

A Framework for Automated Worker Evaluation Based on Free-Text Responses with No Ground Truth

Tomer Geva¹, Anat Goldstein², Inbal Yahav¹

¹Tel Aviv University, ²Ariel University

ABSTRACT

Evaluating workers according to their work quality is an important managerial task. However, such evaluation becomes highly challenging when no ground-truth information is available against which to compare workers' output. Previous work has addressed this problem in settings in which workers produce binary, numerical or multi-categorical labels. Here, we consider the problem of automatically evaluating workers on the basis of their free-text responses to open-ended questions, without ground truth. To address this problem, we propose a new, unsupervised framework for automated worker evaluation. The framework is based on two main ideas: (a) framing the problem of textual response-based worker evaluation as a multidimensional voting problem; and (b) using an iterative reweighting algorithm that benefits from a holistic assessment of workers' inherent capabilities. To evaluate the framework, we empirically test its performance in two separate studies: using a semi-synthetic-data-based evaluation; and using two datasets of real workers' textual responses. In a third study we use a pure numerical simulation to explore the method's operating conditions. Overall, we find that, across multiple settings, the framework consistently obtains superior results compared with a baseline approach, and that its performance is robust in various challenging conditions. Thus, our framework can serve as a useful benchmark for future research on this problem. Additional benefits of our framework include scalability, modularity, and compatibility with existing, advanced, textual representations.

Keywords: Worker Evaluation, Data Science, Machine Learning, Textual Representation.