

ANNOTATION

Theme of the Graduation Thesis: "Designing an automatic text attribution system"

Author: Dzhafarov Islam-Butai Timurovic

Thesis tutor: Candidate of Economic Sciences, Associate Professor Timchenko O.V.

Information about the contracting authority: Individual entrepreneur Izudinova P.M.

Relevance of the research topic: Attribution of the literary text is one of the most intensively developing directions in linguistics, which was formed long ago, uses the methods of various fields of science and continues to develop intensively. The increase in the volume of information presented in the form of textual arrays in natural language opens new avenues in the study of literature and language. Huge opportunities appear in the implementation of statistical methods with the support of computer technology in solving many theoretical problems of linguistics and practical problems of processing text data.

Objective: develop an algorithm and software implementation of the intellectual system of attribution of the text, as well as evaluate the effectiveness of the functioning of this system.

Tasks:

1. To study the literature on the design of intelligent systems;
2. Based on the theoretical materials studied, make a choice of the architecture of the system, the training method, other parameters;
3. To study the methods of attribution of texts and the possibility of using intelligent systems in attribution tasks;
4. Conduct experiments on training the system and determine the effectiveness of its operation.

Theoretical and practical significance of the research: consists in approbation of the proposed model of the intellectual system of pattern recognition

within the framework of the final qualifying work and its application in text attribution tasks.

Results of the research: The intellectual system of automatic determination of authorship of the text is developed, realizing the algorithm based on Markov chains with preliminary "clearing" of the text. The system has been trained. The effectiveness of the system is estimated.

Recommendations: It is necessary to continue training the system, expanding the base of authors and texts. To increase the effectiveness of determining authorship, consider the preliminary cleaning of the text from non-informative words.