QUALITY OF CROWDSOURCED RELEVANCE JUDGMENTS IN ASSOCIATION WITH LOGICAL REASONING ABILITY

Abstract

Crowdsourcing has become a low-cost and scalable alternative to gather relevance assessments through crowdsourcing platforms. However, when gathering subjective human judgments, it can be challenging to enforce data quality due to the lack of vision about how judges make a decision. It is important to determine the attributes that could affect the effectiveness of crowdsourced-judgments in an information retrieval systems evaluation. The purpose of the experiment that is discussed in this paper is to investigate if logical reasoning ability of the crowd workers is related to the quality of the relevant judgments produced through the crowdsource process. The study also evaluates the effect of cognitive characteristics on the quality of relevance judgment compared to the gold standard dataset. Through this experiment, a comparison study is done between the quality of the judgments obtained through the crowdsourcing process and the original baseline judgments generated by the hired experts by TREC. In the study, the systems performances were measured using both of these sets of relevance judgments to see its correlation. The experimentation reveals that quality of relevance judgments is highly correlated with the logical reasoning ability of individuals. The judgment difficulty level reported by the crowdsourcworkers and the confidence level claimed by the workers showed a significant correlation with the quality of...
the judgments. Unexpectedly though, self-reported knowledge about a given topic and demographics data have no correlation with the quality of judgments produced through crowdsourcing.