Google Translate та Яндекс засновані на статистичному обчисленні ймовірності збігів та використовують безліч баз паралельних текстів, у яких попарно зберігаються словосполучення та їх переклади. У процесі перекладу здійснюється статистичний аналіз: система підбирає еквівалент для перекладу, ґрунтуючись на частоті

вживання та підставляє варіант, що має найвищий відсоток збігів. Однак для кожної мови характерні свої особливості лексико-семантичної сполучуваності. Найчастіше переклад, виконаний цими програмами, свідчить про те, що системи не узгоджують слова. Обидві системи мають проблеми з правильним перекладом термінів, розпізнаванням мовних частин. Трапляються і граматичні помилки.

DeerL – один із найточніших онлайн-перекладачів. Його найбільшими перевагами є безпека та точність. Усі тексти, перекладені DeerL, видаляються відразу після перекладу, тому збереження конфіденційних даних гарантовано. Крім того, для забезпечення додаткової безпеки DeerL може здійснювати наскрізне шифрування даних, що повністю відповідає політиці ЄС щодо захисту даних. У порівнянні з конкурентами DeerL перевершив їх у співвідношенні 3:1, тому що працює на основі передової технології штучного інтелекту (AI) та надає безліч варіантів налаштування.

За бажанням користувачі можуть створити свій власний глосарій, щоб визначити переклад конкретних слів та фраз відповідно до їхнього бренду або проекту. Крім того, DeerL має вбудований словник, що дозволяє шукати слова та самостійно вивчати їх переклад.

Всі преміум-плани DeerL пропонують необмежену кількість текстових перекладів та дозволяють додавати двох або більше користувачів до облікового запису для спільної роботи в тандемі. Також ви можете інтегруватися з вашим CAT-інструментом, таким як Trados Studio, memoQ та Across.

Незважаючи на всі переваги штучного інтелекту, на даний момент немає ідеального словника, який міг би з точністю осмислити всі тонкощі перекладу вузькоспеціалізованих текстів. Тому в рамках роботи над перекладом студентам необхідно розвивати навички критичного осмислення перекладу та вміння користуватися кількома джерелами одночасно.

#### **Інформаційні джерела:**

1. Mikolov T., Chen K., Corrado G., Dean, J. Efficient Estimation of Word Representations in Vector Space, 2013. URL: <https://arxiv.org/abs/1301.3781>

2. Методи та технології обчислювального інтелекту: Навчальний посібник [Електронний ресурс] : навч. посіб. для студ. спеціальності 122 «Комп'ютерні науки» / І. В. Федорін; КПІ ім. Ігоря Сікорського. Київ: КПІ ім. Ігоря Сікорського, 2022. 314 с. URL: <https://ela.kpi.ua/bitstream/123456789/50934/1/MethodsCompIntell.pdf>