Knowledge Lab

Birkbeck UNIVERSITY OF LONDON

Research Student
Igor Volzhanin

Supervisors
Dell Zhang
Ulrike Hahn

Birkbeck UNIVERSITY OF LONDON

Department of Computer Science and Information Systems

Similarity Measures and Truth

Although defining optimal measures of similarity have been a research question for a long time, dating back to information retrieval, less research focused on the interplay between accuracy of the different measures and their computational intensity.
While some methods are computationally intensive, with Kemeny-Snell being an NP hard problem with list size of just 4 (Dwork, 2001), other methods such as the mean, median, or majority lists are much easier to compute (Xie, 2015).

Research Methodology

By creating a simulation environment that allows for different models to be implemented and varied accordingly, we will be able to compare the different aggregation methods and evaluate their effectiveness in both getting to the 'truth', as well as their computational costs.

Research Approach

The initial simulation environment has been programmed in R and is capable of drawing lists of various sizes and comparing different aggregation methods to determine accuracy. The model accepts different parameters: aggregation method, total runs, simulation runs, list size, individual competence – the likelihood of picking the correct list among the alternatives, draw method – parameter determining the error distribution
The output of the simulation is a graph that displays various results measuring accuracy as a function of number of judges, but can be altered to display any part of the simulation.
At the moment four aggregation models have been implemented in the simulation: averaging, Kemeny Snell distance measure, Majority and Spearman's footrule.

Preliminary Results

Initial simulations largely focused on manipulating the individual competence measure and number of judges to determine which aggregation models perform best under the different conditions. We have found a curious switch between accuracy majority and averaging rules as a function of competence. We have also found these **computationally simpler rules outperforming Kemeny Snell and Spearman's footrule in nearly all situations.**

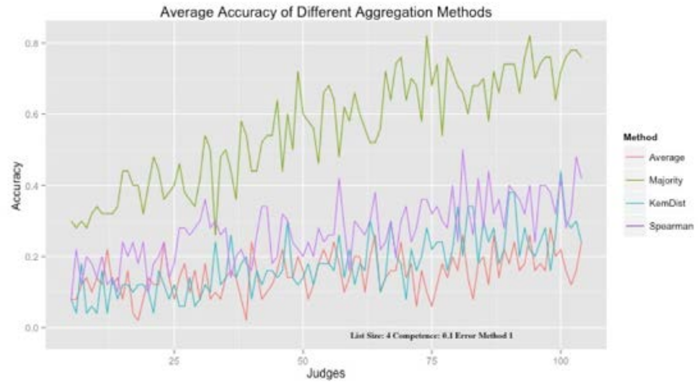

Figure 1. Majority outperforming under high competence

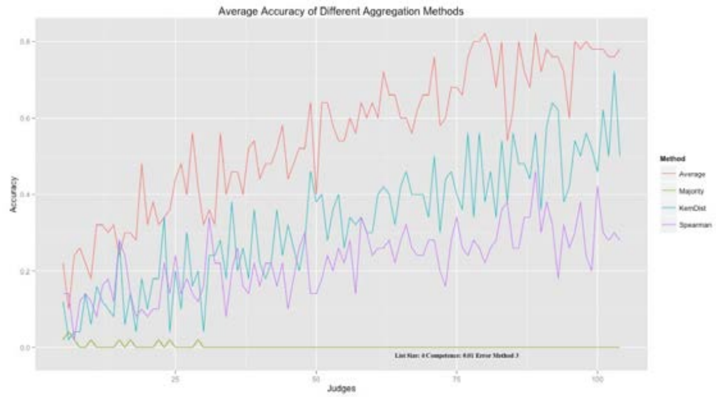

Figure 2. Averaging outperforming under lower competence

Publications

Brancotte, B., Yang, B., Blin, G., & Cohen-boulakia, S. (2015). Rank aggregation with ties : Experiments and Analysis Rank Aggregation problem. *Proceedings of the 41st International Conference on Very Large Data Bases*, 1202–1213.
Duddy, C., & Piggins, A. (2012). A measure of distance between judgment sets. *Social Choice and Welfare*, *39*(4), 855–867.
Dwork, C., Kumar, R., Naor, M., & Sivakumar, D. (2001). Rank aggregation methods for the Web. *Proceedings of the 10th International Conference on World Wide Web*, 613–622.
Vigna, S. (2014). A Weighted Correlation Index for Rankings with Ties, (Ga 288956), 1–19.