Multispecies Ensemble Architecture For Texts Authorship Classification

Trukhanovich I. [1]

Paramonov A. [1]

2023

[1]Belarusian State University of Informatics and Radioelectronics, 6 P. Brovki Street, Minsk 220013 Belarus

Abstract: The problem of text authorship identification is considered. This issue is now one of the most demanded tasks in the field of natural language processing. A review of popular authorship analysis techniques is conducted. The implementation of a hybrid classifier built on the ensemble approach is suggested in the paper. By balancing a variety of methodologies and choosing the best outcomes from them, this is anticipated to produce more accurate authorship identification quality measures. The development of an ensemble approach with a quantum-inspired component is offered as a classifier architecture. The results of computer experiments are presented, which confirmed the hypotheses about the effectiveness of the ensemble approach to the task of text classification. The proposed classifier can be effectively applied to identify borrowing cases. For example, in an educational environment, to better detect student plagiarism.