



## **TECHNOLOGY FOR CREATING A COMPUTER PROGRAM FOR DETECTING PUNCTUATION ERRORS**

**Jumayeva Dilnoza Bakhshulloyevna**

**Navoi State Pedagogical Institute, Faculty of Preschool and Primary  
Education, Senior Teacher of Primary Education f.f.f.d(PhD)**

**Abstract:** This article provides general information about the technology of creating a computer program that detects punctuation errors, and it is emphasized that the detection and correction of punctuation errors is important not only for the grammatical correctness of the language, but also for the intelligibility and communicative effectiveness of the text. It is also explained that punctuation error detection software can serve as an important addition to various text analysis platforms and applications.

**Keywords:** linguistics, information technology, morphological, syntactic and semantic aspects, punctuation, punctuation errors, morphological and syntactic analysis, language models and algorithms.

Today, natural language processing (NLP) issues are rapidly developing in the field of linguistics and information technologies. Technologies for automatic analysis of morphological, syntactic and semantic aspects of the text are of great importance in improving the quality of communication between a person and a computer. Among such technologies, programs that identify punctuation errors are distinguished by their relevance.

Identifying and correcting punctuation errors is important not only for the grammatical correctness of the language, but also for the comprehensibility and communicative effectiveness of the text. While it is important to follow the rules of punctuation in writing literature, journalism, scientific articles and working with everyday business documents, failure to follow these rules can lead to misinterpretations in the communication process. Especially for language learners and professionals who create texts in various fields, the possibility of automatic analysis and correction of errors with the help of such programs significantly increases labor productivity.

Also, punctuation detection software can serve as an important addition to various text analysis platforms and applications. In the conditions of modern information and communication technologies, such software algorithms are one of



the most necessary tools for language processing due to the demand for large-scale processing of texts.

The development of this technology allows, firstly, a deeper study of the syntactic and semantic rules of the language with the help of a computer; secondly, it serves to increase the culture of written communication by applying it in research activities and in the educational process. Therefore, the development of technologies for automatic detection of punctuation errors is today considered as an urgent direction not only in linguistics, but also in the field of software.

Software includes various algorithms, databases, modeling techniques and software tools. This process transforms natural language from an unstructured state into a structured one, and enables the transformation of text into information. As a result, tasks such as identifying semantic connections between texts, following the rules of punctuation in sentences or finding grammatical errors can be performed.

Language processing (NLP) mainly consists of the following steps:

**Text pre-processing:** At this stage, processes such as receiving texts in digital form, tokenization (breaking the text into words or sentences), removing stop words, lemmatization or stemming (returning words to their original form) are performed.

**Morphological and syntactic analysis:** Determining the role of each word in the text and understanding their relationship. Here, the category of words and the grammatical structure are analyzed. This process is essential for spotting punctuation and grammatical errors.

**Semantic Analysis:** This step focuses on determining the meaning of a sentence or text. Semantic analysis is used to understand the relationships between words and phrases, which plays a key role in determining the correct punctuation rules.

**Language models and algorithms:** When identifying punctuation errors, the grammatical structure and rules of words and sentences are controlled with the help of language models. Various software models such as machine learning technologies, rule-based algorithms, and statistical methods are used in this process.

In general, language processing technologies are widely used not only for punctuation, but also for detecting other grammatical errors, understanding the meaning of words and automatic translation. Therefore, the combination of software and language processing significantly facilitates communication between humans and computers, making it possible to analyze texts more efficiently and accurately.

Algorithms for automatic analysis of punctuation rules are based on the following basic principles:

1. Rule-based Approach



This approach involves checking punctuation rules against a predefined set of rules. The punctuation rules for each language are clearly defined and the analysis process is carried out based on these rules. Such algorithms control:

- Correct use of commas (use with conjunctions, conjunctions).

- Correct use of complex punctuation such as semicolons or colons.

- The use of a full stop at the end of a sentence and its use with abbreviations.

- Correct placement of parentheses and quotes.

The advantage of this approach is that it is precise and works on the basis of defined punctuation rules. However, a rule-based approach may have difficulty covering many exceptions to the rule.

## 2. Syntactic analysis (Syntactic Parsing)

Syntactic analysis allows a deep analysis of the grammatical structure of the text. Punctuation marks are used based on the grammatical connections between words and sentences in the text. Algorithms identify sentence components and analyze their grammatical relationship. During this process, the places where punctuation rules are violated are identified:

- Comma requirements between connecting and separating parts.

- Appropriate agreement between possessor and participle.

- Punctuation between first and second sentences.

Syntactic analysis is effective in determining rules based on the grammatical structure of the text, and performs correct punctuation analysis even in very complex sentences.

## 3. Machine Learning Approaches

In machine learning technologies, algorithms are trained on large amounts of text to detect punctuation errors. In this approach, models learn sample punctuation usage from previously trained datasets and try to find errors in new texts. The following technologies are used for this:

**Supervised learning:** In this method, a set of incorrectly and correctly punctuated texts is entered into the system. The algorithm analyzes punctuation rules based on this learned information.

**Recursive Neural Networks (RNNs) and Transform Models:** These methods are very effective in analyzing connections in sequences. Neural networks learn language structures and help determine punctuation rules through semantic connections.

Machine learning technologies are more flexible than rigid rule-based algorithms, able to account for text complexity and different contexts.



#### 4. Statistical methods

Statistical methods are used to analyze large data sets and determine the relationships between the use of punctuation marks and text structures. Algorithms calculate the probability of placement of punctuation marks and determine how they should be used in different parts of the text. These methods are implemented using large volumes of written texts and corpora.

#### 5. Hybrid Approaches

Automatic analysis of punctuation rules often uses a combination of several methods. For example, rule-based approaches first examine simple text punctuation rules, and then use machine learning algorithms for more complex cases. These hybrid approaches increase accuracy and flexibility.

In conclusion, we can say that algorithms for automatic analysis of punctuation rules include different approaches. Although rule-based systems are highly accurate, they suffer from flexibility and exception handling difficulties. And machine learning is effective in analyzing complex texts taking into account different contexts. Technologies for analyzing punctuation errors play an important role in improving text quality and creating automatic editing tools.

#### **Used literature**

1. Jurafsky, D., & Martin, J. H. (2019). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Nashriyot: Pearson. ISBN: 978-0131873216
2. Manning, C. D., & Schütze, H. (1999). Foundations of Statistical Natural Language Processing. Nashriyot: MIT Press. ISBN: 978-0262133609
3. Goldberg, Y. (2017). Neural Network Methods for Natural Language Processing. Nashriyot: Morgan & Claypool Publishers. ISBN: 978-1627052986
4. Charniak, E. (1993). Statistical Language Learning. Nashriyot: MIT Press. ISBN: 978-0262033039
5. Briscoe, T. (1996). Automatic Extraction of Grammar from Corpora. Nashriyot: MIT Press. ISBN: 978-0262527330
6. Lison, P., & Tiedemann, J. (2016). OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles. Nashriyot: LREC Proceedings.



7. Andrej Karpathy (2015). The Unreasonable Effectiveness of Recurrent Neural Networks. Onlayn manba: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
8. Genzel, D., & Charniak, E. (2002). Entropy Rate Constancy in Text. Nashriyot: Association for Computational Linguistics (ACL)
9. Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. Nashriyot: O'Reilly Media. ISBN: 978-0596516499