

NATURAL LANGUAGE PLATFORMS

Paramonova A.E.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Rogachevskaya A.I. – Senior Lecturer

Annotation. Natural language processing is a computer's ability to understand and process human language. In the realm of virtual Assistants, NLP is used to determine a user's intention, extract information from an utterance, and carry on a conversation with the user in order to execute and complete a task. The article discusses the main methods and algorithms of work that are used in NLP, basic functions performed, examples of various bot architectures. The results of the implementation of such systems and their benefits for consumers and companies are considered.

Keywords. NLP (Natural Language Programming), NLU (Natural Language Understanding), AI (artificial intelligence), neural networks, chatbot, platform.

Natural language processing (NLP) is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data. As a result we must have a computer capable of "understanding" the contents of documents, including the contextual nuances of the language within them. Then the technology can accurately extract information and facts contained in the documents as well as categorize and systematize the documents themselves. Challenges in natural language processing frequently involve speech recognition, natural language understanding, and natural-language generation. There are numerous API providers in the chatbot landscape, the majority of them are focusing on Natural Language Programming (NLP) and Natural Language Understanding (NLU). They are the most important steps to be done as they will be handling the important stage in a conversational interface [1].

In the artificial intelligence (AI) context, NLP is the overarching discipline that encompasses several spheres dealing with the interaction between computer systems and human natural languages. NLP includes several sub-disciplines such as discourse analysis, relationship extraction, natural language understanding and a few other language analysis areas.

NLU is a subset of NLP that focuses on reading comprehension and semantic analysis. The combination of NLP and NLU technologies is becoming increasingly relevant in different modern software areas including bot technologies. While there are many vendors and platforms focused on NLP-NLU technologies, such technologies are becoming extremely popular within the bot developer community [2].

Recent systems based on machine-learning algorithms have many advantages over hand-produced algorithms. The learning procedures used during machine learning automatically focus on the most common cases, whereas when writing rules by hand it is often not very obvious where the efforts should be directed. Automatic learning procedures can make use of statistical inference algorithms to produce models that are robust to unfamiliar input (e.g. containing words or structures that have not been seen before) and to erroneous input (e.g. with misspelled words or words accidentally omitted). Usually handling such input data with handwritten rules, or creating systems of handwritten rules that make soft decisions, is extremely difficult, error-prone and time-consuming. Systems based on automatic rules learning can be made more accurate simply by supplying more input data. However, systems based on handwritten rules can only be made more accurate by increasing the complexity of the rules and that is a much more difficult task. In particular, there is a limit to the complexity of the systems based on handwritten rules beyond which the systems become more and more unmanageable. However, generating more input data to machine-learning systems simply requires a corresponding increase in the number of man-hours worked, generally without significant increases in the complexity of the annotation process.

Some performed tasks have direct real-world applications, while others more commonly serve as subtasks that are used to aid in solving larger tasks. Though natural language processing tasks are closely intertwined, they can be subdivided into categories for convenience: text and speech processing, syntactic analysis, lexical semantics (of individual words in context) and so on.

References:

1. Speech and Language Processing/ D.Jurafsky [et al.] // California, 1994
2. Handbook of Natural Language Processing / J. Damerau [et al.] // 2010
3. Text Mining with R/ S. Robinson // 2007
4. Neural Network Methods in Natural Language Processing (Synthesis Lectures on Human Language Technologies)/ Yoav Goldberg , Graeme Hirst// 2017
5. Taming Text/ Grant Ingersoll, Thomas Morton, Drew Farris// 2013.