

Wojciech Włodarczyk, MSc

Models for evaluating the correctness of linguistic data obtained using crowdsourcing

Summary

Crowdsourcing uses the collective intelligence of a large group of people to solve tasks in fields such as artificial intelligence, machine learning and scientific research development through online platforms. Nowadays, research on the development of the crowdsourcing method is focused primarily in three areas: optimization of the quality of acquired data, optimization of the cost of the process and optimization of the duration of the process.

This dissertation focuses on issues related to optimizing the quality of the crowdsourcing process for tasks related to linguistic data. The work describes the author's Dynamic Feedback Model, whose task is to generate feedback automatically. The dissertation verifies the effectiveness of this model for empirical and simulation data.

The analysis of the results of the experiment shows the effectiveness of the author's model in improving the quality of the generated feedback, but the quality of this solution depends on the quality of the markings created by annotators.

Keywords

crowdsourcing, human-based computation, machine teaching, natural language processing

Wojciech Włodarczyk