The Self-Organizing Map for the Analysis of Survey Data

M. Litova[1] and K. Lagus[2]

[1] University of Helsinki, Finland
e-mail: maria.litova@helsinki.fi

[2] University of Helsinki, Finland
e-mail: krista.lagus@helsinki.fi

Abstract

The self-organizing map (SOM) algorithm is implemented to analyze the survey data obtained from the faculty well-being project conducted in spring 2022 at the Faculty of Social Sciences in the University of Helsinki. The self-organizing map is an unsupervised neural network and data analysis method that enables dimensionality reduction, exploration of variable variation and dependencies and visualization of similarity relations (Kohonen, 2001). With its strong clustering capabilities and visualization potential, the self-organizing map can be effectively applied to the analysis of survey data, particularly data collected through questionnaires.

The survey conducted as part of the faculty well-being project comprised three question groups. The first group focused on gathering background information, while the second group consisted of closed Likert-scale questions aimed at capturing individuals' subjective experiences of well-being within the faculty. Lastly, the survey included open-ended text questions that explored various topics, including social interaction, the role of the faculty in promoting well-being, maintaining equality, and other related subjects (Laine et al., 2022). Closed background and Likert-scale questions were chosen to be analyzed with the SOM method.

The utilization of the self-organizing map algorithm for the analysis of closed questions facilitated the identification of seven profiles (clusters) among survey participants. These profiles were obtained based on the varying experiences concerning well-being and the accompanying background information, including gender, position within the faculty and proficiency in the Finnish language. The implementation of the SOM method can be described as an experimental undertaking that typically involves several key steps. These steps include the identification of the optimal set of parameters for the SOM training, the selection of an appropriate approach for encoding variables, and the handling of missing values within the dataset.

Keywords: self-organizing map, survey data, clustering, well-being.

References

Kohonen, T. (2001). Self-Organizing Maps. Springer-Verlag; Berlin, Heidelberg.

Laine, K., Litova, M., and Oikarinen T. (2022). Crowdsourcing social wellbeing non-probability survey.
In*: Baltic-Nordic-Ukrainian workshop on survey statistics 2022,* Tartto, 46.