

**СУЧАСНІ МОДЕЛІ ТА ІНСТРУМЕНТАЛЬНІ ЗАСОБИ У
КОНТЕКСТІ ПЕРСПЕКТИВ ЛІНГВІСТИЧНИХ ДОСЛІДЖЕНЬ.
MODERN MODELS AND TOOLS WITHIN A FRAMEWORK OF
LINGUISTIC RESEARCH.**

Н. Козьмина, Л. Гурєєва

*Старший преподаватель кафедры английского языка технической
направленности №1, Национальный технический университет Украины
«Киевский политехнический институт»*

*Преподаватель кафедры английского языка технической направленности
№1, Национальный технический университет Украины «Киевский
политехнический институт»*

Abstract. Linguistics is interested in the creation and application of advanced models and tools for applied purposes.

Among the most common and widely used are neural networks and artificial intelligence. Artificial neural network, often simply called neural network is a mathematical model, the prototype of which is a biological neural network.

Лінгвістика цікавиться питаннями створення та застосування сучасних моделей та інструментальних засобів з прикладною метою.

Серед найбільш поширених та застосовуваних є нейронні мережі та штучний інтелект. Штучна нейронна мережа, яку часто називають просто нейронною мережею, являє собою математичну модель, прототипом якої є біологічні нейронні мережі. Нейронна мережа складається зі взаємозв'язаних груп штучних нейронів і обробляє інформацію, використовуючи нейромережевий підхід до обчислення. У більшості випадків нейронні мережі представляють собою адаптивну систему, яка змінює свою структуру на послідовних етапах навчання. Нейронні мережі зокрема використовуються для пошуку закономірності в даних[1]. Ця функція цікавить лінгвістів, що вже довгий час намагаються знайти способи, які б надали можливість науковцям у

різних галузях працювати з іншомовними текстами, не знаючи іноземної мови та не вдаючись до допомоги спеціалістів.

Так був виділений напрямок, загальний для лінгвістики і штучного інтелекту, - обробка природної мови (ОПМ). Він вивчає проблеми комп'ютерного аналізу та синтезу природної мови. Стосовно штучного інтелекту аналіз означає розуміння мови, а синтез – створення розумного тексту. Розв'язок цих проблем буде означати створення ефективного інтерфейсу - зручнішої форми взаємодії комп'ютера та людини.

Нижче наводиться список найчастіше досліджуваних задач в ОПМ. Деякі з цих завдань мають пряме застосування, у той час як інші найчастіше служать допоміжними компонентами, що використовуються для вирішення набагато складніших завдань. Головними завданнями для обробки природної мови є: видобування даних, синтез мови, розпізнавання мови, синтез тексту, машинний переклад, питально-відповідальні системи, інформаційний пошук, отримання інформації, спрощення тексту, технологія перекладу та автоматичне реферування[2]. Це тільки незначна кількість завдань, які стоять перед дослідниками, що працюють у цій галузі.

Розуміння природної мови іноді вважають AI-повним завданням («AI-complete»), таким, що легко вирішується людиною, але досі не може вирішуватись комп'ютерними системами, тому що розпізнавання живої мови вимагає дуже повних знань системи про навколишній світ і можливості з ним взаємодіяти. Саме визначення сенсу слова «розуміти» - одна з найголовніших завдань штучного інтелекту. Якість розуміння залежить від безлічі факторів: від мови, від національної культури, від типу тексту або від співрозмовника і т. д.[3]

І хоча реальні дослідження у цій сфері почалися ще у 1950-х, а зацікавленість цим питанням з'явилась набагато раніше, наразі вчені, дослідники, університети та промислові компанії різних країн активно працюють над цією задачею. А її вирішення залежить в першу чергу від плідної співпраці фахівців різних галузей. Це відкриває широкі перспективи для

лінгвістичних досліджень та створення машин, здатність яких вчитися і запам'ятовувати буде нагадувати процес мислення людини, що допоможе вирішенню багатьох прикладних задач.

Серед основних інструментальних засобів, які сьогодні використовують дослідники лінгвістичних завдань, слід виділити мови програмування високого рівня (JAVA, PYTHON та ін.) та системи керування реляційними базами даних (ORACLE, MySQL та ін.).

В реляційних таблицях баз даних формуються вузли (слова) та гілки (відношення слів), які в процесі ОПМ використовуються для формування знань у вигляді семантичних мереж. Семантична мережа є одним із способів представлення знань. У назві сполучені терміни з двох наук: семантика у мовознавстві вивчає сенс одиниць мови, значення слів, а мережа в математиці є різновидом графу — набору вершин, сполучених дугами (гілками, ребрами, хордами)[4]. У семантичній мережі роль вершин виконують головні (опорні) поняття та об'єкти бази знань про предметну царину, а дуги (причому направлені) задають відношення між ними[5]. Таким чином, семантична мережа відображає семантику предметної царини у вигляді понять і відносин між ними. Відносини можуть бути кількісними, функціональними, просторовими, логічними, лінгвістичними та ін. Вперше семантичні мережі з'явилися в області машинної лінгвістики як засіб аналізу змісту (семантики) природної мови. Сама ідея семантичної мережі - тобто, мережі асоціативно пов'язаних понять – дуже стара, і деякі вчені стверджують, що її почав розвивати ще Аристотель. Але сам термін семантична мережа вперше з'явився у дисертації М.Р. Куїлліана (1968), в якій він представив це як спосіб говорити про організацію людської семантичній пам'яті[6].

Таке уявлення змісту предметної царини дозволяє ефективно вирішувати весь спектр лінгвістичних завдань та послідовно формувати відповідну базу знань.

БІБЛІОГРАФІЯ

1. http://en.wikipedia.org/wiki/Artificial_neural_network

2. http://en.wikipedia.org/wiki/Natural_language_processing
3. http://en.wikipedia.org/wiki/Language_model
4. http://en.wikipedia.org/wiki/Semantic_network
5. <http://knhelp.wordpress.com/2012/04/19/>
6. <http://www.cs.bham.ac.uk/research/projects/poplog.html>

ВІДОМОСТІ ПРО АВТОРІВ

1. Гурєєва Л.В., викладач кафедри англійської мови технічного спрямування №1, Національного технічного університету України «КПІ».

2. Козьміна Н.А., ст. викладач кафедри англійської мови технічного спрямування №1, Національного технічного університету України «КПІ».