A framework for leveraging multi-rater data in brain decoding analysis: Prediction of evaluation drawn from population data using sparse probit regression

Shirin Vafaei[1], Hiroaki Yamane[2,3], Ryohei Fukuma[1], Noriaki Yahata[4], Takufumi Yanagisawa[1], Kei Majima[4]

[1]Institute for advanced co-creation studies, Osaka University, Osaka, Japan

[2]RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

[3]Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

[4]National Institutes for Quantum and Radiological Science and Technology, Chiba, Japan

Introduction: Using stimuli (*e.g.*, images, videos, products) labeled by a number of raters has recently become common in brain decoding analysis, where subjective emotion/impression for stimuli felt by the population is predicted from brain responses. However, there remains no established method for constructing a decoder using such multi-rater labels. In previous studies, the variability across multiple raters was assumed to reflect noise, and the answers for a binary judgment were averaged across raters. Then, the average scores (*i.e.*, empirical probabilities) for individual stimuli were predicted using standard regression methods. While this procedure is a simple and popular approach, it is not appropriate because most of these regression methods ignore the fact that probability is the variable to be predicted. To address this in an appropriate manner, we present a new framework in this study.

Methods: Here, we assume that individual answers for a binary judgment about a stimulus follow a Bernoulli distribution. We then predicted the probability of positive answers from the human functional magnetic resonance imaging (fMRI) response to the stimulus using probit regression. We also introduced sparse regularization into probit regression (sparse probit regression) to prevent overfitting.

Results: In both simulation and real fMRI data analysis, sparse probit regression more accurately predicted the probabilities of positive answers for individual stimuli than probit regression without sparse regularization, indicating that sparseness results in better decoding performance. Sparse probit regression also outperformed linear regression using the same type of sparse regularization, reflecting the advantage of our appropriate treatment of probability.

Discussion & Conclusion: Our results suggest that our framework using sparse probit regression provides an effective method for the population prediction of emotion/impression assessment based on brain activity.