

NATURAL LANGUAGE PROCESSING (NLP): BRIDGING THE GAP BETWEEN HUMAN AND MACHINE

Vinnytsia National Technical University

Анотація

Публікація досліджує роль обробки природної мови (Natural Language Processing, NLP) у взаємодії між людиною та комп'ютером. Розглянуто основні принципи роботи NLP, а також його застосування у сучасних технологіях, таких як чат-боти, голосові помічники та автоматичний переклад. Особливу увагу приділено використанню машинного навчання та великих мовних моделей для аналізу і розуміння людської мови..

Ключові слова: обробка природної мови, NLP, машинне навчання, штучний інтелект, чат-боти.

Abstract

This article explores the role of Natural Language Processing (NLP) in enabling interaction between humans and computers. The study discusses the basic principles of NLP and its applications in modern technologies such as chatbots, voice assistants, and automatic translation. Special attention is given to the use of machine learning and large language models for analyzing and understanding human language..

Key words: natural language processing, NLP, machine learning, artificial intelligence, chatbots.

Introduction

In recent years, communication between humans and computers has become increasingly natural thanks to the development of Natural Language Processing (NLP). NLP is a field of artificial intelligence that focuses on enabling machines to understand, interpret, and generate human language. Historically, early NLP systems relied on rigid, rule-based algorithms that required manual coding of grammar rules.

However, human language is highly unstructured, emotional, and full of nuances, which sharply contrasts with binary machine logic. With the rapid growth of digital information, modern NLP has shifted toward statistical and neural network-based approaches, acting as a vital bridge between human flexibility and machine precision.

Research results

Natural Language Processing combines techniques from computer science, applied mathematics, and linguistics to decode human language. The process of bridging the communication gap begins with a complex data pipeline. Before a machine can "understand" text, the raw data must undergo extensive preprocessing. This includes tokenization (breaking text into smaller units like words or subwords), lemmatization (reducing words to their base form), and the removal of syntactic noise[1]. Once the text is cleaned, the central challenge of NLP arises: converting qualitative human language into quantitative data that a computer can process. Modern systems achieve this through Word Embeddings (such as Word2Vec or GloVe). These algorithms transform vocabulary into continuous, high-dimensional numerical vectors.

By mapping words into a vector space, semantic relationships between them become measurable mathematical distances. Consequently, a computer comprehends language through linear algebra, utilizing vector addition and matrix multiplication to calculate context. For years, the industry standard for processing these sequences was Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models. While effective for short sentences, they processed data sequentially, leading to a loss of context in longer texts.

A monumental breakthrough occurred with the introduction of the Transformer architecture, which replaced sequential processing with parallel processing. The core innovation of the Transformer is the "Self-Attention" mechanism. Instead of reading word-by-word, this architecture evaluates the entire sequence simultaneously. This specific matrix operation allows the AI [2] to dynamically assign different weights to

different words, determining which part of the context is the most critical, effectively mimicking human reading comprehension. The practical applications of this mathematical framework are transformative. Beyond standard machine translation and sentiment analysis used in marketing, NLP is revolutionizing software engineering itself. Large Language Models (LLMs) trained on massive repositories of source code can now parse natural language prompts to automatically generate functional scripts. Tools like AI-assisted code generators perform syntax highlighting, predict code completion, automate code refactoring, and even write technical documentation. They treat programming languages (like Python or C++) simply as another highly structured human language.

Despite these massive leaps, bridging the gap remains an ongoing challenge. Current NLP models face significant computational barriers; training massive matrices with billions of parameters requires enormous processing power and energy consumption [3]. Furthermore, advanced AI systems frequently struggle with "hallucinations"—generating mathematically probable but factually incorrect information. They can also inherit cognitive and cultural biases present in their training datasets, raising serious ethical concerns regarding their deployment in critical systems.

Conclusion

To sum up, Natural Language Processing is rapidly and permanently closing the communication gap between humans and machines. By transforming unstructured human text into structured mathematical models of vectors and matrices, NLP allows computers to interact with us in a deeply context-aware manner. From basic linguistic tools to advanced AI coding assistants that reshape the software development life cycle, this technology has integrated into both our daily routines and professional engineering workflows. Addressing the current computational limitations and algorithmic biases will be the next crucial frontier. As these neural architectures become more sophisticated, the boundary between human thought and machine processing will continue to blur, redefining the future of digital interaction.

REFERENCES

1. Natural Language Processing Overview – <https://www.ibm.com/topics/natural-language-processing>
2. Machine Translation in NLP – <https://towardsdatascience.com/machine-translation>
3. Voice Assistants and NLP – <https://www.techtarget.com/whatis/definition/natural-language-processing-NLP>

Коваль Вероніка Анатоліївна – студентка групи 2ПІІ-24б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, e-mail: nika0678910860@gmail.com

Науковий керівник: **Кухарчук Галина Вікторівна** – викладач англійської мови, кафедра іноземних мов, Вінницький національний технічний університет, м. Вінниця, e-mail: galinakuh07@gmail.com.

Koval Veronika – a student of 2PІІ-24b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: nika0678910860@gmail.com

Scientific Supervisor: **Halyna Kukharchuk** – teacher of English, Foreign Languages Department, Vinnytsia National Technical University, Vinnytsia, e-mail: galinakuh07@gmail.com.